SIAM J. CONTROL OPTIM.
Vol. 58, No. 6, pp. 3559–3585

SHARING PROFITS IN THE SHARING ECONOMY*

PAOLO GUASONI† AND GU WANG‡

Abstract. A monopolist platform (the principal) shares profits with a population of affiliates (the agents), heterogeneous in skill, by offering them a common nonlinear contract contingent on individual revenue. The principal cannot discriminate across individual skill but knows its distribution and aims at maximizing profits. This paper identifies the optimal contract, its implied profits, and agents' effort as the unique solution to an equation depending on skill distribution and agents' costs of effort. If skill is Pareto-distributed and agents' costs include linear and power components, then closed-form solutions highlight two regimes: If linear costs are low, the principal's share of revenues is insensitive to skill distribution and decreases as agents' costs increase. If linear costs are high, then the principal's share is insensitive to the agents' costs and increases as inequality in skill increases.

Key words. optimal contracts, principal-agent, hidden type, adverse selection

AMS subject classifications. 91B41, 91B43

DOI. 10.1137/19M1254453

1. Introduction.

The sharing economy, based on peer-to-peer markets, has quickly disrupted industries, most evidently transportation and accommodation, by enabling a myriad of individuals to offer these services. It has also created new industries, such as video-sharing, in which producers of popular content are paid in relation to the advertising revenue that they generate.

Yet, each industry has also seen the emergence of a dominant platform that connects users with service providers—and collects a substantial fraction of revenues. As both users and providers have an incentive to use the most popular platform, network effects cement its dominance and confer it significant market power over affiliates. For example, the dominant ride-sharing service Uber has a tiered fee structure whereby it collects between 20 and 30 percent of drivers' fares and years after its launch has raised its proportional fees in some cities with a mature presence.[1]

Absent competitive pressure, such a platform seeks the sharing contract that maximizes its aggregate profits from the revenues generated by a crowd of affiliates with different productivity (henceforth, skill). It is akin to a platform-specific government that, in contrast to a benevolent social planner, devises an income tax schedule merely to maximize aggregate tax receipts. Such near-monopolistic aspects have drawn public concern and legal scrutiny, with the U.S. Supreme Court scheduled to hear a case involving fees on Apple's App Store, which (like Google's Play Store) levies a 30% fee on the revenue generated by developers of mobile applications.[2]

This paper solves the profit-sharing problem for a monopolist platform, identifying the contract that maximizes its aggregate profits and the optimal effort that each

*Received by the editors April 8, 2019; accepted for publication (in revised form) September 16, 2020; published electronically December 1, 2020.
https://doi.org/10.1137/19M1254453
Funding: The work of the first author is partially supported by the ERC (279582), NSF (DMS-1412529), and SFI (16/IA/4443 and 16/SPP/3347).
†School of Mathematical Sciences, Dublin City University, Glasnevin, Dublin 9, Ireland (paolo.guasoni@dcu.ie).
‡Department of Mathematical Sciences, Worcester Polytechnic Institute, Worcester, MA 01609 USA (gwang2@wpi.edu).
[1]See https://www.wsj.com/articles/uber-tests-30-fee-its-highest-yet-1431989126.
[2]See https://www.wsj.com/articles/app-tax-a-hard-one-to-beat-1535198400.

affiliate chooses, in order to maximize income, net of personal costs. At its core, our problem is closest to the Mirrlees [19] model of optimal taxation, with two related differences. First, our principal maximizes profits, rather than some welfare measure that aggregates agents' own utilities. Second, while in Mirrlees's model tax revenues are an exogenous quantity, profits here are maximized and hence endogenous.

The paper's contribution is threefold: First, we provide a rigorous formulation of a monopolist platform's optimization problem and obtain its solution through variational methods. Importantly, we achieve this goal without making strong regularity assumptions on the class of optimal contracts considered, demonstrating instead that regularity is necessary for optimality. Second, we introduce the notion of a *canonical contract*, a contract that is uniquely constructed to induce a prescribed revenue function for each skill, and which outperforms all other contracts that induce the same revenue. Third, we show how to obtain optimal contracts in closed form in relevant settings, and how different combinations of parameter values lead to significantly different regimes, in which either skill dispersion or personal costs are dominant.

Optimal contracts in our setting share some features with optimal taxation models: the optimal contract implies that agents with the lowest skill do not participate; their optimal effort is null. This phenomenon, known as *bunching* [16, 21], arises because the principal gains more from taking a larger share of the basic income of high-skill agents than it loses from forgoing any share of low-skill incomes, thereby inducing nonparticipation. Furthermore, higher-skilled agents always generate more revenue than the lower skilled and receive higher income. Also, marginal sharing rates are always positive for both parties, in that neither the principal nor the agent receive subsidies. Subsidizing agents to stimulate higher effort costs more than it produces because agents' personal costs are convex (effort has diminishing returns). Vice versa, seizing more than marginal revenue is also counterproductive; mere confiscation always generates as much profit for the principal (and possibly more).

Theorem 4.2, the main result, describes quantitatively the optimal policy for agents with different skills, which arises from the principal's trade-off between increasing its share of profits and spurring agents to more effort through better incentives. It also identifies the optimal contract for the principal through the first-order condition which requires that, at the optimum, the increase in revenue is exactly offset by agents' behavioral responses. Optimality is summarized by the optimization problem (4.1), which reduces to the functional equation (4.3) when the cost of effort is sufficiently smooth. Importantly, such an equation admits closed-form solutions under specific assumptions on skill distribution and cost, which we examine in detail.

We bring to life the results in a tractable setting, focusing on a population of agents with Pareto-distributed skill (i.e., the number of agents with skill greater than or equal to t is proportional to $t^{-\alpha}$), and on linear-power costs $f(y) = by + ay^\lambda$, characterizing explicitly the optimal effort, its income, and the corresponding optimal contract. With power costs alone, the optimal contract is exactly affine. With linear-power costs, the principal's share declines as revenue increases, reaching an asymptotic share equal to the share implied by power costs alone.

The relative magnitude of linear costs is a key driver of the principal's share of revenues under the optimal contract, which displays two rather different regimes (see Figure 1). In the *low costs* regime—when the linear coefficient b in agents' costs is relatively low—the principal's share is insensitive to the distribution of skill but very sensitive to the growth rate λ of such costs. In the case of $b = 0$, the share is exactly $1 - 1/\lambda$.

Vice versa, in the *high costs* regime—when b is relatively high—the principal's

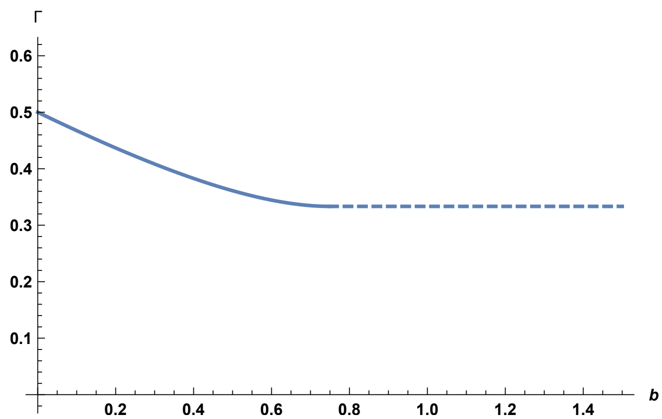

FIG. 1. *Proportion of total revenue (Γ) collected by the principal, against the linear component (b) of agents' costs ($f(y) = by + ay^\lambda$), when agents' skill is Pareto distributed ($M(t) = (t/\underline{t})^{-\alpha}$). As the linear component increases from zero to $\alpha\underline{t}/(1+\alpha)$, the principal's share shifts from $1 - 1/\lambda$ (low-cost regime) to $1/\alpha$ (high-cost regime), where it remains after further increases. Parameters: $a = 1$, $\underline{t} = 1$, $\alpha = 3$, and $\lambda = 2$. See section 5 for details.*

share is insensitive to the growth rate λ but is driven by the distribution of skill, summarized by the Pareto exponent α. Indeed, if b becomes greater than or equal to a fixed threshold, the principal's share is exactly $1/\alpha$, which means that skill inequality increases the principal's overall share of revenues. The intuition is that the bulk of the principal's profits stems from high-skill agents, who become increasingly prevalent as skill inequality increases.

In general, heterogeneity in skill has an ambiguous effect on individual agents. On one hand, it may benefit an individual agent by shifting the common contract away from the one that would be optimal for the principal if all agents had the same skill, potentially leaving a greater share of income to that agent. On the other hand, agents at the bottom of the skill distribution find themselves essentially ignored by the principal, who chooses a contract that induces them not to participate. Such agents clearly suffer from the presence of the higher skilled, without whom the principal would choose to cater to their own ability.

Overall, the above considerations point to two main environments in which a monopolist platform has high market power: either relatively homogeneous affiliates with high linear costs, or affiliates with costs that are negligible for small effort but that increase quickly as effort increases. An example of the former may be drivers, whose fuel and maintenance costs are relatively uniform. Vice versa, the latter environment may encompass those activities that are akin to hobbies when performed sporadically, such as hosting a guest in a spare room on a monthly basis, but that require a substantial commitment of resources when conducted systematically, such as managing several apartments for short-term rentals.

Even though the problem of optimal contract design for a monopolistic platform with a heterogeneous population of agents does not appear to have been considered in the literature,[3] it falls squarely within the class of problems with adverse selection stemming from hidden type [2, 7, 1, 5], which is a mainstay of contract theory (see,

[3]In recent continuous-time models, El Euch et al. [9] find the optimal contract for a trading platform as principal and one market maker as agent, while Elie, Mastrolia, and Possamaï [10] tackle the problem of a principal and infinitely many identical agents.

e.g., [3, 6]). Indeed, adverse selection is the main aspect considered in the literature on optimal taxation (cf. [19, 8, 18, 11]). Similar to this literature, in our model agents' actions are perfectly observable (i.e., moral hazard is absent), which means that the principal trades off the increased profits of higher fees against their efficiency costs of reduced effort. In this setting, we strive to keep the class of admissible contracts as general as possible, so as to make sharp statements on the qualitative properties implied by optimality. Starting from any upper-semicontinuous positive contract, we proceed to derive further properties, including regularity and a priori bounds, by showing that their absence leads to potential improvements.

Finally, we investigate the impact of skill-dependent reservation utility, focusing for simplicity on a high-low regime in which such utility is higher for agents with skill above a certain threshold. The methodology developed in the paper extends to this setting, and we show how the canonical contract requires a careful global construction in the skill space, which entails several cases. In particular, this setting highlights the emergence of *countervailing effects*, whereby the optimal contract may entail two cohorts of nonparticipating agents: as in the main model, the first cohort includes agents with very low skill, for whom participation is surpassed by their low reservation utility. The second cohort includes agents with skill barely above the threshold for the high reservation utility: such agents also shun the optimal contract, as it does not offer a sufficient surplus to compensate for the effort. The principal optimally forgoes the participation of these agents because it would require such a high compensation for them and for those with similar ability that the costs of such increased compensation would more than offset the benefits from their participation.

The rest of the paper is organized as follows: Section 2 introduces the model in detail and describes its assumptions. Section 3 derives a priori properties of optimal contracts, including regularity, monotonicity, and positivity, thereby enabling the use of variational tools to characterize optimality. Section 4 contains the main result, which identifies the optimal contract in terms of the agents' distribution and cost function. Section 5 examines the implications of the main result in the tractable setting of Pareto skill and linear-power costs. Section 6 discusses the extension to high-low reservation utilities, and section 7 concludes. All proofs are in Appendix A.

2. Model. One principal faces a population of agents who differ in productivity, or skill. The skill of each agent remains hidden to the principal, who only knows its distribution, described by a density.

Assumption 2.1. The density $m : [0,\infty) \mapsto [0,\infty)$ is measurable and satisfies $\int_0^\infty m(t)dt < \infty$ (i.e., the population is finite). Let $M(t) = \int_t^\infty m(u)du$.

Following the convention of the optimal taxation literature since [19], we define skill as agent's productivity per unit of effort. Thus, for an agent with skill t, $y \in [0,\infty)$ units of effort produce revenue of yt, split as *net income* $c(ty)$ for the agent and *profit* $ty - c(ty)$ for the principal. The contract c satisfies the following assumption.

Assumption 2.2. The contract $c : [0,\infty) \mapsto \mathbb{R}$ is upper-semicontinuous and nonnegative. C denotes the set of such contracts.

A nonnegative contract is a basic participation (or incentive-compatibility) constraint, which implies that the agent has the alternative of nonparticipation, corresponding to a zero payoff (i.e., income cannot be negative, which excludes the confiscation of other pre-existing wealth).

Upper-semicontinuity means that income cannot increase by a finite amount as effort varies by an infinitesimal amount, effectively excluding penalties that can be avoided with negligible changes in effort, and hence would be irrelevant in practice. (For example, suppose that the principal's share on an income of exactly 1000 is 20%, but it falls to 15% for any income in $[999, 1000) \cup (1000, 1001]$; then any agent who would choose 1000 if the share were at 15% prefers to increase (or overreport) income by by some ε (say, 0.001), thereby securing the 15% rate. Of course, increasing by $\varepsilon/2$ is even better, which leads to the nonexistence of an optimal effort.)

Aside from these well-posedness restrictions, the above setting includes contracts that vary arbitrarily, leaving marginal sharing rates the flexibility to be progressive, regressive, negative, above or below 100%, or even to have jumps, provided that at each jump the income does not drop below its left and right limits.

An agent with skill t who chooses effort y receives the reward

$$c(ty) - f(t, y), \tag{2.1}$$

resulting from the net income $c(ty)$ minus the personal cost $f(t, y)$, which satisfies the following assumption.

Assumption 2.3. The cost $f : [0, \infty) \times [0, \infty) \mapsto [0, \infty)$ satisfies the following:
(i) f is strictly increasing and convex in y, and $\lim_{u\to\infty} f_y(t, u) = \infty$ for every t.
(ii) f is continuously differentiable in y and t.
(iii) $f(t, 0)$ and f_y are nonincreasing in t.
Denote by $\mathcal{F}$ the set of such cost functions.

A convex increasing f (as a function of y) reflects the increasing marginal personal costs of effort, and continuous differentiability excludes that such marginal cost has jumps. The third condition (which is akin to assumption (CS) in [13]) prescribes that, at the same level of effort, higher skill is associated to both a higher baseline (zero effort) utility and equal or lower marginal costs. The intuition is that a more skilled individual is more likely to have more tangible or intangible assets that alleviate the hardships of increased effort (e.g., lower commuting times, or more flexible hours). Assumptions 2.1, 2.2, and 2.3 apply to the rest of the paper, without further notice.

For each contract $c \in C$ and personal cost $f \in \mathcal{F}$, Lemma A.1 in Appendix A shows that for any $t \geq 0$ there exists some $y^c(t) \in [0, \infty]$ that maximizes (2.1). Then the resulting total profit under the contract c is

$$P(c) = \int_0^\infty (ty^c(t) - c(ty^c(t)))m(t)dt, \tag{2.2}$$

and the principal's objective is to choose the contract $c \in C$ that maximizes profits.

Prima facie, one is tempted to solve this optimization problem by focusing on the first-order condition on revenue $s = ty$ for individuals with skill t maximizing the objective (2.1), which is

$$c'(s) = \frac{1}{t} f_y\left(t, \frac{s}{t}\right). \tag{2.3}$$

The problem with such a condition is that it is moot at the points where c is not differentiable, it is ambiguous at the points where the revenue s is discontinuous, and neither case can be excluded. As discussed in section 4, even a quadratic cost function can lead to discontinuity in the optimal marginal sharing rate c' and the corresponding revenue s, depending on the distribution of skill (see Remark 4.4 after Theorem 4.2). As a more careful approach is necessary, the next section starts investigating the a priori properties that optimal contracts must satisfy.

3. Necessary conditions and canonical contracts. This section proves that, notwithstanding the flexibility in the choice of contracts, the ones that maximize profits must be continuously differentiable and entail positive sharing rates for both the principal and the agents. These a priori restrictions simplify the search for the optimal contract by narrowing the focus on schedules with these properties. First, only positive sharing rates are optimal. Second, it is possible to find a *canonical contract* that is at least as profitable as any other contract that induces the same revenue. As such a contract is differentiable by construction, it follows that it is sufficient to consider such contracts. The next proposition reduces the search for optimal contracts to those that are right-continuous and free from subsidies.

PROPOSITION 3.1. *Define the set of contracts*

$$\tilde{C} = \{c \in C : \textit{nondecreasing, right-continuous, } 0 \le c(x) \le x \textit{ for all } x \ge 0\}. \tag{3.1}$$

(i) $\sup_{c\in C} P(c) = \sup_{c\in\tilde{C}} P(c)$*; i.e., in order to maximize profits, it suffices to consider contracts in* $\tilde{C}$*.*
(ii) *Given* $c \in \tilde{C}$*,* $y^c(t)$ *is finite for every* t*, and* $\lim_{t\to 0} y^c(0) = 0$*. Furthermore, if* $0 < t_1 < t_2$*, then* $t_1 y^c(t_1) \le t_2 y^c(t_2)$*.*

Proposition 3.1 shows that any revenue profile $s(t) = ty^c(t)$, generated by a contract $c \in \tilde{C}$, henceforth referred to as "admissible" revenue, belongs to the set

$$S = \big\{s \ge 0 : s(0) = 0, \lim_{t\to 0} s(t)/t = 0, s \text{ is nondecreasing}\big\}. \tag{3.2}$$

The next step is to understand the relation between the contract $c \in \tilde{C}$ and its respective revenue $s(t)$ across skill levels. The following definition constructs, for a given admissible revenue $s \in S$, a "canonical" contract c^s, with the additional property (Proposition 3.3 below) that it maximizes the profits among those contracts c that induce the same revenue. As a result, the optimization problem can focus only on such contracts.

DEFINITION 3.2. *For* $s \in S$*, define the function* $c^s : \mathbb{R}_+ \to \mathbb{R}_+$ *as* $c^s(x) = \int_0^x c_x^s(z)dz$*, where* $c_x^s(z) = \frac{1}{t_z} f_y(t_z, \frac{z}{t_z})$*, and* $t_z = \inf\{t : s(t) > z\}$*, with the convention that if* $\{t : s(t) > z\} = \emptyset$*, then* $t_z = \infty$*.*

This definition constructs a contract c^s that satisfies a version of the above first-order condition, in which t is the right-continuous inverse of the increasing function $s(t)$. Note that, because $s(t)$ converges to 0 as t decreases, for every $z > 0$, $t_z > 0$, and $c_x^s(z)$ is well defined, except for $z = 0$. Also by definition, t_z is increasing in z, and is discontinuous (thus so is c_x^s) only at z, such that $s(t) = z$ for every $t \in (t_1, t_2)$, where $t_1 = \inf\{t : s(t) = z\}$ and $t_2 = \sup\{t : s(t) = z\}$. In the latter case, $c_x^s(z) = \frac{1}{t_2} f_y(t_2, \frac{z}{t_2})$. Thus, c^s is differentiable and increasing on $(0, \sup_{t\ge 0} s(t))$, and c_x^s is right-continuous. At $z = s(t)$ for some t, it satisfies the usual optimality condition that the marginal net income matches the marginal personal cost of the highest skilled agents who produce z. At those revenue points z where a marginal change in skill may entail a large change in income, the condition prescribes a marginal net income equal to marginal cost for the skill t_z that would be just enough to exceed such revenue.

The logic of the definition is simple: at any revenue discontinuity point the marginal profit rate should be as low as required to make the agent's income as high as consistent with the first-order condition. The next result shows that the contract c^s is optimal among all the contracts that induce the same revenue $t \mapsto s(t)$.

PROPOSITION 3.3. *The contract* c^s *defined in Definition* 3.2

(i) *induces the revenue* $s(t)$ *(i.e.,* $s(t) = ty^{c^s}(t)$*) up to countably many* t*, and*
(ii) *satisfies* $P(c^s) \geq P(c)$ *for any contract* $c \in \tilde{C}$ *that satisfies* (i).

Proposition 3.3 implies that the principal only needs to consider canonical contracts c^s corresponding to admissible revenue functions s. In terms of such a revenue function s, the principal's objective becomes

$$\max_{s \in S} P(c^s) = \max_{s \in S} \int_0^\infty (s(t) - c^s(s(t)))m(t)dt. \tag{3.3}$$

Notice that the set of revenues generated by contracts in $\tilde{C}$ may be a strict subset of S, and thus not every c^s belongs to $\tilde{C}$; for example, c^s_x may not be integrable. But it would not affect the maximum profits to consider all c^s's, because those that are not in $\tilde{C}$ are still in C and are certainly suboptimal.

The next theorem strengthens Proposition 3.1 by showing not only that any optimal contract is lower than revenue but also that its marginal increase is also less than the marginal increase in revenue, at the levels actually being produced.

THEOREM 3.4. *If the contract* $c^{\hat{s}}$ *that corresponds to* $\hat{s} \in S$ *is optimal, then* $c^{\hat{s}}_x(\hat{s}(t)) \leq 1$ *for every* t *such that* $M(t) > 0$*.*

4. Optimal contracts.

This section contains the main result, which identifies the optimal contract through its implied revenue $\hat{s}(t)$. In addition to the standing Assumptions 2.1, 2.2, and 2.3, this result relies on the following assumption.

Assumption 4.1. Let f be independent of t (and denote it as $f(y)$), $yf_y(y)$ be convex, and $tm(t)/M(t)$ be increasing (recall that $M(t) = \int_t^\infty m(z)dz$).

The convexity of $yf_y(y)$ is a mild requirement that is satisfied by all examples of interest.[4] Likewise, the condition of an increasing elasticity $tm(t)/M(t)$ stipulates that the density of skill does not decline too slowly; it is satisfied by power laws (constant elasticity), by exponential or Gamma tails (affine elasticity), and by Gaussian tails (quadratic elasticity).[5]

THEOREM 4.2. *Let Assumptions* 2.1, 2.2, 2.3, *and* 4.1 *hold, and define*

$$g(t, s) = (s - f(s/t))m(t) - M(t)\frac{s}{t^2}f_y(s/t), \qquad t, s \geq 0. \tag{4.1}$$

(i) *The revenue* $\hat{s}(t)$ *defined as*

$$\hat{s}(t) = \begin{cases} 0, & t = 0, \\ \inf\{s \geq 0 : s \text{ maximizes } g(t, \cdot)\}, & t > 0, \end{cases} \tag{4.2}$$

is finite for every $t > 0$ *and solves* (3.3).

(ii) *If* f *is twice differentiable, then for* $t > 0$*,* $\hat{s}(t) = 0$ *if* $g_s(t, 0) \leq 0$*. Otherwise,* $\hat{s}(t) = \inf\{s : g_s(t, s) = 0\}$*, where*

$$g_s(t, s) = \left(1 - \frac{1}{t}f_y\left(\frac{s}{t}\right)\right)m(t) - \left(\frac{s}{t}f_{yy}\left(\frac{s}{t}\right) + f_y\left(\frac{s}{t}\right)\right)\frac{M(t)}{t^2}. \tag{4.3}$$

[4]The condition is mild but not redundant. For a counterexample, consider $f(x)$ defined as 1 on $[0, 1]$, as $1/x + \log x$ on $[1, 2]$, and as $1/2 + \log 2 + (x - 2)/4 + (x - 2)^2$ on $[2, \infty)$, which is increasing and convex, while $xf'(x)$ is concave on $[1, 2]$.

[5]For a distribution so leptokurtic that it violates this condition, consider $M(t) = 1/\log t$ for $t \geq 1$, which implies that the elasticity is also $1/\log t$ and hence decreasing. Such a distribution does not have finite moments of any (positive) order.

Remark 4.3. Theorem 4.2 resonates with the conclusion in Theorem 3.4, in that condition (i) implies that if $t < \bar{t} = \sup\{t \geq 0 : M(t) > 0\}$ and $\hat{s}(t) > 0$, then $\hat{s}(t)$ maximizes $g(t,s)/m(t) = (s - f(\frac{s}{t})) - \frac{M(t)}{m(t)t}\frac{s}{t}f_y(\frac{s}{t})$ as a function of s. Since $\frac{M(t)}{m(t)t}\frac{s}{t}f_y(\frac{s}{t})$ is increasing in s, a necessary condition for $\hat{s}(t)$ to be the maximizer is that $s - f(\frac{s}{t})$ is increasing at $\hat{s}(t)$, which implies that $1 \geq \frac{1}{t}f_y(\frac{\hat{s}(t)}{t}) = c_x^{\hat{s}}(\hat{s}(t))$.

Remark 4.4. A direct consequence of Theorem 4.2 is that the optimal revenue profile $\hat{s}$ may not be continuous, as now shown. Consider, for example, the quadratic cost function $f(y) = \frac{1}{2}ay^2 + by$: the solution to the equation $0 = g_s(t,s)$ in (4.3) is $\hat{s}(t) = \frac{m(t)t^2 - bm(t)t - bM(t)}{2aM(t)/t + am(t)}$ (for sufficiently large t). Thus, $\hat{s}$ inherits the regularity properties (or lack thereof) of the density $m(t)$. In particular, if $m(t)$ is discontinuous, then so is $\hat{s}(t)$. (Note that Assumption 4.1 only requires $tm(t)/M(t)$ to be increasing—not necessarily continuous.)

The objective $g(s,t)$ in (4.1) embodies the central trade-off that the principal faces. To better understand it, consider the first term in (4.1), which, aggregated over the population of agents, is $\int_0^\infty (s(t) - f(\frac{s(t)}{t}))m(t)dt = \int_0^\infty (s(t) - c(s(t))\, m(t)dt + \int_0^\infty (c(s(t)) - f(\frac{s(t)}{t}))m(t)dt$. As the two terms on the right-hand side represent, respectively, the aggregate principal's profits and the aggregate agents' utilities, their sum is akin to the objective of a hypothetical social planner that seeks to maximize the total revenue net of agents' personal costs. Such a hypothetical objective would be realized by the "altruistic" contract $c(s) = s$, which is not merely suboptimal for the principal; it is the worst possible contract, with a net profit of zero.

Thus, the second term in (4.1) represents the extent to which the principal's interests depart from those of a social planner disinterested in profits. In view of the first-order condition in (2.3), this term is (again, aggregated over all agents) $\int_0^\infty M(t)\frac{s(t)}{t^2}f_y(\frac{s(t)}{t})dt = \int_0^\infty M(t)\frac{s(t)}{t}c_x^s(s(t))dt$. Here, $c_x^s(s(t))s(t)/t$ represents the marginal cost of a change in compensation paid to agents of skill at least t on the revenues that they generate, with the factor $1/t$ reflecting that such a change is inversely proportional to skill, for the same revenue. Because a change in compensation at the revenue level $s(t)$ affects all agents with skill greater than or equal to t, its impact is proportional to $M(t)$ rather than $m(t)$.

The next section explores the implications of this result in a concrete setting, where the optimal contracts and their implied sharing rates are found explicitly.

5. Application: Pareto skill with linear-power costs. This section brings to life the main results with examples in which the assumptions in Theorem 4.2 hold and the optimal contract admits a closed-form solution. Throughout this section, agents' skill is assumed to be above the lower bound $\underline{t}$ and to follow the Pareto distribution $m(t) = \alpha\underline{t}^\alpha t^{-\alpha-1}$ with $\alpha > 1$, so that $M(t) = (t/\underline{t})^{-\alpha}$ (see, e.g., [19]). Agents' costs are of linear-power type, i.e., $f(y) = ay^\lambda + by$ for positive constants a, b, and $\lambda > 1$, such that $\lambda/(\lambda - 1) < \alpha$.

Note that such fat-tailed distributions arise in empirical work on income distribution [20, 8] and video-sharing [4]. More broadly, they tend to arise in the provision of club goods (nonrival but excludable), such as movies, music, e-books, and other content that can be reproduced at negligible costs.

With explicit results for the optimal contract, its corresponding revenue, and profits, we show that power costs (without a linear component) lead to an affine optimal contract, in which principal and agent share each additional unit of revenue according to fixed proportions that depend on the agents' distribution and their costs.

Of all revenues generated under the optimal contract, the principal collects $1-1/\lambda$ as profits, while all agents share the remaining $1/\lambda$. Then, we add the linear component, which leads the principal to set a fee schedule that declines to an asymptotic rate. When the linear component is large enough, such an optimal schedule implies that the principal takes a constant share of the total revenue, and that such a share is the inverse of the Pareto exponent α.

The analysis begins by noting that, in view of (4.3), the optimal revenue $\hat{s}(t)$ for skill $t \geq \underline{t}$ has the explicit solution

$$\hat{s}(t) = \begin{cases} t(\frac{\alpha t-(\alpha+1)b}{a\lambda(\alpha+\lambda)})^{\frac{1}{\lambda-1}}, & t > \frac{(1+\alpha)b}{\alpha} \vee \underline{t}, \\ 0, & t \leq \frac{(1+\alpha)b}{\alpha} \vee \underline{t}. \end{cases}$$

5.1. Power costs and affine optimal contracts. Without a linear component in the cost function ($b=0$), $\hat{s}(t) = (\frac{\alpha}{a\lambda(\alpha+\lambda)})^{\frac{1}{\lambda-1}} t^{\frac{\lambda}{\lambda-1}} = (t/\underline{t})^{\frac{\lambda}{\lambda-1}}\, \hat{s}(\underline{t})$ for $t \geq \underline{t}$. From Definition 3.2, the associated optimal contract $c^{\hat{s}}$ satisfies the condition $c_x^{\hat{s}}(\hat{s}(t)) = \frac{1}{t} f_y(\frac{\hat{s}(t)}{t}) = \frac{\alpha}{\alpha+\lambda}$, and thus, the optimal contract is affine. For agents with skill $t > \underline{t}$, in addition to the flat fee of $\hat{s}(\underline{t}) - c^{\hat{s}}(\hat{s}(\underline{t}))$, every dollar of revenue above $\hat{s}(\underline{t})$ is shared between the principal and the agent with the proportions of $\frac{\lambda}{\alpha+\lambda}$ and $\frac{\alpha}{\alpha+\lambda}$, respectively.

From (A.8) (which holds for any skill distribution),

$$c^{\hat{s}}(\hat{s}(t)) = \lambda a^{-\frac{1}{\lambda-1}} \left(\frac{\alpha}{\lambda(\alpha+\lambda)}\right)^{\frac{\lambda}{\lambda-1}} t^{\frac{\lambda}{\lambda-1}} - (\lambda-1) a^{-\frac{1}{\lambda-1}} \left(\frac{\alpha}{\lambda(\alpha+\lambda)}\right)^{\frac{\lambda}{\lambda-1}} \underline{t}^{\frac{\lambda}{\lambda-1}}. \tag{5.1}$$

Thus, for agents with skill $t \geq \underline{t}$, the principal collects the profit $\hat{s}(t) - c^{\hat{s}}(\hat{s}(t)) = (\frac{\lambda}{\lambda+\alpha}(\frac{t}{\underline{t}})^{\frac{\lambda}{\lambda-1}} + \frac{(\lambda-1)\alpha}{\lambda(\lambda+\alpha)})\hat{s}(\underline{t})$. The fraction of total revenue that is collected by the principal as profit is (the conditions $\lambda > 1$ and $\lambda/(\lambda-1) < \alpha$ ensure that the total revenue is finite) $\Gamma = \frac{\int_{\underline{t}}^{\infty} (\hat{s}(t) - c^{\hat{s}}(\hat{s}(t))) m(t) dt}{\int_{\underline{t}}^{\infty} \hat{s}(t) m(t) dt} = 1 - \frac{1}{\lambda}$.

Though the principal's marginal share is constant, its average share of profit is declining from $1-\alpha/(\lambda(\alpha+\lambda))$ at $\underline{t}$ to $\lambda/(\alpha+\lambda)$ at ∞. Effectively, the contract is equivalent (in the absence of agents with skill below $\underline{t}$) to a two-tiered structure with higher fees up to $\underline{t}$ and lower fees above that level. In this setting, the principal's total share is insensitive to the parameter α that controls the distribution of agents' skill.

5.2. Linear-power costs. With a linear component in the cost function ($b > 0$), the marginal cost at any positive effort is at least b. If this marginal cost is sufficiently large so that $(1+\alpha)b/\alpha > \underline{t}$, the optimal contract selectively allows only agents with sufficiently high skill to work, while agents with skill below $(1+\alpha)b/\alpha$ do not participate. On the other hand, if $(1+\alpha)b/\alpha \leq \underline{t}$, then all agents participate.

For every $t \geq \hat{t} := \frac{(1+\alpha)b}{\alpha} \vee \underline{t}$, the principal's marginal rate is $c_x^{\hat{s}}(\hat{s}(t)) = \frac{1}{t} f_y(\frac{\hat{s}(t)}{t}) = (\alpha t + (\lambda-1)b)/((\alpha+\lambda)t)$, which decreases in t (thus also in $\hat{s}(t)$) and converges to $\alpha/(\lambda+\alpha)$ as skill t becomes large.

The high-cost regime corresponds to $b \geq \alpha\underline{t}/(1+\alpha)$, in which case $\hat{t} = (1+\alpha)b/\alpha$ and Γ simplifies to $1/\alpha$, which means that the principal shares more with agents if they are less unequal in skill. This observation is counterintuitive in part because one might expect that the more agents differ from one another, the lower the principal's leverage in using a single contract to induce their desired behavior.

However, as the distribution of skill becomes more heavy-tailed, the principal's focus progressively shifts on extracting profits from the highest skilled, while neglecting completely agents of increasingly higher skill. Indeed, the skill threshold for inertia $\hat{t} = (1 + 1/\alpha)\, b$ also increases with inequality, i.e., decreases with α.

For $0 < b < \alpha\underline{t}/(1+\alpha)$, the integrals that define the principal's share Γ are evaluated numerically, and Figure 1 displays the effect of b, with $a = 1$, $\underline{t} = 1$, $\lambda = 2$, and $\alpha = 3$. When the linear cost b is absent ($b = 0$), the principal's share is $\Gamma = 1 - 1/\lambda = 0.5$, as in section 5.1. As b increases to $\alpha\underline{t}/(1+\alpha) = 0.75$, Γ transitions from $1 - 1/\lambda$ to $1/\alpha = 1/3$, the value in the high-cost regime, where it remains for any $b \geq \alpha\underline{t}/(1+\alpha)$.

6. High-low reservation utilities and countervailing effects. The main result assumes that agents can only choose between the contract offered by the principal and the alternative of a zero payoff. In particular, such an alternative is the same for all agents, regardless of their skill. This section investigates how the possibility that agents of different skill t have a different reservation utility $R(t)$ affects the conclusions of the main result.

For tractability, the discussion focuses on a high-low regime, whereby agents below some skill threshold have a certain reservation utility, while agents above that skill have a higher one. Even such an ostensibly simple setting leads to a rather complex solution that entails several cases, depending on the relative values of such utilities.[6]

An agent with skill t and reservation utility $R(t)$ accepts a contract $c \in C$ if and only if it yields the utility $u^c(t) = c(ty^c(t)) - f(t, y^c(t)) \geq R(t)$. Thus, denote by $\chi(c) = \{t \geq 0 : u^c(t) \geq R(t)\}$ the set of skills for which the contract $c \in C$ is acceptable. The principal's goal is to maximize (with the implicit requirement that the integral exists for the contract c) $P^R(c) = \int_{\chi(c)} (ty^c(t) - c(ty^c(t)))m(t)dt$. Note that the main result in the previous sections corresponds to the reservation utility $R(t) = -f(t, 0)$. While in that case, under the optimal contract, only agents below a certain skill do not participate, the presence of a skill-dependent reservation utility $R(t)$ may also lead another layer of agents to shun the optimal contract, as explained below. This phenomenon is known in the economics literature as the countervailing effect [15, 17, 14] and has recently arisen also in electricity pricing [1].

To begin the discussion, the next result shows that the necessary conditions for the optimal contract in section 3 still hold with skill-dependent reservation utility, allowing the principal to focus on the set $\tilde{C}$ of contracts that satisfy the basic a priori restrictions. Such a result is valid in general, in that it does not require the reservation utility to take only two values.

LEMMA 6.1. $\sup_{c \in C} P^R(c) = \sup_{c \in \tilde{C}} P^R(c)$, *where* $\tilde{C}$ *is defined as in* (3.1).

Let the reservation utility take only the values $R_0 < R_1$, where $R(t) = R_0$ for all $t < t_1$ and $R(t) = R_1$ for $t \geq t_1$. The next lemma shows that countervailing effects may arise, so that agents with mediocre skills may not participate, depending on their reservation utility R_1. The intuition behind this assumption is that higher-skilled agents typically have additional opportunities other than the proposed contract and hence require a higher utility to accept it (cf. Lemma A.9).

If R_1 were equal to R_0 or slightly higher, then any agent with skill above the minimal $\underline{t}_0$ would participate, including agents with skill above t_1, who would continue to earn a positive surplus from participation. As R_1 rises, such surplus declines because nonparticipation becomes increasingly attractive. If R_1 is large enough, nonpartici-

[6] The setting of n possible values can be tackled with the same arguments as the ones presented below but would lead to a cumbersome number of cases that does not lend itself to a concise presentation.

pation becomes optimal for agents with skill t_1 and slightly higher, while only agents with skill above the higher threshold $\underline{t}_1$ continue to participate. Henceforth, denote by S_2 the set of revenue profiles that are nonzero on the intervals $[\underline{t}_0, t_1)$ and $[\underline{t}_1, \infty)$:

$$(6.1) \quad S_2 = \{\bar{s} : \mathbb{R}_+ \mapsto \mathbb{R}_+, \bar{s}(t) = s 1_{[\underline{t}_0, t_1) \cup [\underline{t}_1, \infty)} \text{ for } \underline{t}_0, \underline{t}_1 \text{ such that } t_1 \in [\underline{t}_0, \underline{t}_1] \text{ and } s \text{ as in (3.2)}\}.$$

Lemma 6.2. *For any $c \in \tilde{C}$, its corresponding constrained revenue s^R coincides with an element of S_2 up to finitely many points.*

Similar to the result without reservation utility, we now construct a (differentiable) canonical contract $c^{s^R,R}$ for each $s^R \in S_2$, which allows the principal to further narrow down the candidate optimal contract to a member of this class. For z in the range of s^R (i.e., $0 < z = s^R(t)$ for some t), marginal compensation must match marginal disutility (i.e., $c_x^{s^R,R}(z) = \frac{1}{t} f_y\left(t, \frac{z}{t}\right)$). The difficulty is to understand the value of marginal compensation $c_x^{s^R,R}(z)$ outside the range of s^R; specifying the contract's payoff for revenues that are not attained by any agent is important precisely to ensure that no agent has the incentive to generate such revenues.

The next definition summarizes the notation for the quantities that describe the optimal contract in the general case of $R_1 > R_0$.

Definition 6.3. *Let $s^R \in S_2$ and its corresponding $\underline{t}_0$, $\underline{t}_1$ be as in* (6.1). *Denote*

(i) $\bar{s}_0 = \sup_{t<t_1} s^R(t)$;

(ii) $t_z^R = \inf\{t : s^R(t) > z\}$;

(iii) $g_0(t) = \int_0^{s^R(\underline{t}_0)} \frac{1}{t} f_y(t, \frac{z}{t}) dz - f(\underline{t}_0, \frac{s^R(\underline{t}_0)}{\underline{t}_0})$;

(iv) $t_0^* = \begin{cases} \underline{t}_0 & \text{if } g_0(\underline{t}_0) \geq R_0, \\ g_0^{-1}(R_0) & \text{otherwise}; \end{cases}$

(v) $g_1(t) = \int_0^{s^R(\underline{t}_0)} \frac{1}{t_0^*} f_y(t_0^*, \frac{z}{t_0^*}) dz + \int_{s^R(\underline{t}_0)}^{\bar{s}_0} \frac{1}{t_z^R} f_y(t_z^R, \frac{z}{t_z^R}) dz + \int_{\bar{s}_0}^{s^R(\underline{t}_1)} \frac{1}{t} f_y(t, \frac{z}{t}) dz - f(\underline{t}_1, s^R(\underline{t}_1)/\underline{t}_1)$;

(vi) $t_1^* = \begin{cases} \underline{t}_1 & \text{if } g_1(\underline{t}_1) > R_1, \\ g_1^{-1}(R_1) & \text{otherwise.} \end{cases}$

The interpretation of these quantities is as follows: $\bar{s}_0$ is the maximum revenue generated by an agent with the lower reservation utility R_0, while t_z^R is the right-continuous inverse of the revenue function.

The function $g_0(t)$ represents the utility of an agent with skill $\underline{t}_0$ if the marginal compensation between 0 and $s^R(\underline{t}_0)$ equals the marginal disutility of agents with skill t. Thus, t_0^* is the maximum such t (corresponding to the minimum compensation) for which agents of skill $\underline{t}_0$ participate and produce $s^R(\underline{t}_0)$: it coincides with $\underline{t}_0$ when the reservation utility R_0 is small enough and with $g_0^{-1}(R_0) < \underline{t}_0$ otherwise.

Similarly, the function $g_1(t)$ represents the utility of agents with skill $\underline{t}_1$ if the marginal compensation equals the marginal disutility with skill t_0^* in $[0, s^R(\underline{t}_0)]$, the marginal disutility with skill t_z^R in $[s^R(\underline{t}_0), \bar{s}_0]$, and the marginal disutility with skill t in $[\bar{s}_0, s^R(\underline{t}_1)]$. Thus, t_1^* corresponds to the minimum compensation required for the participation of an agent with skill $\underline{t}_1$, who produces $s^R(\underline{t}_1)$. It coincides with $\underline{t}_1$ if the reservation utility R_1 is small enough, and with $g_1^{-1}(R_1) < \underline{t}_1$ otherwise.

The next definition constructs a canonical contract $c^{s^R,R}$ for each $s^R \in S_2$, distinguishing four cases, which correspond to the relative sizes of R_0 and R_1. Note that, at this stage, the contract does not have to be optimal. The goal is simply to construct a contract that induces a prescribed revenue function s^R under the participation constraint.

DEFINITION 6.4. *For each $s^R \in S_2$, the canonical contract is defined as $c^{s^R,R} = d + \int_0^x c_x^{s^R,R}(z)dz$, where the constant d and the function $c_x^{s^R,R}(z)$ are identified as follows: $c_x^{s^R,R}(z) = \frac{1}{t_z^R} f_y(t_z^R, \frac{z}{t_z^R})$ for $z \in [s^R(\underline{t}_0), \bar{s}_0) \cup [s^R(\underline{t}_1) \vee \bar{s}_0, \infty)$, and the following hold:*

(i) *If $\underline{t}_0 = t_1$, then $c_x^{s^R,R}(z) = \frac{1}{t_1^*} f_y(t_1^*, \frac{z}{t_1^*})$ for $0 \le z < s^R(\underline{t}_1)$ and $d = 0$.*

(ii) *If $\underline{t}_0 < t_1 < \underline{t}_1 = \infty$, then $c_x^{s^R,R}(z) = \frac{1}{t_0^*} f_y(t_0^*, \frac{z}{t_0^*})$ for $z \in [0, s^R(\underline{t}_0))$ and $d = 0$.*

(iii) *If $\underline{t}_0 < t_1 \le \underline{t}_1 < \infty$ and $g_1(t_1) \ge R_1$, then $d = 0$ and*

$$c_x^{s^R,R}(z) = \begin{cases} \frac{1}{t_0^*} f_y(t_0^*, \frac{z}{t_0^*}) & \text{if } 0 \le z < s^R(\underline{t}_0), \\ \frac{1}{t_1^*} f_y(t_1^*, \frac{z}{t_1^*}) & \text{if } \bar{s}_0 \le z < s^R(\underline{t}_1). \end{cases}$$

(iv) *If $\underline{t}_0 < t_1 \le \underline{t}_1 < \infty$ and $g_1(t_1) < R_1$, then*

$$c_x^{s^R,R}(z) = \begin{cases} \frac{1}{\underline{t}_0} f_y(\underline{t}_0, \frac{z}{\underline{t}_0}) & \text{if } 0 \le z < s^R(\underline{t}_0), \\ \frac{1}{t_1} f_y(t_1, \frac{z}{t_1}) & \text{if } \bar{s}_0 \le z < s^R(\underline{t}_1), \end{cases} \quad \text{and}$$

$$d = R_1 - \int_0^{s^R(\underline{t}_1)} c_x^{s^R,R}(z)dz - f(\underline{t}_1, s^R(\underline{t}_1)/\underline{t}_1) > 0.$$

The rationale of the above construction is as follows: for $z \in [s^R(\underline{t}_0), \bar{s}_0) \cup [s^R(\underline{t}_1), \infty)$ (i.e., $z = s^R(t)$ for some $t \in [\underline{t}_0, t_1) \cup [\underline{t}_1, \infty)$), the marginal compensation $c_x^{s^R,R}(z)$ matches $f_y(t_z^R, z/t_z^R)$, the marginal disutility of agents who produce z, which is the same as in Definition 3.2, as this is the range of skills in which the participation constraint is not binding.

On the remaining range $z \in [0, s^R(\underline{t}_0)] \cup [\bar{s}_0, s^R(\underline{t}_1))$, the principal needs to choose the marginal compensation $c_x^{s^R,R}(z)$ carefully, so that $c^{s^R,R}$ indeed induces s^R even without a participation constraint for agents with skills $t \in [\underline{t}_0, t_1) \cup [\underline{t}_1, \infty)$, and they (and only they) participate under the constraint; i.e., their utility under $c^{s^R,R}$ achieves the required levels R_0 and R_1, respectively. In this sense, the marginal compensation must be chosen as to reflect the shadow price of the participation constraint. This construction leads to the following four cases, illustrated in Figure 2.

In Case (i), no agents with skill below $\underline{t}_1$ participate. Thus, the marginal compensation between 0 and $s^R(\underline{t}_1)$ needs to guarantee that agents with skill $\underline{t}_1$ participate and that no one with lower skills does. To ensure such an outcome, the principal starts by considering $\frac{1}{\underline{t}_1} f_y(\underline{t}_1, \frac{z}{\underline{t}_1})$, which is the smallest compensation that would induce a $\underline{t}_1$-agent (i.e., an agent with skill $\underline{t}_1$) to produce $s^R(\underline{t}_1)$ while preventing anyone with lower skill from participating, in the baseline case without reservation utility. If the corresponding utility for the $\underline{t}_1$-agent satisfies $g_1(\underline{t}_1) \ge R_1$ (notice that in this case $s^R(\underline{t}_0) = \bar{s}_0 = 0$), then the participation constraint is slack, and the marginal compensation considered is correct. Otherwise, the constraint is binding; hence a higher compensation is required. Such higher compensation is $\frac{1}{t_1^*} f_y(\underline{t}_1^*, \frac{z}{t_1^*})$, which guarantees

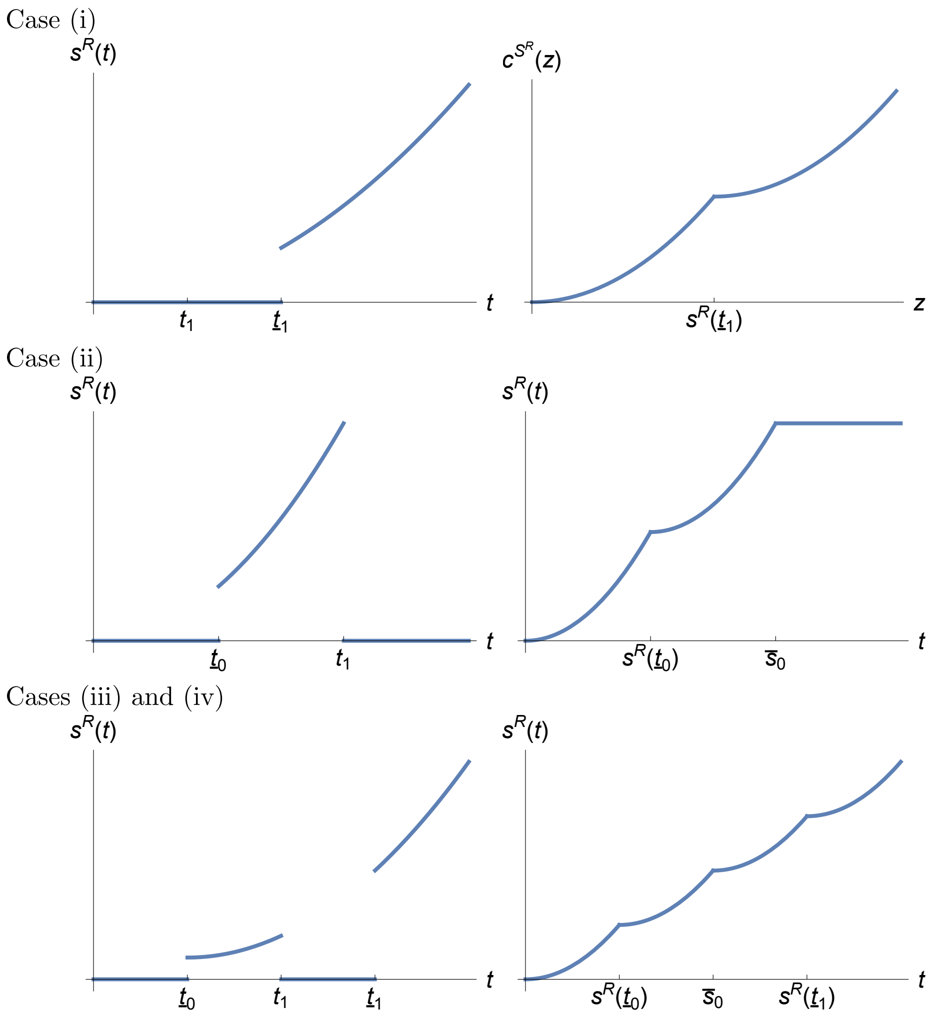

FIG. 2. *Construction of canonical contracts.*

exactly the utility $g_1(\underline{t}_1^*) = R_1$ for $\underline{t}_1$-agents.[7]

In Case (ii), only agents with skill $t \in [\underline{t}_0, t_1)$ participate, and t_0^* plays a similar role as t_1^* in Case (i) by guaranteeing that they produce $s^R(t)$ and that the utility of a $\underline{t}_0$-agent is exactly R_0. As Lemma 6.5 below implies, if s^R is indeed induced by some $c \in \tilde{C}$, then the utility for agents with skill $t \geq t_1$ is below R_1 under $c^{s^R,R}$, and they do not participate.

In Case (iii), agents with skill $t \in [\underline{t}_0, t_1) \cup [\underline{t}_1, \infty)$ participate and t_0^* plays the same role as in Case (ii). Since agents with skills $t \in [t_1, \underline{t}_1)$ do not participate, the principal has flexibility in choosing the marginal compensation $c_x^{s^R,R}(z)$ for $\bar{s}_0 \leq z \leq s^R(\underline{t}_1)$, leading to potentially different utilities for $\underline{t}_1$-agents. $g_1(t)$ represents such a utility

[7]Observe that, as g_1 is decreasing and $g_1(0) = \infty$, $t_1^* \leq \underline{t}_1$, and in particular it exists (similarly for t_0^*). Though the contract induces all agents with skill $\max(t_1, t_1^*) \leq t < \underline{t}_1$ to participate in the baseline case without reservation utility, it is not the case with reservation utility R_1, as the corresponding utility is strictly increasing in skill at $\underline{t}_1$ by Lemma A.9.

if the principal uses as compensation $c_x^{s^R,R}(z) = f_y(t, z/t)$, the marginal disutility of a t-agent. By definition, t_1^* is the highest skill for which $g_1(t_1^*) \geq R_1$, and with reservation utility R_1, only agents with skills $t \geq \underline{t}_1$ participate.

In Case (iv), $g_1(t_1) < R_1$, which means that even with the largest allowable marginal compensation $\frac{1}{t_1} f_y(t_1, \frac{z}{t_1})$ for $\bar{s}_0 \leq z \leq s^R(\underline{t}_1)$,[8] a $\underline{t}_1$-agent still does not achieve the required utility R_1. In order to ensure their participation (and thus that of all agents with higher skills), the principal needs to provide an additional subsidy. To avoid changing the behavior of agents with skill $t \in [\underline{t}_0, t_1)$, such a subsidy entails a lump sum d, which, in view of Proposition 6.6, is the cheapest compensation scheme to induce s^R under the participation constraint.

Furthermore, with a lump sum $d > 0$, the marginal compensation for $z \in [0, s^R(\underline{t}_0))$ also has to be modified. If $t_0^* < \underline{t}_0$ and the principal continues to pay marginally at $\frac{1}{t_0^*} f_y(t_0^*, \frac{z}{t_0^*})$, which guarantees that $\underline{t}_0$-agents achieve exactly utility R_0, then the lump sum induces agents with skill sufficiently close to $\underline{t}$ to participate, contradicting s^R. Thus, the marginal compensation must be equal to $\frac{1}{\underline{t}_0} f_y(\underline{t}_0, \frac{z}{\underline{t}_0})$, the marginal disutility of $\underline{t}_0$-agents, which guarantees that even in the baseline case without reservation utility, agents with skill $t \leq \underline{t}_0$ produce nothing. Finally, with all the marginal compensation decided, the lump sum d is determined by guaranteeing exactly the utility R_1 for $\underline{t}_1$-agents.

The next lemma identifies more properties of the constrained revenue s^R that is actually induced by some $c \in \tilde{C}$.

LEMMA 6.5. *Let $c \in \tilde{C}$ and s^R be its constrained revenue.*

(i) *If $\underline{t}_0 < t_1 < \underline{t}_1 < \infty$, then $g_1(\underline{t}_1) \leq R_1$.*

(ii) *If $\underline{t}_0 < t_1 < \underline{t}_1 = \infty$, then $c^{s^R,R}(\bar{s}_0) - \lim_{t\to\infty} f\left(t, \frac{\bar{s}_0}{t}\right) \leq R_1$.*

The next result shows that the principal can search for the optimal contract among all $c^{s,R}$'s with $s \in \hat{S}_2$, a subset of S_2, which satisfies more constraints, including those in Lemma 6.5, because they generate more profits than any contracts in $\tilde{C}$ that induce the same constrained revenue. The profits maximization problem can then be formulated as a calculus of variation problem over $\hat{S}_2$.

PROPOSITION 6.6.

(i) *Let $c \in \tilde{C}$ and its constrained revenue be s^R. Then, the constrained revenue induced by the canonical contract $c^{s^R,R}$ coincides with s^R, and $P^R(c) \leq P^R(c^{s^R,R})$.*

(ii) *Set $h(u,t) = s^R(t) - \int_u^t (\frac{s^R(v)}{v^2} f_y(v, \frac{s^R(v)}{v}) - f_t(v, \frac{s^R(v)}{v}))dv - f(t, \frac{s^R(t)}{t})$, and the maximum constrained profit $\max_{c\in C} P^R(c)$ equals*

$$\max_{s^R\in\hat{S}_2} \int_{\underline{t}_0}^{t_1} h(\underline{t}_0, t)m(t)dt + \int_{\underline{t}_1}^{\infty} h(\underline{t}_1, t)m(t)dt - R_0(M(\underline{t}_0) - M(t_1)) - \tilde{R}_1 M(\underline{t}_1),$$

where $\hat{S}_2$ is the collection of $s \in S_2$ which satisfies (a) *the two properties in Lemma* 6.5*, and* (b) *if $\underline{t}_0 < t_1 < \underline{t}_1 < \infty$, then $g_1(t_1) \geq R_1$, where*

$$\tilde{R}_1 = \begin{cases} R_1 & \text{if } \underline{t}_1 > t_1, \\ R_0 + \int_{\underline{t}_0}^{t_1} (\frac{s^R(v)}{v^2} f_y(v, \frac{s^R(v)}{v}) - f_t(v, \frac{s^R(v)}{v}))dv & \text{if } \underline{t}_1 = t_1. \end{cases}$$

Remark 6.7. (i) If the reservation utility is constant ($R_0 = R_1$) and $\underline{t}_0 < t_1$, then $\underline{t}_1 = t_1$. The canonical contracts are as in Case (iii) of Definition 6.4.

[8]With larger compensation, the agents with skills less than and sufficiently close to t_1 are better off producing $s^R(\underline{t}_1)$ instead of $s^R(t)$, contradicting the definition of s^R.

(ii) If R takes $n > 2$ values $(R_0, R_1, \dots, R_{n-1})$, for each constrained revenue s^R, the canonical contracts can be constructed similarly to Definition 6.4, by matching marginal compensation with the appropriate marginal disutility at the revenue levels which are actually being produced, which guarantees that the induced revenue coincides with s^R without reservation utility.

Then, in the gaps $[t_i, \underline{t}_i)$ of s^R, where $t_i = \min\{t > 0 : R(t) = R_i, 0 \le i \le n-1\}$ (with $t_0 = 0$ and $t_n = \infty$) and $\underline{t}_i = \min\{t_i \le t < t_{i+1} : s^R(t) > 0\}$ (equal to t_{i+1} if this set is empty), i.e., where the participation constraint is binding, apply the smallest marginal compensation so that R_i is achieved at $\underline{t}_i$.

This procedure is carried out from $i = 0$ to $i = n-1$, and, if for some i, R_i cannot be obtained at $\underline{t}_i$ by applying the largest possible marginal compensation (which guarantees s^R as the induced revenue), then the marginal compensations at all lower gaps $[t_j, \underline{t}_j)$ for $0 \le j < i$ should be revised to $\frac{1}{\underline{t}_j} f(\underline{t}_j, \frac{z}{\underline{t}_j})$, and a lump sum compensation needs to be added to guarantee R_i at $\underline{t}_i$, which also ensures R_j at $\underline{t}_j$ for $0 \le j < i$ (as demonstrated in Case (iv) of Definition 6.4). Note, however, that if this case indeed happens, as shown in Proposition 6.6, the corresponding s^R is suboptimal and can be actually excluded from the admissible set.

7. Conclusion. As peer-to-peer markets have emerged in several service sectors over the last decade, so have near-monopolistic platforms that connect users with service providers in exchange for a significant fraction of revenues.

This paper finds the optimal contract for a monopolist principal who offers a common nonlinear contract to a population of agents that differ in skill. The optimal contract always entails positive sharing rates without subsidies and implies nonparticipation for those agents with the lowest skill.

Affine contracts (with constant sharing rates) are optimal when the skill is Pareto-distributed and individual costs are of power type. When costs include a linear component, the principal's share of revenues is characterized by a low-cost regime, in which its main determinant is the growth rate of agents' costs, and by a high-cost regime, in which the main determinant of the principal's share is the dispersion of agents' skill.

Appendix A. Proofs. The next lemma shows that Assumptions 2.2 and 2.3 guarantee the existence of a (possibly infinite) optimal effort. When multiple optimizers exist, one can choose the one requiring the minimum effort.

LEMMA A.1. *Given a contract* $c \in C$, *for any* $t \ge 0$ *there exists some* $y^c(t) \in [0, \infty]$ *that maximizes* (2.1), *and such that* $y^c(t) = \min\{y \ge 0 : y \text{ maximizes } (2.1)\}$.

Proof. Let $\{y_i\}_{i=1}^{\infty}$ be a maximizing sequence for skill t. If any such sequence is unbounded, then $y^c = \infty$ by definition. Otherwise, up to a subsequence $\lim_{i\to\infty} y_i = \hat{y} < \infty$ with $\lim_{i\to\infty} u^c(t, y_i) = \sup_{y\ge 0} u^c(t, y)$. By the continuity of f and the upper-semicontinuity of c, $\hat{y}$ is a maximizer because $u^c(t,\hat{y}) = c(t\hat{y}) - f(t,\hat{y}) \ge \limsup_{i\to\infty} c(ty_i) - \lim_{i\to\infty} f(t, y_i) \ge \sup_y u^c(t,y) \ge u^c(t,\hat{y})$. Finally, $y^c(t) = \inf\{y : u^c(t,y) = \sup_{y\ge 0} u^c(t,y)\}$ is finite as the set is not empty, and is also a maximizer, as the limit of a maximizing sequence, by the same argument. ∎

Remark A.2. Note that an optimal infinite effort at this point cannot be ruled out, as the contract c could be arbitrarily generous. Such contracts are obviously suboptimal for the principal, and section 3 identifies restrictions for optimality, under which the optimal effort is necessarily finite also for the agent.

However, if $y^c(t) < \infty$, its optimality implies that, for any $y \geq 0$,

$$\int_y^{y^c(t)} f_y(t,u)du = f\left(t, y^c(t)\right) - f\left(t, y\right) \leq c(ty^c(t)) - c(ty), \tag{A.1}$$

which means that any change in effort implies more cost than income increase.

Proof of Proposition 3.1. The result proceeds through several steps: (i) is proved through Lemmas A.3 and A.4, while (ii) is proved through Lemma A.5 below. ∎

The next lemma shows that negative sharing rates for the agent can be excluded, as their profits are the same as those obtained by capping such rates at zero. More formally, any contract c does not generate more profit than its increasing envelope.

LEMMA A.3. *For any contract $c \in C$, let $c^*(x) = \sup_{0\leq y\leq x} c(y)$. Then $c^* \in C$, $y^c(t) = y^{c^*}(t)$ for any $t \geq 0$, and $P(c) = P(c^*)$.*

Proof. If $c \in C$, then $c^* \in C$, because c^* is increasing and right-continuous and hence upper-semicontinuous. For any x such that $c(x) < c^*(x)$, choose x^* so that $x^* < x$, $c(x^*) > c(x)$, and $c^*(x^*) = c^*(x)$. Then, for any $t > 0$, $c(x^*) - f(t, x^*/t) > c(x) - f(t, x/t)$. Thus, under contract c, t-agents do not choose $y = x/t$. For the same reason, they do not choose $y = x/t$ under c^*. For agents with 0 skill, the optimal choice is always $y = 0$. Thus, under both c and c^*, t-agents would only consider an effort y's such that $c(ty) = c^*(ty)$, and therefore, $y^c(t) = y^{c^*}(t)$. Hence, $P(c) = P(c^*)$. ∎

As the previous lemma excludes negative sharing rates for the agent, the next one shows that negative rates for the principal do not maximize profits. Equivalently, for any optimal contract, the agent's income must be lower than the total revenue.

LEMMA A.4. *For any nondecreasing $c \in C$, the contract $\tilde{c}(x) = \min(c(x), x) \in C$ and $P(\tilde{c}) \geq P(c)$.*

Proof. To prove that $\tilde{c}(x) \in C$, it suffices to show that $\tilde{c}$ is upper-semicontinuous. Consider a sequence $\{x_i\}_{i=1}^{\infty}$ that converges to x. Since c is upper-semicontinuous, $\limsup_{i\to\infty} \tilde{c}(x_i) = \limsup_{i\to\infty} \min(c(x_i), x_i) \leq \limsup_{i\to\infty} c(x_i) \wedge \limsup_{i\to\infty} x_i \leq \min\left(c(x), x\right) = \tilde{c}(x)$. Thus, $\tilde{c}$ is also upper-semicontinuous.

To see that $P(\tilde{c}) \geq P(c)$, first note that $y^c(t) = y^{\tilde{c}}(t)$ for all $t \in A = \{t : y^c(t) < \infty, c(ty^c(t)) \leq ty^c(t)\}$. Indeed, for any $y \geq 0$,

$$\tilde{c}(ty^c(t)) - f(t, y^c(t)) = c(ty^c(t)) - f(t, y^c(t)) \geq c(ty) - f(t, y) \geq \tilde{c}(ty) - f(t, y), \tag{A.2}$$

where the first inequality follows from (A.1) and is strict for $y < y^c(t)$.

Conversely, $y^{\tilde{c}}(t)$ may differ from $y^c(t)$ for $t \in B = \{t : y^c(t) < \infty, c(ty^c(t)) > ty^c(t)\}$. Yet, any such difference must satisfy $ty^{\tilde{c}}(t) - \tilde{c}(ty^{\tilde{c}}(t)) \geq 0 > ty^c(t) - c(ty^c(t))$, which means that for the principal the contract $\tilde{c}$ is superior to c within B.

Let D be the set of skills t for which any maximizing sequence $\{y_i(t)\}_{i=1}^{\infty}$ diverges to ∞ under c. For any $t \in D$, since f is convex and $\lim_{u\to\infty} f_y(t, u) = \infty$, it follows that $f(t, y_i(t)) \geq f(t, 0) - c(0) + ty_i(t)$ for sufficiently large i. In view of $c(ty_i(t)) - f(t, y_i(t)) \geq c(0) - f(t, 0)$ (also for sufficiently large i), this means that $c(ty_i(t)) \geq ty_i(t)$. As a result, $\limsup_{i\to\infty} \left(ty_i(t) - c(ty_i(t))\right) \leq 0 \leq ty^{\tilde{c}}(t) - \tilde{c}(ty^{\tilde{c}}(t))$.

In summary, the total profits under $\tilde{c}$ and c satisfy

$$\text{(A.3)}\quad P(\tilde{c}) = \int_A \left(ty^{\tilde{c}}(t) - \tilde{c}(ty^{\tilde{c}}(t))\right) m(t)dt + \int_B \left(ty^{\tilde{c}}(t) - \tilde{c}(ty^{\tilde{c}}(t))\right) m(t)dt$$
$$+ \int_D \left(ty^{\tilde{c}}(t) - \tilde{c}(ty^{\tilde{c}}(t))\right) m(t)dt \geq \int_A (ty^c(t) - c(ty^c(t)))\, m(t)dt$$
$$+ \int_B (ty^c(t) - c(ty^c(t)))\, m(t)dt + \int_C \limsup_{i\to\infty} \left(ty^i(t) - c(ty^i(t))\right) m(t)dt \geq P(c). \quad \square$$

The following lemma summarizes the intuitively natural but mathematically nontrivial property that higher skill commands higher revenue and hence higher income.

LEMMA A.5. *Given* $c \in \tilde{C}$, $y^c(t)$ *is finite for every* t, *and* $\lim_{t\to 0} y^c(0) = 0$. *Furthermore, if* $0 < t_1 < t_2$, *then* $t_1 y^c(t_1) \leq t_2 y^c(t_2)$.

Proof. Since f is convex in y and $\lim_{y\to\infty} f_y(t, y) = \infty$, for any $t > 0$ there exists $\bar{y}(t)$ such that $f(t, y) - f(t, 0) > ty$ for all $y \geq \bar{y}(t)$. Thus, for any $y > \bar{y}(t)$, $u^c(t, y) = c(ty) - f(t, y) \leq ty - f(t, y) < -f(t, 0) = u^c(t, 0)$. Therefore, for t-agents the revenue ty is not optimal, and Lemma A.1 implies that $y^c(t)$ is finite.

For any $0 < t_1 < t_2$, (A.1) implies $f(t_1, y^c(t_1)) - f(t_1, t_2 y^c(t_2)/t_1) \leq c(t_1 y^c(t_1)) - c(t_2 y^c(t_2))$. Suppose $0 \leq t_2 y^c(t_2) < t_1 y^c(t_1)$. Because $t_2 > t_1$,

$$y^c(t_2) = t_2 y^c(t_2)/t_2 \leq t_2 y^c(t_2)/t_1,$$
$$\frac{t_1 y^c(t_1)}{t_2} - y^c(t_2) = \frac{t_1 y^c(t_1) - t_2 y^c(t_2)}{t_2} < \frac{t_1 y^c(t_1) - t_2 y^c(t_2)}{t_1} = y^c(t_1) - \frac{t_2 y^c(t_2)}{t_1}.$$

Then, since f is strictly increasing and convex in y and f_y is nonincreasing in t,

$$f(t_2, \frac{t_1 y^c(t_1)}{t_2}) - f(t_2, y^c(t_2)) = \int_{y^c(t_2)}^{\frac{t_1 y^c(t_1)}{t_2}} f_y(t_2, u)du \leq \int_{y^c(t_2)}^{\frac{t_1 y^c(t_1)}{t_2}} f_y(t_1, u)du$$
$$< \int_{\frac{t_2 y^c(t_2)}{t_1}}^{y^c(t_1)} f_y(t_1, u)du = f(t_1, y^c(t_1)) - f\left(t_1, \frac{t_2 y^c(t_2)}{t_1}\right) \leq c(t_1 y^c(t_1)) - c(t_2 y^c(t_2)),$$

which implies that $c(t_2 y^c(t_2)) - f(t_2, y^c(t_2)) < c(t_1 y^c(t_1)) - f(t_2, t_1 y^c(t_1)/t_2)$. This contradicts the optimality of $y^c(t_2)$ and implies that $t_2 y^c(t_2) \geq t_1 y^c(t_1)$.

For $\lim_{t\to 0} y^c(t)$ at $t = 0$, first consider the limit of $s^c(t) = ty^c(t)$. Suppose there exists $\epsilon > 0$, such that for every $t > 0$, $s^c(t) > \epsilon$. Assumption 2.3 implies that $f(t, s^c(t)/t) \geq f(t, \epsilon/t)$, which increases to infinity as $t \to 0$. On the other hand, since both $s^c(t)$ and $c(x)$ are increasing, $c(s^c(t))$ is bounded in a neighborhood of $t = 0$. Thus, there must exist a sufficiently small $t > 0$, such that $f(t, s^c(t)/t)$ is sufficiently large and $u^c(t) = c(s^c(t)) - f(t, s^c(t)/t) < c(0) - f(t, 0)$, where the right-hand side is bounded from below because $f(t, 0)$ is decreasing in t. But this contradicts the optimality of $y^c(t)$, whence $\lim_{t\to 0} s^c(t) = 0$.

Now suppose that there exists $\epsilon > 0$, such that for every $t > 0$, $y^c(t) > \epsilon$; then $u^c(t) = c(s^c(t)) - f(t, y^c(t)) < c(s^c(t)) - f(t, \epsilon)$. Because $s^c(t) \to 0$ as $t \to 0$, $\lim_{t\to 0} u^c(t) \leq c(0) - f(0, \epsilon) < c(0) - f(0, 0)$. On the other hand, from the optimality of $y^c(t) > 0$, $u^c(t) > c(0) - f(t, 0)$ for every t, and $\lim_{t\to 0} u^c(t) \geq c(0) - f(0, 0)$. This contradiction shows that $\lim_{t\to 0} y^c(t) = 0$. $\square$

Proof of Proposition 3.3. (i) Since s in increasing, by construction, for $u \geq s(t)$, $t_u \geq t$, and for $u < s(t)$, $t_u < t$. Also, since f_y is decreasing in t and increasing in y, $\frac{1}{t}f_y(t,\frac{x}{t})$ is decreasing in t. Thus, for any $x > s(t)$, $f(t,\frac{x}{t}) - f(t,\frac{s(t)}{t}) \geq \int_{s(t)}^{x} \frac{1}{t_u} f_y(t_u,\frac{u}{t_u})du = \int_{s(t)}^{x} c^s_x(u)du = c^s(x) - c^s(s(t))$, and the first inequality is strict if $x < s(t)$ and s is continuous at t, which implies that $s(t)$ is optimal for t-agents. If s has a jump at t, then this agent is indifferent between $s(t)$ and $s(t-) \leq x < s(t)$. However, s has only countably many discontinuous points.

(ii) For any $c \in \tilde{C}$, let $s(t)$ be the induced revenue. If $c^s < c$ at some $x \geq 0$, consider $\min(c^s, c) \in \tilde{C}$: for any $t > 0$, if $\min(c^s,c)(s(t)) = c^s(s(t))$, then for any $y \geq 0$, $u^{\min(c^s,c)}(t, s(t)/t) \geq u^{c^s}(t,y) \geq u^{\min(c^s,c)}(t,y)$, which follows from the optimality of $s(t)$ and implies that $s(t)$ is optimal for $\min(c^s,c)$. The same holds if $\min(c^s,c)(s(t)) = c(s(t))$. Thus, $\min(c^s,c)$ induces the same revenue and generates at least as much profit as c. Therefore, without loss of generality, assume $c \leq c^s$ in the following, which implies that $c(0) = 0$.

Suppose $P(c) > P(c^s)$; then there must exist $t^* > 0$ such that $c(s(t^*)) < c^s(s(t^*))$. However, with $c(0) = c^s(0) = 0$, it contradicts Lemma A.6 below. ☐

LEMMA A.6. *For any $s \in S$, the corresponding canonical contract c^s, and any $c \in \tilde{C}$ which induces the revenue s, $c(s(\bar{t}) - c(s(\underline{t})) \geq c^s(s(\bar{t})) - c^s(s(\underline{t}))$ for every $\bar{t} > \underline{t} \geq 0$, such that $M(\bar{t}) > 0$.*

Proof. Let the collection of all discontinuous points of s between $\underline{t}$ and $\bar{t}$ be $T = \{t_1, \ldots, t_K\}$, where K could be ∞. For each $t_i \in T$, let $\bar{x}_i = \inf\{s(t), t > t_i\}$, and $\underline{x}_i = \sup\{s(t), t < t_i\}$, so that $s(t_i) \in D_i = [\underline{x}_i, \bar{x}_i]$. Let $D = \bigcup_{i=1}^{K} D_i$.

For every $1 \leq i \leq K$, since $s(t)$ is increasing, there exists a decreasing sequence $\{t_{ij}\}_{j=1}^{\infty}$ converging to t_i and such that $s(t_{ij})$ converges to $\bar{x}_i$. Thus, from the optimality of $s(t_{ij})$, $c(s(t_{ij})) - c(\underline{x}_i) \geq f(t_{ij}, \frac{s(t_{ij})}{t_{ij}}) - f(t_{ij}, \frac{\underline{x}_i}{t_{ij}}) = \int_{\underline{x}_i}^{s(t_{ij})} \frac{1}{t_{ij}} f_y(t_{ij}, \frac{u}{t_{ij}})du$. Because c is right-continuous and f_y is continuous, Fatou's lemma implies that

$$c(\bar{x}_i) - c(\underline{x}_i) \geq \lim_{j\to\infty} \int_{\underline{x}_i}^{s(t_{ij})} \frac{1}{t_{ij}} f_y(t_{ij}, \frac{u}{t_{ij}})du \geq \int_{\underline{x}_i}^{\bar{x}_i} c^s_x(u)du. \tag{A.4}$$

Furthermore, from the definition of $\underline{x}_i$'s and $\bar{x}_i$'s, if $t_i < t_j$, then $\bar{x}_i \leq \underline{x}_j$, because otherwise, since $s(t)$ is nondecreasing, $s(t) \geq \bar{x}_i > \underline{x}_j \geq s(t)$ for every $t \in (t_i, t_j)$. Thus the only intersections between D_i and D_j are their boundary points.

Because c is increasing, it is differentiable almost everywhere, and for any interval $[a,b]$, $\int_a^b c_x(u)du \leq c(b) - c(a)$, where c_x is its derivative. For every positive integer n, let

$$f_n(u) = \begin{cases} c^s_x(u), & u \in \bigcup_{i=1}^{n} D_i, \\ c_x(u), & u \in [\underline{t}, \bar{t}] \setminus \bigcup_{i=1}^{n} D_i. \end{cases}$$

Then f_n is nonnegative and converges to

$$f(u) = \begin{cases} c^s_x(u), & u \in D, \\ c_x(u), & u \in [\underline{t}, \bar{t}] \setminus D. \end{cases}$$

Since $[s(\underline{t}), s(\bar{t})] \setminus D_1 = C_{11} \bigcup C_{12}$, where $C_{11} = [s(\underline{t}), \underline{x}_1)$ is disjoint with $C_{12} = (\bar{x}_1, s(\bar{t})]$, $c(s(\bar{t})) - c(s(\underline{t})) \geq \int_{C_{11}} c_x(u)du + \int_{D_1} c^s_x(u)du + \int_{C_{12}} c_x(u)du = \int_{s(\underline{t})}^{s(\bar{t})} f_1(u)du$, following (A.4). Similarly, since D_1 and D_2 only intersect at their end points (without loss of generality, assume $\bar{x}_1 \leq \underline{x}_2$), $[s(\underline{t}), s(\bar{t})] \setminus (D_1 \bigcup D_2) = C_{21} \bigcup C_{22} \bigcup C_{23}$, where $C_{21} = [s(\underline{t}), \underline{x}_1)$, $C_{22} = (\bar{x}_1, \underline{x}_2)$, and $C_{23} = (\bar{x}_2, s(\bar{t})]$ are pairwise disjoint, and

$c(s(\bar{t})) - c(s(\underline{t})) \geq \int_{C_{21}} c_x(u)du + \int_{D_1} c_x^s(u)du + \int_{C_{22}} c_x(u)du + \int_{D_2} c_x^s(u)du + \int_{C_{23}} c_x(u)du = \int_{s(\underline{t})}^{s(\bar{t})} f_2(u)du$. The same argument holds for every n and implies $c(s(\underline{t})) - c(s(\bar{t})) \geq \int_{s(\underline{t})}^{s(\bar{t})} f_n(u)du$. Then by Fatou's lemma,

$$c(s(\bar{t})) - c(s(\underline{t})) \geq \liminf_{n\to\infty} \int_{s(\underline{t})}^{s(\bar{t})} f_n(u)du \geq \int_{s(\underline{t})}^{s(\bar{t})} f(u)du. \tag{A.5}$$

Finally, for $u \in [s(\underline{t}), s(\bar{t})]\backslash D$, because there exists $t \geq 0$ such that $s(t) = u$, by the optimality of $y = u/t$ for t-agents, for any $h > 0$, (A.1) implies that $\frac{c(u)-c(u-h)}{h} \geq \frac{1}{t}\frac{f(t,u/t)-f(t,(u-h)/t)}{h/t}$. Thus, for almost every $u \in [s(\underline{t}), s(\bar{t})]\backslash D$, $c_x(u) \geq \frac{1}{t} f_y\left(t, \frac{u}{t}\right) = c_x^s(u)$, which implies that $f \geq c_x^s$. Together with (A.5), this implies that $c(s(\bar{t})) - c(s(\underline{t})) \geq \int_{s(\underline{t})}^{s(\bar{t})} f(u)du \geq \int_{s(\underline{t})}^{s(\bar{t})} c_x^s(u)du = c^s(s(\bar{t})) - c^s(s(\underline{t}))$. ∎

LEMMA A.7. *For a differentiable contract $c \in \tilde{C}$ and its induced revenue s, let $\bar{t} = \sup\{t : M(t) > 0\}$ and $s(\bar{t}) = \lim_{t\to\bar{t}} s(t)$. If there exist $0 \leq t_1 < t_2 \leq \bar{t}$ and a constant a such that $\inf_{s(t_1)\leq z<s(t_2)} c_x(z) \geq a$, and in the case of $t_2 < \bar{t}$, $c_x(s(t_2)) \leq a$, then $\tilde{c} = \int_0^x \tilde{c}_x(z)dz$, where*

$$\tilde{c}_x(z) = \begin{cases} c_x(z), & z < s(t_1) \text{ or } x \geq s(t_2), \\ a, & s(t_1) \leq z < s(t_2), \end{cases}$$

induces the revenue $\tilde{s}$ such that $\tilde{s}(t) = s(t)$ if $t \leq t_1$ or $t \geq t_2$, and $s(t_1) \leq \tilde{s}(t) \leq s(t)$ if $t_1 < t < t_2$.

Proof. If $s(t_2) = 0$, the monotonicity of s implies that $s(t_1) = s(t_2)$, and by definition $c = \tilde{c}$. In the following, assume that $s(t_2) > 0$ and $s(t_2) > s(t_1)$. We prove the case of $t_2 < \bar{t}$, and the case of $t_2 = \bar{t}$ follows similarly.

For any $t > 0$, the optimality of $s(t)$ implies that for any x,

$$c(s(t)) - c(x) \geq f\left(t, \frac{s(t)}{t}\right) - f\left(t, \frac{x}{t}\right), \tag{A.6}$$

and the inequality is strict if $x < s(t)$.

For $t \leq t_1$, (A.6) still holds for $x \leq s(t_1)$ if c is replaced by $\tilde{c}$, because $c = \tilde{c}$ at both x and $s(t) \leq s(t_1)$. On the other hand, for $x > s(t_1)$, since $\tilde{c}_x(u) \leq c_x(u)$ for every $u \geq s(t_1)$, $c(x) - c(s(t)) \geq \tilde{c}(x) - \tilde{c}(s(t))$, which indicates that $\tilde{c}(s(t)) - \tilde{c}(x) \geq c(s(t)) - c(x) \geq f(t, \frac{s(t)}{t}) - f\left(t, \frac{x}{t}\right)$. Thus $\tilde{s}(t) = s(t)$.

For $t \geq t_2$, (A.6) still holds for $x \geq s(t_2)$ if c is replaced by $\tilde{c}$, because the two contracts coincide marginally at every $x \geq s(t_2)$. Thus, any $x \geq s(t_2)$ other than $s(t)$ is not optimal for t-agents under $\tilde{c}$. Together with Lemma A.5, this implies that the optimal revenue under $\tilde{c}$ at t_2 is constrained to $s(t_1) \leq x \leq s(t_2)$. The optimality of $s(t_2)$ under c implies that $c_x(s(t_2)) = \frac{1}{t_2} f(t_2, \frac{s(t_2)}{t_2}) \leq a$. Furthermore, for any $s(t_1) < x < s(t_2)$, $\tilde{c}(s(t_2)) - \tilde{c}(x) = \int_x^{s(t_2)} \tilde{c}_x(u)du = \int_x^{s(t_2)} adu > \int_x^{s(t_2)} \frac{1}{t_2} f_y(t_2, \frac{u}{t_2})du = f(t_2, s(t_2)) - f(t_2, x)$, where the inequality follows from the strict convexity of $f(t_2, \cdot)$ and implies that x is not optimal. Thus $\tilde{s}(t_2) = s(t_2)$. From Lemma A.5, the choice for agents with skill $t > t_2$ is further constrained to $x \geq s(t_2)$, which implies $\tilde{s}(t) = s(t)$.

Finally, for $t_1 < t < t_2$, because $\tilde{s}(t_1) = s(t_1)$ and $\tilde{s}(t_2) = s(t_2)$, Lemma A.5 implies that the choice for t-agents is constrained to $[s(t_1), s(t_2)]$. For any $s(t) < x < s(t_2)$, because $\tilde{c}$ grows more slowly than c from $s(t)$ to x, (A.6) implies that

$\tilde{c}(s(t)) - \tilde{c}(x) \geq c(s(t)) - c(x) \geq f(t, s(t)/t) - f(t, x/t)$, and thus x is not optimal under $\tilde{c}$ for agents with skill t. Therefore, $\tilde{s}(t) \leq s(t)$. ∎

Proof of Theorem 3.4. Let $\bar{t} = \sup\{t : M(t) > 0\}$, and suppose there exists $t_1 < \bar{t}$ such that $c_x^{\hat{s}}(\hat{s}(t_1)) = 1 + \epsilon$ for some $\epsilon > 0$. Let $t_2 = \inf\{t_1 < t < \bar{t} : c_x^{\hat{s}}(\hat{s}(t)) \leq 1\}$, with the convention that $t_2 = \bar{t}$ if this set is empty. Since $c_x^{\hat{s}}$ is right-continuous and only has negative jumps, from the definition of t_2, $t_2 > t_1$, and if $t_2 < \bar{t}$, $c_x^{\hat{s}}(\hat{s}(t_2)) \leq 1$. Let $s(\bar{t}) = \lim_{t \to \bar{t}} s(t)$; then Lemma A.7 implies that there exists a contract $\tilde{c}$, with

$$\tilde{c}_x(z) = \begin{cases} c_x^{\hat{s}}(z), & z < s(t_1) \text{ or } x \geq s(t_2), \\ 1, & s(t_1) \leq z < s(t_2), \end{cases}$$

and the corresponding revenue $\tilde{s}$ satisfies $\tilde{s}(t) = \hat{s}(t)$ if $t \leq t_1$ or $t \geq t_2$, and $\hat{s}(t_1) < \tilde{s}(t) \leq \hat{s}(t)$ if $t_1 < t < t_2$. Thus

$$\begin{aligned} P(\tilde{c}) - P\left(c^{\hat{s}}\right) &= \int_{t \leq t_1, t \geq t_2} \left(c^{\hat{s}}(\hat{s}(t)) - \tilde{c}(\hat{s}(t))\right) m(t) dt \\ &+ \int_{t_1}^{t_2} \left(\tilde{s}(t) - \tilde{c}(\tilde{s}(t)) - \hat{s}(t) + c^{\hat{s}}\left(\hat{s}(t)\right)\right) m(t) dt \end{aligned}$$

$$\text{(A.7)} \quad = \int_{t \geq t_2} \left(c^{\hat{s}}(\hat{s}(t)) - \tilde{c}(\hat{s}(t))\right) m(t) dt + \int_{t_1}^{t_2} \left(\tilde{s}(t) - \tilde{c}(\tilde{s}(t)) - \hat{s}(t) + c^{\hat{s}}\left(\hat{s}(t)\right)\right) m(t) dt,$$

where the last equation follows from the fact that $\tilde{c}(x) = c^{\hat{s}}(x)$ for $x \leq s(t_1)$.

If $t > t_2$, then $\hat{s}(t) \geq \hat{s}(t_2)$ and $c^{\hat{s}}(\hat{s}(t)) - \tilde{c}(\hat{s}(t)) = \int_{\hat{s}(t_1)}^{\hat{s}(t_2)} \left(c_x^{\hat{s}}(z) - 1\right) dz > 0$. Thus the first term in (A.7) is positive.

If $t_1 \leq t \leq t_2$, Lemmas A.5 and A.7 imply that $\hat{s}(t_1) \leq \tilde{s}(t) \leq \hat{s}(t) \leq \hat{s}(t_2)$. Furthermore, since $\tilde{c}_x(z) = 1$ for every $\hat{s}(t_1) < z < \hat{s}(t_2)$, $\tilde{s}(t) - \tilde{c}(\tilde{s}(t)) = \hat{s}(t) - \tilde{c}(\hat{s}(t))$. Thus, by the definition of t_2, $\tilde{s}(t) - \tilde{c}(\tilde{s}(t)) - \hat{s}(t) + c^{\hat{s}}\left(\hat{s}(t)\right) = -\tilde{c}(\hat{s}(t)) + c^{\hat{s}}\left(\hat{s}(t)\right) = c^{\hat{s}}\left(\hat{s}(t)\right) - c^{\hat{s}}\left(\hat{s}(t_1)\right) - \tilde{c}(\hat{s}(t)) + \tilde{c}(\hat{s}(t_1)) = \int_{\hat{s}(t_1)}^{\hat{s}(t)} \left(c_x^{\hat{s}}(z) - 1\right) \geq 0$. This implies that the second term in (A.7) is nonnegative, and $P(\tilde{c}) > P\left(c^{\hat{s}}\right)$.

Thus, from Proposition 3.3, for $c^{\tilde{s}}$ defined in Definition 3.2 corresponding to $\tilde{s}$, $P\left(c^{\tilde{s}}\right) > P\left(c^{\hat{s}}\right)$, which contradicts the optimality of $\hat{s}$ and shows that $c_x^{\hat{s}}\left(\hat{s}(t_1)\right) \leq 1$. ∎

The next lemma represents $u^{c^s}(s(t))$, or, equivalently, $c^s(s(t))$ as an integral with respect to t. Such a representation is key to the identification of the optimal contract in the main result Theorem 4.2.

Lemma A.8. *Let c^s be the canonical contract in Definition* 3.2 *for some admissible revenue $s \in S$, and let $y(t) = s(t)/t$. Then the corresponding utility $u^{c^s}(t) = c^s(ty(t)) - f(t, y(t))$ satisfies $\frac{du^{c^s}(t)}{dt} = \frac{y(t)}{t} f_y(t, y(t)) - f_t(t, y(t))$ almost everywhere, and $u^{c^s}(t) - u^{c^s}(0) = \int_0^t \left(\frac{y(v)}{v} f_y(v, y(v)) - f_t(v, y(v))\right) dv$.*

Proof. Let $t_0 = \inf\{t : s(t) > 0\}$. For $0 \leq t < t_0$, $u^{c^s}(t) = c^s(0) - f(t, 0) =$

$-f(t,0)$ and is differentiable, with $u_t^{c^s}(t) = -f_t(t,0)$. For $t > t_0$ and $h > 0$,

$$\begin{aligned}
&\frac{u^{c^s}(t+h) - u^{c^s}(t)}{h} \\
=&\frac{c^s((t+h)y(t+h)) - c^s(ty(t)) - f(t,(t+h)y(t+h)/t) + f(t,y(t))}{h} \\
&+ \frac{f(t,(t+h)y(t+h)/t) - f(t+h,y(t+h))}{h} \\
\leq&\frac{f\left(t,(t+h)y(t+h)/t\right) - f(t+h,y(t+h))}{h} \\
=&\frac{f(t,\frac{(t+h)y(t+h)}{t}) - f(t,y(t+h))}{hy(t+h)/t}\frac{y(t+h)}{t} + \frac{f(t,y(t+h)) - f(t+h,y(t+h))}{h},
\end{aligned}$$

following from the optimality of $y(t)$. Similarly, the optimality of $y(t+h)$ implies that

$$\begin{aligned}
\frac{u^{c^s}(t+h) - u^{c^s}(t)}{h} &\geq \frac{f(t,y(t)) - f\left(t+h, ty(t)/(t+h)\right)}{h} \\
&= \frac{f(t,y(t)) - f(t+h,y(t))}{h} + \frac{f\left(t+h,y(t)\right) - f\left(t+h,ty(t)/(t+h)\right)}{hy(t)/(t+h)}\frac{y(t)}{t+h},
\end{aligned}$$

and for $h < 0$ the reverse inequalities hold.

Because $f(0,t)$ and f_y are nonincreasing in t by Assumption 2.3(iii), it follows that $f(t_1,y) \geq f(t_2,y)$ for any $y \geq 0$ if $t_1 \leq t_2$, and $f_t \leq 0$. Thus u is an increasing function by (10). Furthermore, the following hold.

(i) If s is continuous at t, because f_y, f_t is continuous, both the lower and upper bounds converge as h goes to 0, and $\frac{du^{c^s}(t)}{dt} = \frac{y(t)}{t}f_y(t,y(t)) - f_t(t,y(t))$.

(ii) If s is discontinuous at t, since $y(t)$ is finite, $\limsup_{h\uparrow 0}\frac{u^{c^s}(t+h)-u^{c^s}(t)}{h} = \frac{y(t)}{t}f_y(t,y(t)) - f_t(t,y(t)) < \infty$ and $\limsup_{h\downarrow 0}\frac{u^{c^s}(t+h)-u^{c^s}(t)}{h} \leq \frac{y(t)}{t}f_y(t,y(t+)) - f_t(t,y(t+)) < \infty$, which implies that u is continuous at t.

Finally, since $s \in S$, $y(t)$ converges to 0 as t decreases from Lemma A.5. Therefore, $\lim_{t\to 0} u^{c^s}(t) = \lim_{t\to 0}\left(c^s(s(t)) - f(t,y(t))\right) = c^s(0) - f(0,0) = u^{c^s}(0)$, and u is continuous at $t = 0$.

Thus u is continuous at every $t \geq 0$ and is differentiable except for countably many points (where s is discontinuous). Furthermore, since u is increasing, its derivative is integrable. Then, Theorem 6.27 in [12] yields $u^{c^s}(t) - u^{c^s}(0) = \int_0^t \frac{du^{c^s}(v)}{dv}dv = \int_0^t(\frac{y(v)}{v}f_y(v,y(v)) - f_t(v,y(v)))dv$. ∎

Proof of Theorem 4.2. Without loss of generality, assume that $m(t) > 0$ and $M(t) > 0$ for every $t \geq 0$; otherwise t is not relevant for the optimal contract.

(i) First, from Lemma A.8, for every c^s, the individual utility function with gross income s is $u^{c^s}(t) = \int_0^t \frac{s(v)}{v^2}f_y(\frac{s(v)}{v})dv + u^{c^s}(0)$. Because $u^{c^s}(0) = -f(0)$,

$$c^s(s(t)) = \int_0^t \frac{s(v)}{v^2}f_y\left(\frac{s(v)}{v}\right)dv + f\left(\frac{s(t)}{t}\right) - f(0). \tag{A.8}$$

Then, by Fubini's theorem, the problem of profits maximization can be written as $\max_{s\in S} P(c^s) = \max_{s\in S}\int_0^\infty (s(t) - c^s(s(t)))m(t)dt = \max_{s\in S}\int_0^\infty g\left(t,s(t)\right)dt + f(0)$.

Because yf_y and f are convex, $g(t,s)$ is concave in s. Furthermore, since $f_y(\infty) = \infty$, for every $t > 0$, $g(t,\infty) = -\infty < g(t,0) = -f(0)m(t)$, which implies that the maximizer of $g(t,s)$ is finite. Let $\hat{s}(t) = \inf\{s \geq 0 : s \text{ is a maximizer of } g(t,\cdot)\}$, which, by the continuity of g, is also a maximizer $g(t,\cdot)$.

Now, because $\hat{s}(0) = 0$, in order to prove that $\hat{s}$ solves (3.3), it suffices to show that $\hat{s}$ is nondecreasing and that $\hat{s}(t)/t$ converges to 0 as t decreases to 0.

Suppose for some $t_2 > t_1 > 0$ such that $M(t_2) > 0$, $0 \leq \hat{s}(t_2) < \hat{s}(t_1)$. Then, $\hat{s}(t_2)/t_2 < \hat{s}(t_1)/t_1$ and $\epsilon/t_2 < \epsilon/t_1$ for any $\epsilon > 0$. Since yf_y is convex and f is strictly convex in y, and $M(t)/(m(t)t)$ is decreasing in t, for a sufficiently small $\epsilon > 0$,

$$
\begin{aligned}
&f\left(\frac{\hat{s}(t_2)+\epsilon}{t_2}\right) - f\left(\frac{\hat{s}(t_2)}{t_2}\right) + \left(\frac{\hat{s}(t_2)+\epsilon}{t_2} f_y\left(\frac{\hat{s}(t_2)+\epsilon}{t_2}\right) - \frac{\hat{s}(t_2)}{t_2} f_y\left(\frac{\hat{s}(t_2)}{t_2}\right)\right)\frac{M(t_2)}{m(t_2)t_2}\\
&< f\left(\frac{\hat{s}(t_1)}{t_1}\right) - f\left(\frac{\hat{s}(t_1)-\epsilon}{t_1}\right) + \left(\frac{\hat{s}(t_1)}{t_1} f_y\left(\frac{\hat{s}(t_1)}{t_1}\right) - \frac{\hat{s}(t_1)-\epsilon}{t_1} f_y\left(\frac{\hat{s}(t_1)-\epsilon}{t_1}\right)\right)\frac{M(t_1)}{m(t_1)t_1}.
\end{aligned}
$$

On the other hand, since $\hat{s}(t_1)$ is the smallest maximizer of $g(t_1,\cdot)$, $g(t_1,\hat{s}(t_1)) > g(t_1,\hat{s}(t_1)-\epsilon)$, which implies that

$$
f\left(\frac{\hat{s}(t_1)}{t_1}\right) - f\left(\frac{\hat{s}(t_1)-\epsilon}{t_1}\right) + \left(\frac{\hat{s}(t_1)}{t_1} f_y\left(\frac{\hat{s}(t_1)}{t_1}\right) - \frac{\hat{s}(t_1)-\epsilon}{t_1} f_y\left(\frac{\hat{s}(t_1)+\epsilon}{t_1}\right)\right)\frac{M(t_1)}{m(t_1)t_1} < \epsilon,
$$

and $g(t_2,\hat{s}(t_2)) - g(t_2,\hat{s}(t_2)+\epsilon)$

$$
\begin{aligned}
&= -\epsilon + f\left(\frac{\hat{s}(t_2)+\epsilon}{t_2}\right) - f\left(\frac{\hat{s}(t_2)}{t_2}\right) + \left(\frac{\hat{s}(t_2)+\epsilon}{t_2} f_y\left(\frac{\hat{s}(t_2)+\epsilon}{t_2}\right) - \frac{\hat{s}(t_2)}{t_2} f_y\left(\frac{\hat{s}(t_2)}{t_2}\right)\right)\frac{M(t_2)}{m(t_2)t_2}\\
&< -\epsilon + f\left(\frac{\hat{s}(t_1)}{t_1}\right) - f\left(\frac{\hat{s}(t_1)-\epsilon}{t_1}\right) + \left(\frac{\hat{s}(t_1)}{t_1} f_y\left(\frac{\hat{s}(t_1)}{t_1}\right) - \frac{\hat{s}(t_1)-\epsilon}{t_1} f_y\left(\frac{\hat{s}(t_1)-\epsilon}{t_1}\right)\right)\frac{M(t_1)}{m(t_1)t_1},
\end{aligned}
$$

which is less than 0, contradicting the optimality of $\hat{s}(t_2)$, and therefore $\hat{s}(t_2) \geq \hat{s}(t_1)$.

Furthermore, similar to the proof of Lemma A.5, $\hat{s}(t) \to 0$ as $t \to 0$. Otherwise if $\hat{s}(t) \geq \epsilon > 0$ for every t, then $g(t,\hat{s}(t))$ decreases to $-\infty$, which contradicts that $\hat{s}(t)$ is the maximizer for $g(t,\cdot)$ with sufficiently small t. Then suppose $\hat{y}(t) = \hat{s}(t)/t \geq \epsilon > 0$ for every t, since $M(t)/(m(t)t)$ decreases in t, and f is strictly increasing and convex;

$$
\begin{aligned}
\lim_{t\to 0} g(t,\hat{s}(t)) &= \lim_{t\to 0}\left((\hat{s}(t) - f(\hat{y}(t)))m(t) - M(t)\frac{\hat{y}(t)}{t} f_y(\hat{y}(t))\right)\\
&\leq \lim_{t\to 0}\left((\hat{s}(t) - f(\epsilon))m(t) - M(t)\frac{\epsilon}{t} f_y(\epsilon)\right) < -f(\epsilon)m(0) < -f(0)m(0).
\end{aligned}
$$

On the other hand, $\hat{s}(t)$ being the smallest maximizer of $g(t,\cdot)$ on $[0,\infty]$ implies that $\lim_{t\to 0} g(t,\hat{s}(t)) \geq \lim_{t\to 0} g(t,0) = \lim_{t\to 0} -f(0)\,m(t) = -f(0)m(0)$. This contradiction implies that $\hat{y}(t)$ converges to 0.

(ii) Since f is twice differentiable, g is differentiable in s. Then $g(t,s)$ being concave in s implies that for any $t > 0$, if $g_s(t,0) > 0$, then all maximizers of $g(t,\cdot)$ are solutions to $g_s = \left(1 - \frac{1}{t} f_y\left(\frac{s}{t}\right)\right)m(t) - \left(\frac{s}{t} f_{yy}\left(\frac{s}{t}\right) + f_y\left(\frac{s}{t}\right)\right)\frac{M(t)}{t^2} = 0$. By the continuity of g, $\hat{s}(t) = \inf\{s : g_s(t,s) = 0\}$ is the smallest maximizer.

If $g_s(t,0) \leq 0$, then for any $s \geq 0$, $g_s(t,s) \leq 0$, and $\hat{s}(t) = 0$ maximizes $g(t,\cdot)$. According to part (i), $\hat{s}(t)$ is the solution to (3.3). ∎

Proof of Lemma 6.1. Proposition 3.1 implies that, without reservation utility, for any $c \in C$, the agents only choose to produce where c coincides with its nondecreasing envelope c^*; hence their behavior under the constraint on reservation utility R is also the same under these two contracts. Similarly, with $\tilde{c} = \min(c,x)$, agents who choose to produce at x where $c = \tilde{c}$ make the same choice under the latter without the constraint and thus behave in the same way with the constraint. Otherwise they may choose differently under c and $\tilde{c}$ under the constraint, with the former providing nonpositive profits and the latter nonnegative profits. Thus $\tilde{c}$ generates larger profits. ∎

LEMMA A.9. *Let $c \in \tilde{C}$. The corresponding utility u^c is right-continuous and nondecreasing in t and is strictly increasing if $y^c(t) > 0$.*

Proof. Suppose that the utility from the optimal effort satisfies $u^c(t_2) < u^c(t_1)$ for $t_1 < t_2$. Assumption 2.3 implies that $c(t_2 y^c(t_2)) - f(t_2, y^c(t_2)) = u^c(t_2) < u^c(t_1) = c(t_1 y^c(t_1)) - f(t_1, y^c(t_1)) \leq c(t_1 y^c(t_1)) - f(t_2, t_1 y^c(t_1)/t_2)$, which contradicts the optimality of $s(t_2)$ for agents with skill t_2. Thus, u^c is nondecreasing in t for every $c \in \tilde{C}$. u^c is strictly increasing at t where $y^c(t) > 0$, following a similar argument.

To establish right-continuity, suppose by contradiction that u^c is not right-continuous at some t. Since c is right-continuous in revenue, and f is continuous in both arguments, $s(t) < \inf_{v>t} s(v) = \bar{s}$. Let $\inf_{v>t} u^c(v) - u^c(t) = \epsilon$. Then by the continuity of f, there exists t' greater than and sufficiently close to t, such that $c(\bar{s}) - f(t', \bar{s}/t') - (c(\bar{s}) - f(t, \bar{s}/t)) < \epsilon/2$. On the other hand, because $s(v)$ converges to $\bar{s}$ for $v \downarrow t$, t' can also be chosen to be sufficiently close to t, such that $c(\bar{s}) - f(t', \bar{s}/t') > u^c(t') - \epsilon/2$. Thus $c(\bar{s}) - f(t, \bar{s}/t) > u^c(t') - \epsilon > u^c(t)$, which contradicts the optimality of $y^c(t) = s(t)/t$ for agents with skill t. ∎

Proof of Lemma 6.2. For any contract $c \in \tilde{C}$ and its unconstrained revenue $s \in S$, the agents' utility u^c is increasing and right-continuous by Lemma A.9. Thus, with the constraint that the reservation utility R is a two-step function, the constrained revenue s^R satisfies

$$s^R(t) = \begin{cases} 0 & \text{if } t < \underline{t}_0, \\ s(t) & \text{if } \underline{t}_0 \leq t < t_1, \\ 0 & \text{if } t_1 \leq t < \underline{t}_1, \\ s(t) & \text{if } \underline{t}_1 \leq t, \end{cases}$$

where $\underline{t}_0 = \min\{t : u^c(t) \geq R_0\}$ and $\underline{t}_1 = \min\{t : u^c(t) \geq R_1\}$. The minimum exists by the right-continuity of u^c. ∎

Proof of Lemma 6.5. (i) Let $g_1(\underline{t}_1) > R_1$, and denote the unconstrained revenue induced by c and $c^{s^R,R}$ as s and $\tilde{s}$, respectively. Then Lemma A.10 implies that s and $\tilde{s}$ coincide at $t \in [\underline{t}_0, t_1) \cup [\underline{t}_1, \infty)$. Since $g_1(\underline{t}_1) > R_1$, $t_1^* = \underline{t}_1$, and by the construction of $c^{s^R,R}$ and Proposition 3.1(ii), $\tilde{s}(t) = \bar{s}_0$ for $t_1 \leq t < \underline{t}_1$. Similarly, $\bar{s}_0 \leq s(t) \leq s^R(\underline{t}_1)$ for $t_1 \leq t < \underline{t}_1$. Since by construction $u^{c^{s^R,R}}(\underline{t}_0) = R_0$, $c(s^R(\underline{t}_0)) \geq c^{s^R,R}(s^R(\underline{t}_0))$, and Lemma A.6 and the continuity of $c^{s^R,R}$ imply that $c(\bar{s}_0) \geq c^{s^R,R}(\bar{s}_0)$. Thus, from the optimality of $s(t)$, $u^c(t) \geq c(\bar{s}_0) - f(t, \bar{s}_0/t) \geq c^{s^R,R}(\bar{s}_0) - f(t, \bar{s}_0/t) = u^{c^{s^R,R}}(t)$. Since in this case $u^{c^{s^R,R}}(\underline{t}_1) > R_1$, and under $c^{s^R,R}$ $\underline{t}_1$-agents are indifferent between $\bar{s}_0$ and $s^R(\underline{t}_1)$, there exists $t \in [t_1, \underline{t}_1)$ and sufficiently close to $\underline{t}_1$ such that $u^{c^{s^R,R}}(t) > R_1$. Then $u^c(t) > R_1$, which contradicts that s^R is the constrained revenue under c.

(ii) Similar to (i), under $c^{s^R,R}$, $\tilde{s}(t) = \bar{s}_0$ for every $t \geq t_1$, and $u^c(t) \geq u^{c^{s^R,R}}(t) = c^{s^R,R}(\bar{s}_0) - f(t, \bar{s}_0/t)$. Since $\underline{t}_1 = \infty$, $u^c(t) < R_1$ for every $t \geq t_1$, which implies that $c^{s^R,R}(\bar{s}_0) - f(t, \bar{s}_0/t) < R_1$, and since $f(t, \bar{s}_0/t)$ is decreasing in t, it follows that $c^{s^R,R}(\bar{s}_0) - \lim_{t\to\infty} f(t, \bar{s}_0/t) \leq R_1$. ∎

Proof of Proposition 6.6. (i) Lemma A.10 implies that $c^{s^R,R}$ induces the same revenue s^R at $t \in [0, t_1) \cup [\underline{t}_1, \infty)$, without the participation constraint. It suffices to show that $\chi(c^{s^R,R}) = [\underline{t}_0, t_1) \cup [\underline{t}_1, \infty)$.

In Case (i) of Definition 6.4, if $\underline{t}_1 < \infty$, then $u^{c^{s^R,R}}(\underline{t}_1) = R_1$, and Lemma A.9 implies that $\chi(c^{s^R,R}) = [\underline{t}_1, \infty)$. Thus the constrained revenue under $c^{s^R,R}$ coincides with s^R. Similarly, in Case (ii), $u^{c^{s^R,R}}(\underline{t}_0) = R_0$, and according to Lemma 6.5(ii), $u^{c^{s^R,R}}(t) < R_1$ for $t \geq t_1$. Thus $\chi(c^{s^R,R}) = [\underline{t}_0, t_1)$. For Case (iv), by definition $u^{c^{s^R,R}}(\underline{t}_1) = R_1$. On the other hand,

$$u^{c^{s^R,R}}(\underline{t}_0) = d + \int_0^{s^R(\underline{t}_0)} \frac{1}{\underline{t}_0} f_y\left(\underline{t}_0, \frac{z}{\underline{t}_0}\right) dz - f\left(\underline{t}_0, \frac{s^R(\underline{t}_0)}{\underline{t}_0}\right)$$

$$= R_1 - \int_0^{s^R(\underline{t}_1)} c_x^{s^R,R}(z) dz + f\left(\underline{t}_1, \frac{s^R(\underline{t}_1)}{\underline{t}_1}\right) + \int_0^{s^R(\underline{t}_0)} \frac{1}{\underline{t}_0} f_y\left(\underline{t}_0, \frac{z}{\underline{t}_0}\right) dz - f\left(\underline{t}_0, \frac{s^R(\underline{t}_0)}{\underline{t}_0}\right)$$

$$> \int_0^{s^R(\underline{t}_0)} \frac{1}{t_0^*} f_y\left(t_0^*, \frac{z}{t_0^*}\right) dz + \int_{s^R(\underline{t}_0)}^{\bar{s}_0} \frac{1}{t_z^R} f_y\left(t_z^R, \frac{z}{t_z^R}\right) dz + \int_{\bar{s}_0}^{s^R(\underline{t}_1)} \frac{1}{t_1} f_y\left(t_1, \frac{z}{t_1}\right) dz$$

$$- \int_{s^R(\underline{t}_0)}^{s^R(\underline{t}_1)} c_x^{s^R,R}(z) dz - f\left(\underline{t}_0, \frac{s^R(\underline{t}_0)}{\underline{t}_0}\right) = \int_0^{s^R(\underline{t}_0)} \frac{1}{t_0^*} f_y\left(t_0^*, \frac{z}{t_0^*}\right) dz - f\left(\underline{t}_0, \frac{s^R(\underline{t}_0)}{\underline{t}_0}\right),$$

which equals R_0, and based on the marginal compensation, the optimal effort for agents with skill $t < \underline{t}_0$ is 0. Thus, similar to Cases (ii) and (iii), $\chi(c^{s^R,R}) = [\underline{t}_0, t_1) \cup [\underline{t}_1, \infty)$. For Case (iii), if $\underline{t}_1 > t_1$, $u^{c^{s^R,R}}(\underline{t}_1) = R_1$, then the argument is the same as above, and the constrained revenue under $c^{s^R,R}$ coincides with s^R. Otherwise $\underline{t}_1 = t_1$ and $u^{c^{s^R,R}}(\underline{t}_1) \geq R_1$. Thus $\chi(c^{s^R,R}) = [\underline{t}_0, t_1) \cup [\underline{t}_1, \infty)$.

For the comparison between $P^R(c)$ and $P^R(c^{s^R,R})$, it suffices to show that for every $t \in [\underline{t}_0, t_1) \cup [\underline{t}_1, \infty)$, $c(s^R(t)) \geq c^{s^R,R}(s^R(t))$. In Case (i) of Definition 6.4, since $u^{c^{s^R,R}}(\underline{t}_1) = R_1$, the constraint implies that $c(s^R(\underline{t}_1)) \geq c^{s^R,R}(s^R(\underline{t}_1))$. Lemma A.6 implies that for any $t \in [\underline{t}_1, \infty)$, $c(s^R(t)) \geq c^{s^R,R}(s^R(t))$. Case (ii) follows from a similar argument, replacing $\underline{t}_1$ with $\underline{t}_0$.

In Case (iii), since $u^{c^{s^R,R}}(\underline{t}_0) = R_0$, $\underline{t}_0 \leq t < t_1$ follows from arguments similar to those above. If $\underline{t}_1 = t_1$, the continuity of $c^{s^R,R}$ and Lemma A.6 imply that $c^{s^R,R}(s^R(\underline{t}_1)) - c^{s^R,R}(s^R(\underline{t}_0)) \geq c(s^R(\underline{t}_1)) - c(s^R(\underline{t}_0))$, whence $c^{s^R,R}(s^R(\underline{t}_1)) \geq c(s^R(\underline{t}_1))$. If $\underline{t}_1 > t_1$, the definition of $c^{s^R,R}$ and Lemma 6.5 imply that $u^{c^{s^R,R}}(\underline{t}_1) = R_1$, and thus $c(s^R(\underline{t}_1)) \geq c^{s^R,R}(s^R(\underline{t}_1))$. Then Lemma A.6 implies that for every $t \geq \underline{t}_1$, $c(s^R(t)) \geq c^{s^R,R}(s^R(t))$.

In Case (iv), since $u^{c^{s^R,R}}(\underline{t}_1) = R_1$, the argument for $t \geq \underline{t}_1$ is the same as for $\underline{t}_0 \leq t < t_1$ in Case (iii). For $\underline{t}_0 \leq t < t_1$, notice that $c(s^R(\underline{t}_1)) \geq c^{s^R,R}(s^R(\underline{t}_1))$. Furthermore, $c(s^R(\underline{t}_1)) - c(\bar{s}_0) \leq f(t_1, \frac{s^R(\underline{t}_1)}{t_1}) - f(t_1, \frac{\bar{s}_0}{t_1}) = c^{s^R,R}(s^R(\underline{t}_1)) - c^{s^R,R}(\bar{s}_0)$; otherwise there would exist some t sufficiently close to t_1, such that $c(s^R(\underline{t}_1)) - c(s^R(t)) > f(t, \frac{s^R(\underline{t}_1)}{t}) - f(t, \frac{s^R(t)}{t})$, which contradicts the optimality of $s^R(t)$ under c. Thus $c(\bar{s}_0) \geq c^{s^R,R}(\bar{s}_0)$. Finally, c must be left-continuous at $\bar{s}_0$; otherwise there would exist t less than and sufficiently close to t_1, such that the t-agent prefers to produce $\bar{s}_0$ to $s^R(t)$ due to the jump in c, which contradicts the optimality of $s^R(t)$. Then Lemma A.11 implies that $c(s^R(t)) \geq c^{s^R,R}(s^R(t))$ for every $\underline{t}_0 \leq t \leq t_1$.

(ii) From (i), it is sufficient for the principal to focus on $c^{s^R,R}$, where s^R is the constrained revenue generated by some $c \in \tilde{C}$. According to Lemma 6.5, such s^Rs must satisfy the first two properties in the definition of $\hat{S}_2$. Also, Case (iv) in

Definition 6.4 means that in order to induce s^R, the most economic contract for the principal is $c^{s^R,R} \notin \tilde{C}$, and thus is suboptimal. Therefore, the principal can focus on $\hat{S}_2$, which excludes this case, and $\max_{c\in C} P^R(c) = \max_{s^R\in\hat{S}_2}\int_{[\underline{t}_0,t_1)\cup[\underline{t}_1,\infty)}(s^R(t) - c^{s^R,R}(s^R(t)))m(t)dt$.

For $s^R \in \hat{S}_2$, Lemma A.8 implies that for every $t \in (\underline{t}_0, t_1) \cup (\underline{t}_1, \infty)$, $\frac{du^{c^s}(t)}{dt} = \frac{s^R(t)}{t^2} f_y(t, \frac{s^R(t)}{t}) - f_t(t, \frac{s^R(t)}{t})$. Since $u^{c^{s^R,R}}(\underline{t}_0) = R_0$, for $t \in [\underline{t}_0, t_1)$, $c^{s^R,R}(s^R(t)) = f(t, \frac{s^R(t)}{t}) + R_0 + \int_{\underline{t}_0}^t (\frac{s^R(v)}{v^2} f_y(v, \frac{s^R(v)}{v}) - f_t(v, \frac{s^R(v)}{v}))dv$. If $\underline{t}_1 > t_1$, $u^{c^{s^R,R}}(\underline{t}_1) = R_1$, and if $\underline{t}_1 = t_1$, $u^{c^{s^R,R}}(\underline{t}_1) = R_0 + \int_{\underline{t}_0}^{t_1} (\frac{s^R(v)}{v^2} f_y(v, \frac{s^R(v)}{v}) - f_t(v, \frac{s^R(v)}{v}))dv$. Thus for $t \in [\underline{t}_1, \infty)$, $c^{s^R,R}(s^R(t)) = f(t, \frac{s^R(t)}{t}) + \int_{\underline{t}_1}^t (\frac{s^R(v)}{v^2} f_y(v, \frac{s^R(v)}{v}) - f_t(v, \frac{s^R(v)}{v}))dv + \tilde{R}_1$,

$$\begin{aligned}&\int_{[\underline{t}_0,t_1]\cup[\underline{t}_1,\infty)} \left(s^R(t) - c^{s^R,R}(s^R(t))\right) m(t)dt \\ &\quad= \int_{\underline{t}_0}^{t_1} h(\underline{t}_0, t)m(t)dt + \int_{\underline{t}_1}^{\infty} h(\underline{t}_1, t)m(t)dt - R_0(M(\underline{t}_0) - M(t_1)) - \tilde{R}_1 M(\underline{t}_1). \quad \square\end{aligned}$$

LEMMA A.10. *Up to countably many points, $c^{s^R,R}$ induces the same revenue as s^R at $t \in [\underline{t}_0, t_1) \cup [\underline{t}_1, \infty)$ and $M(t) > 0$, without participation constraint.*

Proof. The proof is similar to that of Proposition 3.3, by comparing the marginal compensation and marginal disutility, and we omit the details here. □

LEMMA A.11. *If a contract $c \in \tilde{C}$ induces the constrained revenue s^R, then for any $\underline{t}_0 \le \underline{t} < \bar{t} < t_1$, $c(s^R(\bar{t})) - c(s^R(\underline{t})) = c^{s^R,R}(s^R(\bar{t})) - c^{s^R,R}(s^R(\underline{t}))$.*

Proof. If $c_x^{s^R,R}$ is not integrable, Lemma A.6 implies that c also explodes, which contradicts that $c \in \tilde{C}$. Thus, without loss of generality, in the following assume that $c_x^{s^R,R}$ is integrable. Also, without loss of generality, assume $\underline{t} > \underline{t}_0$, and the conclusion for $\underline{t}_0$ follows from the right-continuity of c.

Let the collection of all discontinuity points of s^R between $\underline{t}$ and $\bar{t}$ be $T = \{t_1, \ldots, t_K\}$, where K could be ∞. For each $t_i \in T$, let $\bar{x}_i = \inf\{s^R(t), t > t_i\}$ and $\underline{x}_i = \sup\{s^R(t), t < t_i\}$, so that $s^R(t_i) \in D_i = [\underline{x}_i, \bar{x}_i]$. Let $D = \bigcup_{i=1}^K D_i$. The proof of Lemma A.6 shows that D_i's only intersect at their end points, and $c(\bar{x}_i) - c(\underline{x}_i) \ge \int_{\underline{x}_i}^{\bar{x}_i} \frac{1}{t_i} f_y(t_i, \frac{z}{t_i})dz = c^{s^R,R}(\bar{x}_i) - c^{s^R,R}(\underline{x}_i)$ for every i. On the other hand, if $c(\bar{x}_i) - c(\underline{x}_i) > \int_{\underline{x}_i}^{\bar{x}_i} \frac{1}{t_i} f_y(t_i, \frac{u}{t_i})du$, since f is continuous, there must exist $t < t_i$, so that $s^R(t)$ is sufficiently close to $\underline{x}_i$, and $c(\bar{x}_i) - c(s^R(t)) > \int_{s^R(t)}^{\bar{x}_i} \frac{1}{t} f_y\left(t, \frac{u}{t}\right) du$, which contradicts the optimality of $s^R(t)$ and implies that $c(\bar{x}_i) - c(\underline{x}_i) = \int_{\underline{x}_i}^{\bar{x}_i} c_x^{s^R,R}(u)du$.

Furthermore, for every n, rearrange $t_1, \ldots, t_n$ in ascending order, relabeling them as $t_{n_1} < \cdots < t_{n_n}$. Let $\bar{x}_{n_0} = s^R(\underline{t})$ and $\underline{x}_{n_{n+1}} = s^R(\bar{t})$; then $c(s^R(\bar{t})) - c(s^R(\underline{t})) = \lim_{n\to\infty} \sum_{i=1}^{n+1} \left((c(\underline{x}_{n_i}) - c(\bar{x}_{n_{i-1}})) + (c(\bar{x}_{n_i}) - c(\underline{x}_{n_i}))\right)$.

Define a function $\tilde{c}$ on $[s^R(\underline{t}), s^R(\bar{t})]$, with $\tilde{c}(x) = c(x)$ for every $x \in [s^R(\underline{t}), s^R(\bar{t})] \setminus \cup_{i=1}^{\infty}[\underline{x}_i, \bar{x}_i)$. For any $x \in [\underline{x}_i, \bar{x}_i)$, for some i, let $\tilde{c}(x) = \tilde{c}(\bar{x}_i) - \int_x^{\bar{x}_i} \frac{1}{t_i} f_y(t_i, \frac{z}{t_i})dz$, so that $\tilde{c}(\underline{x}_i) = c(\underline{x}_i)$ and $\tilde{c}$ is nondecreasing. Thus, $\tilde{c}(s^R(\bar{t})) - \tilde{c}(s^R(\underline{t})) = c(s^R(\bar{t})) - c(s^R(\underline{t})) = \lim_{n\to\infty} \sum_{i=1}^{n+1} \left((\tilde{c}(\underline{x}_{n_i}) - \tilde{c}(\bar{x}_{n_{i-1}})) + (\tilde{c}(\bar{x}_{n_i}) - \tilde{c}(\underline{x}_{n_i}))\right)$. Focusing on $\tilde{c}$, since it is nondecreasing, it is differentiable almost everywhere. For each $z \in (\underline{x}_i, \bar{x}_i)$, for some i, $\tilde{c}_x(z) = \frac{1}{t_i} f_y(t_i, \frac{z}{t_i}) = c_x^{s^R,R}(z)$. For $z \in [s^R(\underline{t}), s^R(\bar{t})] \setminus D$, $z = s^R(t)$ for some t, and there are two cases to discuss.

If s^R is strictly increasing at t, then for any $h>0$, if $z-h=s^R(t')$ for some t', then the optimality of $s^R(t)$ under c implies that $\frac{\tilde c(z)-\tilde c(z-h)}{h}=\frac{c(z)-c(z-h)}{h}\geq \frac{1}{t}\frac{f\left(t,\frac{z}{t}\right)-f\left(t,\frac{z-h}{t}\right)}{\frac{h}{t}}$. Otherwise $z-h\in D_i$ for some i, and the definition of $\tilde c$ implies

$$\frac{\tilde c(z)-\tilde c(z-h)}{h}=\frac{c(z)-c(\bar x_i)+\tilde c(\bar x_i)-\tilde c(z-h)}{h}$$
$$\geq \frac{f(t,\frac{z}{t})-f(t,\frac{\bar x_i}{t})+f(t_i,\frac{\bar x_i}{t_i})-f(t_i,\frac{z-h}{t_i})}{h}\geq \frac{1}{t}\frac{f(t,\frac{z}{t})-f(t,\frac{z-h}{t})}{h/t},$$

where the second inequality follows from the fact that f_y is nonincreasing in t. Let h go to 0, and it follows that $\tilde c_x(z-)\geq \frac{1}{t}f_y\left(t,\frac{z}{t}\right)$.

For the upper bound, if $z-h=s^R(t')$ for some t', then from the optimality of $z-h$ for t'-agents, $\frac{c'(z)-c'(z-h)}{h}\leq \frac{1}{t'}\frac{f\left(t',z/t'\right)-f\left(t',(z-h)/t'\right)}{h/t'}$. If $z-h\in D_i$, then

$$\frac{\tilde c(z)-\tilde c(z-h)}{h}=\frac{c(z)-c(s^R(t_i))+\tilde c(s^R(t_i))-\tilde c(z-h)}{h}$$
$$\leq \frac{f\left(t_i,\frac{z}{t_i}\right)-f\left(t_i,\frac{s^R(t_i)}{t_i}\right)+f\left(t_i,\frac{s^R(t_i)}{t_i}\right)-f\left(t_i,\frac{z-h}{t_i}\right)}{h}=\frac{f\left(t_i,\frac{z}{t_i}\right)-f\left(t_i,\frac{z-h}{t_i}\right)}{h},$$

where the inequality follows from the optimality of $s^R(t_i)$ under c, and from that both $s^R(t_i)$ and $z-h\in D_i$. Since, as h decreases to 0, both t' and t_i converge to t, $\tilde c_x(z-)\leq \frac{1}{t}f_y\left(t,\frac{z}{t}\right)$, and thus equality holds. A similar argument shows that $\tilde c_x(z+)=\frac{1}{t}f_y\left(t,\frac{z}{t}\right)$, and therefore $\tilde c$ is differentiable at z, with $\tilde c_x(z)=c_x^{s^R,R}(z)$.

Finally, if $z\notin D$ and s^R is not strictly increasing at z, then $t^*>t_*$, where $t^*=\inf\{t: s^R(t)>z\}$ and $t_*=\sup\{t: s^R(t)<z\}$. Following the same argument as above, for every $t_*<t<t^*$, $\frac{1}{t}f_y\left(t,\frac{z}{t}\right)\leq \tilde c_x(z-)\leq \frac{1}{t_*}f_y(t_*,\frac{z}{t_*})$, which implies that $\tilde c_x(z-)=\frac{1}{t_*}f_y(t_*,\frac{z}{t_*})$. Similarly, $\tilde c_x(z+)=\frac{1}{t^*}f_y\left(t^*,\frac{z}{t^*}\right)$. Thus, $\tilde c$ is not differentiable at z, and the latter belongs to a set of Lebesgue measure 0.

In terms of the continuity of $\tilde c$, it is continuous at each $\bar x_i$ by construction, because c is right-continuous, and also at every $z\in[s^R(\underline{t}),s^R(\bar t)]\setminus D$, because its right and left derivatives exist and are finite. Furthermore, for each $\underline{x}_i$, it suffices to check the case that $\underline{x}_i\neq \bar x_j$ for any $j\neq i$. By construction, $\tilde c$ is right-continuous at $\underline{x}_i$. Suppose there exists some i such that $\lim_{x\uparrow \underline{x}_i}\tilde c(x)<\tilde c(\underline{x}_i)$; then $\lim_{x\uparrow\underline{x}_i}c(x)<c(\underline{x}_i)=\tilde c(\underline{x}_i)$. Choose $\tilde x$ sufficiently close to $\underline{x}_i$, so that $f(t_i,\frac{\underline{x}_i}{t_i})-f(t_i,\frac{\tilde x}{t_i})<(c(\underline{x}_i)-\lim_{x\uparrow\underline{x}_i}c(x))/2$. Since f is continuous, there exists $\tilde t<t_i$ such that for every $\tilde t\leq t<t_i$ and every $\tilde x\leq u<\underline{x}_i$, $f\left(t,\underline{x}_i/t\right)-f\left(t,u/t\right)<c(\underline{x}_i)-\lim_{x\uparrow\underline{x}_i}c(x)\leq c(\underline{x}_i)-c(u)$, which implies that $s^R(t)\notin[\tilde u,\underline{x}_i)$. This is a contradiction to the definition of $\underline{x}_i$.

To summarize, $\tilde c$ is continuous on $[s^R(\underline{t}),s^R(\bar t)]$, and $\tilde c_x=c_x^{s^R,R}$ almost everywhere, which is integrable. From Theorem 6.27 in [12], $c(s^R(\bar t))-c(s^R(\underline{t}))=\tilde c(s^R(\bar t))-\tilde c(s^R(\underline{t}))=\int_{s^R(\underline{t})}^{s^R(\bar t)}c_x^{s^R,R}(z)dz=c^{s^R,R}(s^R(\bar t))-c^{s^R,R}(s^R(\underline{t}))$. ∎

Acknowledgments. For helpful comments, we thank Jakša Cvitanic, Kasper Larsen, and seminar participants at Rutgers University, Bachelier Colloquium in Metabief, Tianfu workshop in Financial Mathematics at SWUFE, and the Mathematical Economics and Finance conference at the University of Manchester. We are especially grateful to two anonymous referees, whose stimulating observations led to significant improvements in the paper.

REFERENCES

[1] C. Alasseur, I. Ekeland, R. Élie, N. Hernández Santibáñez, and D. Possamaï, *An adverse selection approach to power pricing*, SIAM J. Control Optim., 58 (2020), pp. 686–713, https://doi.org/10.1137/19M1260578.

[2] D. P. Baron and B. Holmström, *The investment banking contract for new issues under asymmetric information: Delegation and the incentive problem*, J. Finance, 35 (1980), pp. 1115–1138.

[3] P. Bolton and M. Dewatripont, *Contract Theory*, MIT Press, Cambridge, MA, 2005.

[4] X. Cheng, C. Dale, and J. Liu, *Statistics and social network of YouTube videos*, in Proceedings of the 16th International Workshop on Quality of Service, IEEE, Washington, DC, 2008, pp. 229–238.

[5] J. Cvitanic and J. Hugonnier, *Optimal Fund Menus*, CEPR Discussion Paper DP13127, Centre for Economic Policy Research, London, UK, 2018.

[6] J. Cvitanic and J. Zhang, *Contract Theory in Continuous-Time Models*, Springer, Berlin, 2013.

[7] M. N. Darrough and N. M. Stoughton, *Moral hazard and adverse selection: The question of financial structure*, J. Finance, 41 (1986), pp. 501–513.

[8] P. Diamond and E. Saez, *The case for a progressive tax: From basic research to policy recommendations*, J. Econom. Perspectives, 25 (2011), pp. 165–90.

[9] O. El Euch, T. Mastrolia, M. Rosenbaum, and N. Touzi, *Optimal Make-Take Fees for Market Making Regulation*, Tech. report, 2018, https://doi.org/10.2139/ssrn.3174933.

[10] R. Elie, T. Mastrolia, and D. Possamaï, *A tale of a principal and many, many agents*, Math. Oper. Res., 44 (2019), pp. 377–766, https://doi.org/10.1287/moor.2018.0931.

[11] M. Golosov, A. Tsyvinski, and I. Werning, *New dynamic public finance: A user's guide*, in NBER Macroeconomics Annual 2006, NBER Macroenonomics Annual 21, MIT Press, Cambridge, MA, 2007, pp. 317–388.

[12] R. A. Gordon, *The Integrals of Lebesgue, Denjoy, Perron, and Henstock*, Grad. Stud. Math. 4, AMS, Providence, RI, 1994.

[13] R. Guesnerie and J.-J. Laffont, *A complete solution to a class of principal-agent problems with an application to the control of a self-managed firm*, J. Public Econom., 25 (1984), pp. 329–369.

[14] B. Jullien, *Participation constraints in adverse selection models*, J. Econom. Theory, 93 (2000), pp. 1–47.

[15] T. R. Lewis and D. E. Sappington, *Countervailing incentives in agency problems*, J. Econom. Theory, 49 (1989), pp. 294–313.

[16] S. Lollivier and J.-C. Rochet, *Bunching and second-order conditions: A note on optimal tax theory*, J. Econom. Theory, 31 (1983), pp. 392–400.

[17] G. Maggi and A. Rodriguez-Clare, *On countervailing incentives*, J. Econom. Theory, 66 (1995), pp. 238–263.

[18] N. G. Mankiw, M. Weinzierl, and D. Yagan, *Optimal taxation in theory and practice*, J. Econom. Perspectives, 23 (2009), pp. 147–74.

[19] J. A. Mirrlees, *An exploration in the theory of optimum income taxation*, Rev. Econom. Stud., 38 (1971), pp. 175–208.

[20] E. Saez, *Using elasticities to derive optimal income tax rates*, Rev. Econom. Stud., 68 (2001), pp. 205–229.

[21] J. A. Weymark, *Bunching properties of optimal nonlinear income taxes*, Soc. Choice Welf., 3 (1986), pp. 213–232.