DE GRUYTER The B.E. Journal of Theoretical Economics. 2022; 22(1): 233–266

Research Article

Alexander Zimper* and Mpoifeng Molefinyane

When is Knowledge Acquisition Socially Beneficial in the Laffont–Tirole Regulatory Framework?

https://doi.org/10.1515/bejte-2020-0069
Received April 27, 2020; accepted March 1, 2021

Abstract: The Laffont–Tirole regulator observes the accounting costs of a firm but she can neither observe its true cost-type nor its chosen effort level. This paper considers a Laffont–Tirole regulator who could employ an expert to obtain better, albeit not perfect, knowledge about the firm's true cost type. Both the welfare gains through superior allocations from better knowledge but also the knowledge acquisition costs increase in the 'marginal deadweight losses from taxes' parameter $\lambda \geq 0$. We derive a closed-form expression of the overall welfare benefits from knowledge acquisition as a function in λ. We characterize parameter conditions such that knowledge acquisition could improve social welfare in dependence on the value of λ. For this case we show that knowledge acquisition strictly increases social welfare if and only if λ falls into the interval (λ^*, ∞) whereby we present a sharp characterization of the critical threshold-value $\lambda^* \geq 0$. In other words, information acquisition through a regulator only increases welfare for economies with comparatively high deadweight losses from taxation whereas welfare is decreased whenever these deadweight losses are low.

Keywords: regulation, asymmetric information, knowledge

JEL Classification: D82, H21, L51

Article note: We are grateful to comments and suggestions from Stefano Comino, Alex Ludwig, Nicky Nicholls, and Eric Picard. We are especially grateful to the valuable comments and suggestions from two anonymous referees.

***Corresponding author: Alexander Zimper**, Department of Economics, University of Pretoria, Private Bag X20, Hatfield 0028, South Africa, E-mail: alexander.zimper@up.ac.za. https://orcid.org/0000-0002-3730-3086
Mpoifeng Molefinyane, Department of Economics, University of Pretoria, Private Bag X20, Hatfield 0028, South Africa, E-mail: mpoifeng.molefinyane@up.ac.za

1 Introduction

1.1 Motivation and Overview

In a series of influential articles Laffont and Tirole[1] have developed a theoretical framework for the regulation of a firm under the assumption that "Regulators face a double asymmetry of information, called adverse selection and moral hazard respectively" (Tirole 2015, p. 1670). Whereas the Laffont–Tirole regulator can observe the firm's accountancy costs she can neither observe the firm's true cost-type (adverse selection) nor its chosen effort level (moral hazard). Incentive compatible contracts have to ensure that cost-efficient firms do not mimic the accounting costs of less cost-efficient firms through the choice of inefficiently low effort levels. Within the Laffont–Tirole regulatory framework any benefits from better information strictly decrease in the marginal deadweight losses from taxes, denoted $\lambda \geq 0$, whereby these benefits become zero if there are no deadweight losses from taxes, i.e. if $\lambda = 0$. The reason for this peculiar feature of the Laffont–Tirole regulatory framework is the assumption of an utilitarian regulator who maximizes the overall social welfare, which includes the firm's utility. Whenever the regulator has to transfer money to the firm in order to incentivize her to reveal her true type in a second best contract, such transfers would only decrease the overall aggregate welfare if they are financed through taxes that come with deadweight losses. Without such deadweight losses from taxation publicly funded transfers required for incentive compatibility would simply redistribute aggregate welfare towards the firm. In other words, the primary issue of welfare losses from asymmetric information in the Laffont–Tirole regulatory framework is not this asymmetric information per se but rather the inefficiencies of a distortive tax system which makes incentive transfers from the regulator to the firm socially costly.

Because any information acquisition by the regulator must, eventually, also be paid for with taxes, the question about the optimal trade-off between benefits and costs of better information takes on the following specific form within the Laffont–Tirole regulatory framework:

Question: *For which values of the'marginal deadweight losses from taxes' parameter λ do the benefits of better information outweigh the associated social costs of information acquisition and vice versa?*

1 Compare, e.g. Laffont and Tirole (1986, 1990a, 1990b, 1991). This regulatory framework culminated in the Laffont and Tirole (1993) textbook which has since become "the bible in the field" (Salanié 2005, p. 47).

We address this research question through the construction of a stylized model of 'knowledge acquisition' as a special case of 'information acquisition'. General models of information acquisition would consider 'signals' that are used to update a prior distribution about the firm's possible types towards a more precise posterior distribution. Our notion of knowledge acquisition corresponds to the formal special case of information acquisition according to which Bayesian updating can rule out some cost-type(s) as impossible. In our opinion, better knowledge – rather than just more precise posteriors in general – describes the bulk of real-life expert advice. For instance, when people ask a medical doctor for advice about their symptoms they expect the medical expert to narrow down the possible causes for these symptoms. Arguably, 'experts' who cannot reduce the support of the prior distribution but rather declare that 'anything remains possible but with updated posteriors' would be soon out of business as they do not inspire much trust in non-experts. The reason is that any statistical ex post evaluation makes it so much harder to judge the quality of an expert than the simple observation whether the expert's opinion is verified or falsified.

More specifically, we assume that the Laffont–Tirole regulator of our model could employ an expert who possesses a knowledge-generating technology which is characterized by the following two features:[2]

Assumptions on knowledge-generating expert information.

1. *The technology strictly reduces the uncertainty about the firm's possible cost-types by excluding some cost-type(s) as impossible;*
2. *The technology is imperfect in the sense that it cannot reduce all uncertainty about the firm's true cost-type.*

To capture simultaneously both features of the expert's technology in the most parsimonious way, we distinguish between three possible cost-types $\beta_1 < \beta_2 < \beta_3$ of the firm. Without knowledge acquisition the regulator is completely uncertain about the firm's true cost-type whereby she resolves this uncertainty through the prior belief μ such that $\mu(\beta_i) > 0$ for all i. If the regulator acquires better knowledge through the expert, she becomes able to rule out exactly one cost-type as impossible whereby the prior μ is accordingly updated. For the sake of analytical rigor – as well as possible future extensions – we explicitly construct a state space such that better knowledge formally corresponds to a refined information partition of

2 In contrast to the *expert advice* models by Ottaviani and Sørensen (2006a, 2006b) – where an expert of different degrees of experience plays a strategic game to enhance his reputation – this truthful expert would simply share his knowledge with the regulator.

the state space. This information partition reveals in any given state of the world exactly two cost-types β_i and β_j with $i < j$ as possible whereas the cost-type β_k, $k \neq i$, j, is ruled out as impossible.

Based on our parsimonious model of knowledge acquisition, we derive a closed-form expression for the welfare gains from knowledge acquisition within the Laffont–Tirole regulatory framework. Denote by $E_\mu(W^{\text{Unc}})$ the expected welfare from Laffont–Tirole regulation for the situation in which the regulator decides to remain completely uncertain about the firm's possible cost-types. This expected welfare is pinned down by the second-best regulatory contract with the firm for the three cost-types $\{\beta_1, \beta_2, \beta_3\}$, which we explicitly derive in this paper. Next, denote by $E_\mu(W^{\text{Kno}})$ the expected welfare that the Laffont–Tirole regulator achieves when she decides to hire the expert in order to acquire better knowledge about the firm's cost-type. This expected welfare is pinned down as the expectation over the (expected) welfare from the three second-best regulatory contracts for all combinations of two possible costs types, i.e. for $\{\beta_1, \beta_2\}$, $\{\beta_1, \beta_3\}$, and $\{\beta_2, \beta_3\}$, respectively. Such optimal contracts for two cost-types are standard in the literature. The expected welfare gain from better knowledge is then formally defined as the difference between the expected welfare from these two scenarios

$$E_\mu(\Delta W) := E_\mu\left(W^{\text{Kno}}\right) - E_\mu\left(W^{\text{Unc}}\right). \tag{1}$$

To keep our analysis as simple as possible, we impose two technically convenient restrictions on the parameter space. Firstly, we assume that all states of the world are equally likely, implying $\mu(\beta_i) = \frac{1}{3}$ for all i. Secondly, we assume that the three possible cost-types come with a fixed difference

$$\Delta\beta = \beta_{i+1} - \beta_i \leq \frac{1}{4} \text{ for } i \in \{1, 2\}.$$

Under these simplifying assumptions, we derive the following closed-form expression for the expected welfare gain $E_\mu(\Delta W)$ from better knowledge as a function in the 'deadweight losses from taxes' parameter λ:

$$E_\mu(\Delta W)[\lambda] = \frac{1}{3}\lambda\Delta\beta\left(1 - \frac{\lambda}{1+\lambda}\Delta\beta\right). \tag{2}$$

Within the Laffont–Tirole regulatory framework, the acquisition of better knowledge must always result into some non-negative expected welfare gain because better knowledge helps to save on expected deadweight losses from taxes associated with incentive transfers to the firm. These savings are the greater the greater the deadweight losses from taxes. Beyond the general insight that expected welfare gains from knowledge acquisition increase in the deadweight losses from taxes, the expression (2) provides a specific functional form for this relationship. By (2), expected welfare gains are a strictly increasing, unbounded, and strictly concave function in the marginal deadweight losses from taxes.

We are now in the position to answer our research question about the values of the 'marginal deadweight losses from taxes' parameter for which the benefits from knowledge acquisition outweigh the corresponding costs and vice versa. The benefits of knowledge acquisition are given – under our modeling assumptions – as the closed-form expression (2) for expected welfare gains. To capture the costs of knowledge acquisitions, we assume that the regulator can hire the expert at the fixed fee $L \geq 0$. Because this fee has to be financed through taxes, the overall social costs from knowledge acquisition – including deadweight losses from taxes – are given as $(1+\lambda)L$. For a fixed value of the 'marginal deadweight losses from taxes' parameter λ, the overall expected welfare gain from the acquisition of knowledge is thus given as the difference

$$E_\mu(\Delta W)\,[\lambda] - (1+\lambda)L.$$

In other words, the social benefits from knowledge acquisition strictly outweigh the corresponding social costs if and only if

$$E_\mu(\Delta W) > (1+\lambda)L \quad \Leftrightarrow \quad \frac{1}{3}\lambda\Delta\beta\left(1-\frac{\lambda}{1+\lambda}\Delta\beta\right) > (1+\lambda)L. \tag{3}$$

Analyzing inequality (3) gives us – for our specific modeling assumptions – the following answer to our research question:

Answer to our Question:

(i) *If the expert fee for knowledge acquisition is too costly in the specific sense that*

$$L \geq \frac{1}{3}\left(\Delta\beta - (\Delta\beta)^2\right), \tag{4}$$

then knowledge acquisition is never socially beneficial irrespective of the value of the'marginal deadweight losses from taxes' parameter λ.

(ii) *If we have instead that*

$$L < \frac{1}{3}\left(\Delta\beta - (\Delta\beta)^2\right), \tag{5}$$

then knowledge acquisition is strictly socially beneficial for all values of the'marginal deadweight losses from taxes' parameter $\lambda \in (\lambda^*, \infty)$ *such that the critical threshold is given as*

$$\lambda^* = \frac{\sqrt{(\Delta\beta - 6L)^2 + 4(\,(\Delta\beta - (\Delta\beta)^2) - 3L)3L} - (\Delta\beta - 6L)}{2(\,(\Delta\beta - (\Delta\beta)^2) - 3L)} > 0. \tag{6}$$

1.2 Contribution and Relationship to the Literature

The main benefits of the Laffont–Tirole regulatory framework are its theoretical insights into the basic trade-offs for regulation under asymmetric information. According to Tirole (2015) this framework highlights "two broad principles":

> The first is obvious. Authorities should attempt at reducing the asymmetry of information: by collecting data of course [...]. The second principle is that one size does not fit all: one should let the regulated firm make use of its information. (p. 1671)

This paper takes a critical look at the first principle whereby we argue that the overall benefits of better information are less 'obvious' whenever information acquisition is costly. In the Laffont–Tirole regulatory framework both the benefits and the costs from knowledge acquisitions strictly increase in the value of the 'marginal deadweight losses from taxes' parameter λ. In general, both curves – of benefits and costs, respectively – might therefore never intersect or intersect arbitrarily many times. Under our modeling assumptions, we show the following:

- If the knowledge acquisition cost- and the cost-type-difference parameters – L and $\Delta\beta$, respectively – satisfy inequality (4), the regulator should abstain from knowledge acquisition and rather regulate under complete uncertainty;
- Else, there exists exactly one critical threshold value $\lambda^* > 0$ – as pinned down by Equation (6) – such that regulation under better knowledge is (strictly) superior to regulation under complete uncertainty if and only if the 'marginal deadweight losses from taxes' parameter satisfies $\lambda > \lambda^*$ (That is, if inequality (5) holds, the cost- and benefit curves intersect exactly once, namely at λ^*).

To bring our theoretical analysis closer to regulatory practice, one would need empirical estimates about the relevant parameters – L, $\Delta\beta$, and λ – in question. On the one hand, there exists a large literature in public finance which works with plausible values or/and estimates for deadweight losses from taxes, typically within discussions about possible benefits (or the lack thereof) of selected tax reforms.[3] To quote from Jacobs (2018):

> Many applied cost–benefit analyses multiply the cost of public projects with a measure for the marginal cost of public funds that is larger than one. As a result, public projects are less likely to pass a cost–benefit test. For example, Heckman et al. (2010) evaluate the Perry Preschool Program and add 50 cents per dollar spent to account for the deadweight costs of taxation. Many other examples can be given, but the message is clear: The marginal cost of public funds has a tremendous impact on how governments should evaluate the desirability of public policies. (p.884)

On the other hand, we are not aware about any plausible values or/and estimates for relevant parameters of knowledge acquisitions costs L and cost-type differences $\Delta\beta$. The problem from an empirical perspective is that the relevance of the Laffont–Tirole regulatory framework for regulatory practice appears to be rather limited. Compare, e.g. Rogerson (2003) who writes about the Laffont–Tirole regulatory framework:

> Two related problems with applying this theory in practice have been that the economic logic and the underlying mathematics involved in calculating the optimal menu are quite complex, and the principal must be able to specify the agent's entire disutility of effort function in order to calculate the optimal menu. As a result, the model has not been widely used in practice to either calculate actual incentive contracts or even to develop useful qualitative guidance about the nature of the optimal solution and how it is affected by various economic factors. (p. 919)

Whereas our analysis provides a better theoretical understanding of the welfare trade-offs implicit to the Laffont–Tirole regulatory framework, our model is, admittedly, rather far away from any real-life applications in regulatory practice. Nevertheless, we hope that our main insight – according to which information acquisition through a regulator is only socially beneficial for economies with sufficiently high deadweight losses from taxation – might serve as guidance for future research which is more applied in nature. For example, Laffont (2003, 2005) argues that optimal regulation for developing countries should be different from regulation in developed countries (also see Estache and Wren-Lewis 2009). As one important driver for such difference in regulation, Laffont identifies the inefficient fiscal system in developing countries. Translated into our model, greater fiscal inefficiency of a developing country naturally corresponds to a larger value of the

---

3 For early contributions, see Harberger (1962, 1964). For more recent literature that addresses empirical questions about marginal dead-weight losses from taxes, see, e.g. Heckman et al. (2010), Yagan (2015), Serrato and Zidar (2016), Blomquist and Simula (2019), and references therein. For a critical discussion, see Jacobs (2018) and references therein.

'marginal deadweight losses from taxes' parameter. According to our main insight, it would then be beneficial for regulator in developing rather than in developed countries to acquire better knowledge.

Our theoretical approach is also related to two different strands in the theoretical literature that look into the role of better information within a regulatory context.[4] The first strand includes the articles by Cremer and Khalil (1992), Crémer, Khalil, and Rochet (1998a, 1998b), and Szalay (2009). These authors take up the Baron and Myerson (1982) model about the regulation of a monopoly with an unknown cost-structure and ask when is it optimal for the regulator to provide the firm with incentives to acquire better information about its own cost-structure. The differences to this literature are that (i) in our model the regulator would acquire better information – through the expert – directly and (ii) that the regulator is utilitarian in that she also cares about the firm's welfare, which brings in the deadweight losses from taxation as social cost drivers.

The second strand – represented by Lewis (1996) and Magesan and Turner (2010) – introduces asymmetric information into economic policy models of regulation by Stigler (1971) and Becker (1985). For example, Magesan and Turner (2010) consider a situation where a regulator faces firms with different negative externality (i.e. pollution) and profitability characteristics that are unknown to the regulator. In their model the (constrained) regulator has to decide whether she rather acquires information about the social costs of pollution or of the profitability of the firm. This decision matters for the regulatory success because firms with higher profitability might be able to block any regulatory measures through legal processes. In contrast to our approach – where better information is only valuable through better 'contracts' with the regulated firm – information is valuable in Magesan and Turner (2010) because it increases the chances that the regulation of a firm gets actually implemented through the political process.

Although the drivers for the costs and benefits from better information in these two strands of models are quite different from the Laffont–Tirole regulatory framework, this literature shares with our paper the motivation to investigate the benefits of better information in connection with the associated costs of information acquisition.

The remainder of this paper is structured as follows. Section 2 recalls the Laffont–Tirole regulatory framework. Section 3 derives the optimal allocation under complete uncertainty. Section 4 provides the general (Blackwell) argument why expert information must always result in (weakly) superior allocations. Section 5 solves for the optimal allocation under the assumption that expert information generates better knowledge. In Section 6 we compare the expected

4 We are grateful to two anonymous referees for pointing us to these two literature strands.

welfare gain from knowledge acquisition with the corresponding acquisition costs. Section 7 concludes. Formal proofs are relegated to the Appendix.

2 The Laffont–Tirole Regulatory Framework

We consider a firm (e.g. a natural monopoly) that provides a service to the government which yields revenue $R >> 0$. Producing this service comes with the accounting costs

$$C = \beta - e \geq 0$$

where $\beta \geq 1$ denotes the firm's cost-type and $e \in [0, 1]$ denotes the firm's chosen effort level out of all physically available effort levels. The Laffont–Tirole regulator can observe the firm's accounting costs but she cannot observe the cost-components β and e, respectively. To keep the analysis simple, we consider three possible cost-types $\beta_1 < \beta_2 < \beta_3$ such that

$$\Delta\beta = \beta_{i+1} - \beta_i \text{ for } i \in \{1, 2\}.$$

The regulator compensates the firm by transferring the amount $t \geq 0$. If the firm is of cost-type β and chooses effort level e, the firm's overall utility is given as

$$U(t, e; \beta) = t - (\beta - e) - \varphi(e) \tag{7}$$

whereby the firm's disutility from effort $e \in [0, 1]$ is

$$\varphi(e) = \frac{1}{2} e^2.$$

To transfer the amount t to the firm, the regulator has to spend the amount $(1 + \lambda)t$ where the 'marginal deadweight losses from taxes' parameter $\lambda \in [0, \infty)$ measures the social costs from taxation. By convention, the government appropriates the revenue R generated by the firm so that the government's payoff becomes

$$\prod(t) = R - (1 + \lambda)t. \tag{8}$$

The Laffont–Tirole regulator is utilitarian in the sense that she cares about the aggregate welfare given as the sum of (7) and (8), i.e.

$$W(t, e; \beta) = \Pi(t) + U(t, e; \beta) = R - \lambda t - (\beta - e) - \varphi(e).$$

Observe that any transfer $t > 0$ to the firm can only impact negatively on the aggregate welfare in the Laffont–Tirole regulatory framework if the fiscal system is inefficient, i.e. if $\lambda > 0$.

3 Optimal Allocation Under Complete Uncertainty

The regulator resolves her uncertainty about the firm's true cost-type through the prior μ such that

$$\mu(\beta_i) > 0 \text{ for } i \in \{1, 2, 3\}.$$

The Laffont–Tirole regulator maximizes the expected aggregate welfare

$$E_\mu(W(\mathbf{t}, \mathbf{e})) = R - \sum_{i=1}^{3} (\lambda t_i + (\beta_i - e_i) + \varphi(e_i))\mu(\beta_i)$$

over allocations $(\mathbf{t}, \mathbf{e}) = (t_i, e_i)_{i=1,2,3}$ that can be implemented through a cost-type revealing contract $(\mathbf{t}, \mathbf{C} = \boldsymbol{\beta} - \mathbf{e})$. More precisely, the allocation $(\mathbf{t}, \mathbf{e})$ can be implemented if and only if it satisfies the following participation and incentive compatibility constraints.

The participation constraints for $(\mathbf{t}, \mathbf{e})$ are simply given as

$$U(t_i, e_i; \beta_i) \geq 0 \text{ for all } i.$$

To determine the relevant incentive compatibility constraints (=ICCs), note that cost-type i could mimic the contracted costs $C_j = \beta_j - e_j$ of cost-type j if and only if there exists some $e_i' \in [0, 1]$ satisfying[5]

$$(\beta_i - e_i') = C_j$$
$$\Leftrightarrow$$
$$e_i' = e_j - (\beta_j - \beta_i).$$

For any pair of cost-types $i \neq j$ that gives rise to such $e_i' \in [0, 1]$ the allocation $(\mathbf{t}, \mathbf{e})$ is thus incentive compatible if and only if

$$t_i - (\beta_i - e_i) - \varphi(e_i) \geq t_j - (\beta_i - e_i') - \varphi(e_i').$$

The following proposition (proved in the Appendix) characterizes the optimal allocation $(\mathbf{t}^*, \mathbf{e}^*)$ subject to parameter restrictions which ensure that $\mathbf{e}^*$ is an interior optimum pinned down by first-order conditions.

Proposition 1. *Suppose that the model parameters satisfy the following two inequalities*

$$\frac{\mu(\beta_2)}{\mu(\beta_1)} \geq \frac{\mu(\beta_3)}{\mu(\beta_1) + \mu(\beta_2)}, \tag{9}$$

5 In case that $e_j^* - (\beta_j - \beta_i) \notin [0, 1]$ there is no effort level $e_i' \in [0, 1]$ available to i through which i could mimic j. Consequently, no ICC is then needed to keep i from imitating j.

$$\frac{\mu(\beta_3) + \mu(\beta_3)\lambda}{\mu(\beta_3) + \lambda} \geq \Delta\beta. \tag{10}$$

Then the optimal allocation $(\mathbf{t}^*, \mathbf{e}^*)$ *that can be implemented through a contract under complete uncertainty is given as follows:*

$$\begin{aligned} e_1^* &= 1, \\ e_2^* &= 1 - \frac{\lambda}{1+\lambda} \frac{\mu(\beta_1)}{\mu(\beta_2)} \Delta\beta, \\ e_3^* &= 1 - \frac{\lambda}{1+\lambda} \frac{1-\mu(\beta_3)}{\mu(\beta_3)} \Delta\beta; \end{aligned}$$

and

$$\begin{aligned} t_3^* &= \beta_3 - e_3^* + \frac{1}{2}(e_3^*)^2, \\ t_2^* &= \beta_2 - e_2^* + \frac{1}{2}(e_2^*)^2 + \Delta\beta e_3^* - \frac{1}{2}(\Delta\beta)^2, \\ t_1^* &= \beta_1 - \frac{1}{2} + \Delta\beta(e_2^* + e_3^*) - (\Delta\beta)^2. \end{aligned}$$

Next we characterize the maximal expected welfare that obtains under complete uncertainty.

Corollary 1. *The expected welfare that corresponds to the optimal allocation of Proposition 1 is given as*

$$\begin{aligned} &E_\mu\left(W^{\mathrm{Unc}}\right) := E_\mu(W(\mathbf{t}^*, \mathbf{e}^*)) = \\ &R - \left[\left((1+\lambda)\left(\beta_1 - \frac{1}{2}\right) + \lambda(\Delta\beta(e_2^* + e_3^*) - (\Delta\beta)^2)\right)\mu(\beta_1)\right. \\ &+ \left((1+\lambda)\left(\beta_2 - e_2^* + \frac{1}{2}(e_2^*)^2\right) + \lambda\left(\Delta\beta e_3^* - \frac{1}{2}(\Delta\beta)^2\right)\right)\mu(\beta_2) \\ &\left. + (1+\lambda)\left(\beta_3 - e_3^* + \frac{1}{2}(e_3^*)^2\right)\mu(\beta_3)\right]. \end{aligned} \tag{11}$$

3.1 Discussion: Parameter Conditions

The literature derives the optimal allocation of the Laffont–Tirole regulatory framework either for a continuous interval of cost-types or for two different cost-types only.[6] The challenge in deriving Proposition 1 for three different cost-types is to identify the parameter conditions (9) and (10) that ensure the existence of an internal optimum.

The optimal allocation of Proposition 1 is derived under the assumption that the local downward incentive constraints (=LDICs) – according to which cost-type β_i has no strict incentive to mimic the contracted costs of type β_{i+1} – are binding. By a standard argument, the Spence–Mirrlees single crossing condition ensures that binding LDICs also imply the remaining downward incentives constraints (=DICs) to hold, which means in our case that cost-type β_1 has no strict incentive to mimic cost-type β_3. However, as the Spence–Mirrlees single crossing condition is not applicable to our model, we have to ensure directly that this DIC also holds for the allocation of Proposition 1. This is done by the likelihood ratio condition (9) which is equivalent to $e_2^\star \geq e_3^\star$ so that the optimal effort levels become monotonic in the firm's cost-type.

Turn now to the size of $\Delta\beta$ which needs to be sufficiently small for the optimal allocation to be characterized by Proposition 1. Condition (10) is mathematically equivalent to

$$e_3^\star \leq \Delta\beta$$

which implies, together with $e_2^\star \geq e_3^\star$, that

$$e_{i+1}^\star \leq (\beta_{i+1} - \beta_i) \text{ for } i \in \{1, 2\}.$$

Consequently, inequality (10) ensures the existence of some $e_i' \in [0, 1]$ such that $(\beta_i - e_i') = C_{i+1}^\star$ for $i \in \{1, 2\}$ so that the more efficient cost-type β_i is physically able to mimic the contracted costs $C_{i+i}^\star$ of the less efficient type β_{i+1}. Without condition (10) the LDICs would become superfluous to the effect that these LDICs are no longer binding in an optimum as assumed by the optimal allocation $(\mathbf{t}^\star, \mathbf{e}^\star)$ of Proposition 1.

We conclude this section with an illustrating example.

6 For textbook treatments of both standard cases see Salanié (2005) and Bolton and Dewatripont (2005).

Example 1. Assume that all cost-types are equally likely, i.e.

$$\mu(\beta_i) = \frac{1}{3} \text{ for } i \in \{1, 2, 3\}.$$

Clearly, condition (9) is satisfied, implying

$$e_2^\star = 1 - \frac{\lambda}{1+\lambda}\Delta\beta$$
$$> 1 - \frac{\lambda}{1+\lambda} 2\Delta\beta = e_3^\star.$$

Condition (10) concerning the maximal size of $\Delta\beta$ is satisfied if and only if

$$\frac{1+\lambda}{1+3\lambda} \geq \Delta\beta.$$

Because of

$$\lim_{\lambda\to\infty} \frac{1+\lambda}{1+3\lambda} = \frac{1}{3},$$

Proposition 1 characterizes the optimal allocation for any value $\lambda \geq 0$ of the 'marginal deadweight losses from taxes' parameter as long as we ensure that $\Delta\beta \leq \frac{1}{3}$. □

4 Superior Allocations Under Expert Information

According to Blackwell a given information (i.e. signal) structure is *more informative* than another if the decision maker cannot be worse of under the former structure (cf. Blackwell 1953; Blackwell and Girshick 1954; Campbell 2004; Lehrer and Shmaya 2008; Leshno and Spector 1992).[7] This section discusses in some detail how expert information results in ex ante superior transfer-effort allocations compared to the less informative information structure of a Laffont–Tirole regulator who is completely uncertain.

Consider a finite state space Ω with generic element ω which captures all dimensions of uncertainty that are relevant to the decision maker. We refer to the members in the powerset of Ω, denoted 2^Ω, as *events*. To capture the uncertainty that is relevant to our Laffont–Tirole regulator, we need some state space Ω such that the three possible costs types β_1, β_2, and β_3 are events in 2^Ω that form a partition

7 For a detailed analysis of the relationship between better signal structures and refined information partitions see Green and Stokey (1978).

of Ω. An *information* (i.e. *signal*) *structure* on Ω, denoted Π, is a partition of Ω into $n \geq 1$ different *information cells* (i.e. *signals*), i.e.

$$\Pi = \{I_1, \ldots, I_n\}$$

with $I_k \neq \emptyset$, $I_i \cap I_j = \emptyset$ for $i \neq j$, and $\bigcup_{k=1,..,n} I_k = \Omega$. The interpretation is that the decision maker with information structure Π receives in the state $\omega \in \Omega$ the information (i.e. signal) I_k such that $\omega \in I_k$. Based on this information, the decision maker can then exclude in state $\omega \in \Omega$ all events $A \in 2^\Omega$ as impossible for which $A \cap I_k = \emptyset$ holds. The trivial partition $\Pi^{\text{Unc}} = \{I_1 = \Omega\}$ is *non-informative* because it does not allow the decision maker to exclude any events as impossible. The information structure of the Laffont–Tirole regulator of the previous section who acted under complete uncertainty corresponds to this non-informative information structure Π according to which the regulator only receives the trivial signal Ω. In contrast, any expert would come, by definition, with some *informative* information structure

$$\Pi^{\text{Exp}} = \{I_1, \ldots, I_n\}$$

in the specific sense that $n \geq 2$ so that she receives in any state $\omega \in \Omega$ the non-trivial signal $I_k \subsetneq \Omega$ with $\omega \in I_k$.

We call an allocation $(\mathbf{t}, \mathbf{e})$ Π-*measurable* if and only if all transfers and effort levels are constant across the states in any given information cell $I_k \in \Pi$. That is, $(\mathbf{t}, \mathbf{e})$ is Π-measurable if and only if

$$\omega, \omega' \in I_k \text{ implies } (\mathbf{t}(\omega), \mathbf{e}(\omega)) = (\mathbf{t}(\omega'), \mathbf{e}(\omega'))$$

for all $I_k \in \Pi$. If is Π-measurable, we simply write $(\mathbf{t}, \mathbf{e})[I_k]$ for $(\mathbf{t}(\omega), \mathbf{e}(\omega))$ with $\omega \in I_k$. Because the expert information partition Π^{Exp} is a strict refinement of the trivial partition Π^{Unc}, any Π^{Unc}-measurable allocation $(\mathbf{t}, \mathbf{e})$ must also be Π^{Exp}-measurable. The optimal allocation $(\mathbf{t}^*, \mathbf{e}^*)$ under complete uncertainty characterized in Proposition 1 is Π^{Unc}-measurable as it is constant across all states $\omega \in \Omega$. In other words, under complete uncertainty the regulator can condition her optimal contract only on the non-informative event Ω. If the regulator could use instead the information structure Π^{Exp} of an expert, she can condition her optimal contract on the signals $I_k \in \Pi^{\text{Exp}}$. The corresponding (second best) allocation in terms of type-dependent transfers and efforts

$$(\mathbf{t}^{**}, \mathbf{e}^{**})[I_1], \ldots, (\mathbf{t}^{**}, \mathbf{e}^{**})[I_n]$$

becomes Π^{Exp}-measurable in the sense that $(\mathbf{t}^{**}, \mathbf{e}^{**})[I_k]$ has to be constant across all states $\omega \in I_k$ whereby we allow for the possibility that

$$(\mathbf{t}^{**}, \mathbf{e}^{**})[I_j] \neq (\mathbf{t}^{**}, \mathbf{e}^{**})[I_k]$$

for $j \neq k$. That is, all allocations that the Laffont–Tirole regulator can achieve under complete uncertainty form a strict subset of the allocations that she can achieve through contracts under the expert information Π^{Exp}. As a consequence, the optimal Π^{Unc}-measurable allocation $(\mathbf{t}^*, \mathbf{e}^*)$ can never do better as the optimal Π^{Exp}-measurable allocation.

To be specific, recall that the Π^{Unc}-measurable allocation $(\mathbf{t}^*, \mathbf{e}^*)$ from Proposition 1 maximizes the aggregate expected welfare function

$$E_\mu(W(\mathbf{t},\mathbf{e})) = R - \sum_{i=1}^{3} (\lambda t_i + (\beta_i - e_i) + \varphi(e_i))\mu(\beta_i) \tag{12}$$

where the prior belief μ captures the regulator's complete uncertainty. Equipped with the expert information $\Pi^{\text{Exp}} = \{I_1, \ldots, I_n\}$, the Laffont–Tirole regulator would choose the Π^{Exp}-measurable allocation

$$(\mathbf{t}^{**}, \mathbf{e}^{**})\,[I_1], \ldots, (\mathbf{t}^{**}, \mathbf{e}^{**})\,[I_n] \tag{13}$$

that maximizes the following expectation over the I_k-conditional expected aggregate welfare

$$E_\mu(W(\mathbf{t},\mathbf{e})) = R - \sum_{k=1}^{n} \mu(I_k) \sum_{i=1}^{3} (\lambda t_i[I_k] + (\beta_i - e_i[I_k]) + \varphi(e_i[I_k]))\mu(\beta_i|I_k). \tag{14}$$

As the regulator could always choose the Π^{Unc}-measurable allocation $(\mathbf{t}^*, \mathbf{e}^*)$ from Proposition 1, the maximal expected welfare that can be achieved through the Π^{Exp}-measurable allocation (13) can never be worse than the maximal expected welfare under complete uncertainty achieved through $(\mathbf{t}^*, \mathbf{e}^*)$. Moreover, whenever there is some I_k such that

$$\mu(I_k) > 0 \text{ and } \mu(\beta_i) \neq \mu(\beta_i|I_k),$$

the maximization problems (12) and (14) are different to the effect that (13) will result in strictly greater expected aggregate welfare than $(\mathbf{t}^*, \mathbf{e}^*)$.

To be even more specific, observe that maximizing (14) is formally equivalent to maximizing, for every I_k, the S_k-conditional expected aggregate welfare

$$E_{\mu(\cdot|S_k)}(W((\mathbf{t},\mathbf{e})\,[S_k])) = R - \sum_{i=1}^{3} (\lambda t_i[I_k] + (\beta_i - e_i[I_k]) + \varphi(e_i[I_k]))\mu(\beta_i|I_k).$$

Thus, by simply substituting posteriors for priors, we can immediately generalize Proposition 1 to characterize the optimal allocation (13) under expert information for suitable parameter conditions.[8]

8 In other words, Proposition 1 is the special case of Proposition 1* for which it holds that $I_k = \Omega$.

Proposition 1*. *Consider the situation of a Laffont–Tirole regulator who has acquired the expert information*

$$\Pi^{\text{Exp}} = \{I_1, \ldots, I_n\} \tag{15}$$

resulting in posterior beliefs

$$\mu(\beta_i|I_k) > 0 \text{ for all } \beta_i \text{ and } I_k.$$

Suppose that the model parameters satisfy, for all I_k, the following two inequalities

$$\frac{\mu(\beta_2|I_k)}{\mu(\beta_1|I_k)} \geq \frac{\mu(\beta_3|I_k)}{\mu(\beta_1|I_k) + \mu(\beta_2|I_k)},$$
$$\frac{\mu(\beta_3|I_k) + \mu(\beta_3|I_k)\lambda}{\mu(\beta_3|I_k) + \lambda} \geq \Delta\beta.$$

Then the optimal allocation

$$(\mathbf{t}^{**}, \mathbf{e}^{**}), [I_1], \ldots, (\mathbf{t}^{**}, \mathbf{e}^{**}), [I_n]$$

that can be implemented through a contract under the expert information (15) is given as follows: For all I_k,

$$e_1^\star[I_k] = 1,$$
$$e_2^\star[I_k] = 1 - \frac{\lambda}{1+\lambda}\frac{\mu(\beta_1|I_k)}{\mu(\beta_2|I_k)}\Delta\beta,$$
$$e_3^\star[I_k] = 1 - \frac{\lambda}{1+\lambda}\frac{1 - \mu(\beta_3|I_k)}{\mu(\beta_3|I_k)}\Delta\beta;$$

and

$$t_3^\star[I_k] = \beta_3 - e_3^\star[I_k] + \frac{1}{2}(e_3^\star[I_k])^2,$$
$$t_2^\star[I_k] = \beta_2 - e_2^\star[I_k] + \frac{1}{2}(e_2^\star[I_k])^2 + \triangle\beta e_3^\star[I_k] - \frac{1}{2}(\triangle\beta)^2,$$
$$t_1^\star[I_k] = \beta_1 - \frac{1}{2} + \triangle\beta(e_2^\star[I_k] + e_3^\star[I_k]) - (\triangle\beta)^2.$$

The optimal allocation of Proposition 1* is best understood as the allocation that corresponds to expert information which results in updated posteriors that remain 'close' to the original priors of Proposition 1. In particular, these posteriors have all full support on the possible cost-types.[9] When we consider in the next Section the special case of 'better information' defined as 'better knowledge', Proposition 1* is no longer applicable because the posteriors change rather drastically in that one posterior will assign zero to some cost-type.

5 Optimal Allocations Under 'Better Knowledge'

In line with our two assumptions about knowledge-generating expert information (cf. Section 1), we now consider an expert who is not only able to generate better information but (i) who can generate better knowledge about the firm's cost-type whereby (ii) some uncertainty still remains. More precisely, 'better knowledge' in our sense means that the regulator can exclude in every state of the world exactly one cost-type as impossible whereby two cost-types remain possible.[10]

To capture our notion of 'better knowledge', introduce the following state space which consists of six different states of the world

$$\Omega = \{(\beta_1^*, \beta_2), (\beta_1, \beta_2^*), (\beta_1^*, \beta_3), (\beta_1, \beta_3^*), (\beta_2^*, \beta_3), (\beta_2, \beta_3^*)\}.$$

We define the events according to which the firm's true cost-type is β_i, $i = 1, 2, 3$, respectively, as the following subsets of Ω

$$\begin{aligned} \beta_1 &:= \{(\beta_1^*, \beta_2), (\beta_1^*, \beta_3)\}, \\ \beta_2 &:= \{(\beta_1, \beta_2^*), (\beta_2^*, \beta_3)\}, \\ \beta_3 &:= \{(\beta_1, \beta_3^*), (\beta_2, \beta_3^*)\}. \end{aligned}$$

Next introduce the following expert information partition of Ω

$$\Pi^{\text{Exp}} = \{I_{[1,2]}, I_{[1,3]}, I_{[2,3]}\} \tag{16}$$

such that

$$I_{[i,j]} = \{(\beta_i^*, \beta_j), (\beta_i, \beta_j^*)\}.$$

9 Alternative models of expert information economic applications that result in posteriors with full support are *Gaussian-quadratic* models where normally distributed priors are updated to normally distributed posteriors in the light of normally distributed signals (cf. Angeletos and Pavan 2007; Colombo, Femminis, and Pavan 2014).

10 In a previous version of this paper, we had formally defined our notion of 'better knowledge' in terms of a set-theoretic knowledge operator (cf., e.g. Battigalli and Bonanno 1999, or Chapter 5.1 in Osborne and Rubinstein 1994). The reviewers convinced us, however, that this was unnecessarily complicated as only the updating of priors to posteriors matters for the Bayesian regulator.

The interpretation is as follows. If the expert learns – in either state $(\beta_i^{\star}, \beta_j)$ or $(\beta_i, \beta_j^{\star})$ – information

$$I_{[i,j]} = \left\{ (\beta_i^{\star}, \beta_j), \left(\beta_i, \beta_j^{\star}\right) \right\} \in \Pi^{\mathrm{Kno}},$$

she can exclude the cost-type β_k, $k \neq i, j$, as impossible whereas the cost-types β_i and β_j remain possible. More precisely, in state $(\beta_i^{\star}, \beta_j)$ the true cost-type is β_i whereas in state $(\beta_i, \beta_j^{\star})$ the true cost-type is β_j.[11] Table 1 lists, for every state of the world, the cost-types that the expert perceives as possible in this state of the world; (only these cost-types will have a strictly positive probability after the regulator updates her prior).

Suppose now that the regulator has acquired better knowledge from the expert. The regulator is a Bayesian decision maker who resolves her uncertainty about the firm's true cost-type through the prior μ defined on $(\Omega, 2^{\Omega})$. If the true state of the world is, e.g. $(\beta_i^{\star}, \beta_j)$ she updates her prior in the light of the expert information

$$I_{[i,j]} = \left\{ (\beta_i^{\star}, \beta_j), \left(\beta_i, \beta_j^{\star}\right) \right\} \in \Pi^{\mathrm{Exp}} \tag{17}$$

to the posterior $\mu(\cdot | I_{[i,j]})$ such that

$$\mu(\beta_i | I_{[i,j]}) = \frac{\mu(\beta_i^{\star}, \beta_j)}{\mu(I_{[i,j]})},$$

$$\mu(\beta_j | I_{[i,j]}) = \frac{\mu\left(\beta_i, \beta_j^{\star}\right)}{\mu(I_{[i,j]})},$$

$$\mu(\beta_k | I_{[i,j]}) = 0.$$

Table 1: Cost types that are perceived as possible by the expert.

State of the world	Possible cost-types
$(\beta_1^{\star}, \beta_2)$	$[\beta_1, \beta_2]$
$(\beta_1, \beta_2^{\star})$	$[\beta_1, \beta_2]$
$(\beta_1^{\star}, \beta_3)$	$[\beta_1, \beta_3]$
$(\beta_1, \beta_3^{\star})$	$[\beta_1, \beta_3]$
$(\beta_2^{\star}, \beta_3)$	$[\beta_2, \beta_3]$
$(\beta_2, \beta_3^{\star})$	$[\beta_2, \beta_3]$

11 The asterisk notation indicates that the cost type β_i is actually the firm's true cost type in state $(\beta_i^{\star}, \beta_j)$.

The optimal allocation under 'better knowledge' in our sense is not covered by Proposition 1* because the posterior constraints for the optimal allocation of Proposition 1* would be violated by zero-probability posteriors. Knowing that only the cost-types $[\beta_i, \beta_j]$ are possible, the regulator designs a second best allocation $(\mathbf{t}^*, \mathbf{e}^*)\,[i, j]$ that maximizes the conditional expected welfare

$$E_{\mu(\cdot|I_{[i,j]})}(W(\mathbf{t}, \mathbf{e}))$$

$$= R - \left(\lambda t_i + (\beta_i - e_i) + \frac{1}{2}(e_i)^2\right) \frac{\mu(\beta_i^*, \beta_j)}{\mu(I_{[i,j]})}$$

$$-\left(\lambda t_j + (\beta_j - e_j) + \frac{1}{2}(e_j)^2\right) \frac{\mu(\beta_i, \beta_j^*)}{\mu(I_{[i,j]})}$$

subject to incentive compatibility and participations constraints. Observe that the cost-type i can physically mimic the optimal costs C_j^* of cost-type j if and only if the optimal allocation satisfies

$$e_j^* \geq \beta_j - \beta_i.$$

For only two cost-types we obtain the following standard result.

Proposition 2. *Fix any possible cost-types* $[\beta_i, \beta_j]$ *with* $i < j$. *Suppose that the model parameters satisfy*

$$\frac{(1+\lambda)\mu(\beta_i, \beta_j^*)}{(1+\lambda)\mu(\beta_i, \beta_j^*) + \lambda\mu((\beta_i^*, \beta_j))} \geq \beta_j - \beta_i. \tag{18}$$

Then the optimal allocation $(\mathbf{t}^*, \mathbf{e}^*)\,[i, j]$ *that can be implemented through a contract under better knowledge is given as:*

$$e_i^* = 1,$$

$$e_j^*[i,j] = 1 - \frac{\lambda}{1+\lambda} \frac{\mu((\beta_i^*, \beta_j))}{\mu(\beta_i, \beta_j^*)} (\beta_j - \beta_i)$$

and

$$t_j^* = \beta_j - e_j^*[i,j] + \frac{1}{2}\left(e_j^*[i,j]\right)^2,$$

$$t_i^* = \beta_i - \frac{1}{2} + (\beta_j - \beta_i)e_j^*[i,j] - \frac{1}{2}(\beta_j - \beta_i)^2.$$

Let us illustrate the parameter condition (18) through an example.

Example 2. Suppose that all states are equally likely, i.e.

$$\mu\left(\beta_i, \beta_j^{\star}\right) = \mu\left(\beta_i, \beta_j^{\star}\right) = \frac{1}{6} \text{ for } i < j.$$

Then for any $[\beta_i, \beta_j]$ the optimal effort level of the less efficient type j is given as

$$e_j^{\star} = 1 - \frac{\lambda}{1+\lambda}(\beta_j - \beta_i).$$

The parameter condition (18) becomes for $[\beta_1, \beta_2]$ and $[\beta_2, \beta_3]$

$$\frac{1+\lambda}{1+2\lambda} \geq \Delta\beta$$

whereas it becomes for $[\beta_1, \beta_3]$

$$\frac{1+\lambda}{2+4\lambda} \geq \Delta\beta.$$

Taking the limit

$$\lim_{\lambda\to\infty} \frac{1+\lambda}{2+4\lambda} = \frac{1}{4}$$

shows that Proposition 2 pins down the optimal allocations for any value $\lambda \geq 0$ of the 'marginal deadweight losses from taxes' parameter as long as we ensure that $\Delta\beta \leq \frac{1}{4}$. □

To characterize the maximal expected welfare achievable under better knowledge, we have to take the expectation over the conditional expected welfares corresponding to the optimal allocations of Proposition 2.

Corollary 2. *The expected welfare that corresponds to the optimal allocations* $(\mathbf{t}^{\star}, \mathbf{e}^{\star})[i, j]$ *of Proposition 2 is given as*[12]

12 In what follows, we will see that the rather incomprehensible formula (19) will greatly simplify under the parameter assumption of equally likely states of the world.

$$
\begin{aligned}
E_\mu\left(W^{\text{Kno}}\right) &:= \sum_{[i,j]\in\{[1,2],[1,3],[2,3]\}} E_{\mu}(\cdot|I_{[i,j]}) W((\mathbf{t}^\star, \mathbf{e}^\star)[i,j])\mu(I_{[i,j]}) \\
&= R- \\
&\Big[\Big((1+\lambda)\Big(\beta_1-\frac{1}{2}\Big)+\lambda\Big(\Delta\beta e_2^\star[1,2]-\frac{1}{2}(\Delta\beta)^2\Big)\Big)\mu(\beta_1^\star,\beta_2) \\
&+(1+\lambda)\Big(\beta_2-e_2^\star[1,2]+\frac{1}{2}(e_2^\star[1,2])^2\Big)\mu(\beta_1,\beta_2^\star) \\
&+\Big((1+\lambda)\Big(\beta_1-\frac{1}{2}\Big)+\lambda\Big(2\Delta\beta e_3^\star[1,3]-\frac{1}{2}(2\Delta\beta)^2\Big)\Big)\mu(\beta_1^\star,\beta_3) \\
&+(1+\lambda)\Big(\beta_3-e_3^\star[1,3]+\frac{1}{2}(e_3^\star[1,3])^2\Big)\mu(\beta_1,\beta_3^\star) \\
&+\Big((1+\lambda)\Big(\beta_2-\frac{1}{2}\Big)+\lambda\Big(\Delta\beta e_3^\star[2,3]-\frac{1}{2}(\Delta\beta)^2\Big)\Big)\mu(\beta_2^\star,\beta_3) \\
&+(1+\lambda)\Big(\beta_3-e_3^\star[2,3]+\frac{1}{2}(e_3^\star[2,3])^2\Big)\mu(\beta_2,\beta_3^\star)\Big]
\end{aligned}
\tag{19}
$$

such that

$$
\begin{aligned}
e_2^\star[1,2] &= 1-\frac{\lambda}{1+\lambda}\,\frac{\mu((\beta_1^\star,\beta_2))}{\mu(\beta_1,\beta_2^\star)}\,\Delta\beta, \\
e_3^\star[1,3] &= 1-\frac{\lambda}{1+\lambda}\,\frac{\mu((\beta_1^\star,\beta_3))}{\mu(\beta_1,\beta_3^\star)}\,2\Delta\beta, \\
e_3^\star[2,3] &= 1-\frac{\lambda}{1+\lambda}\,\frac{\mu((\beta_2^\star,\beta_3))}{\mu(\beta_2,\beta_3^\star)}\,\Delta\beta.
\end{aligned}
$$

6 When is Knowledge Acquisition Socially Beneficial?

Before we turn to the question for which values of the ‘marginal deadweight losses from taxes’ parameter it would be optimal for the regulator to acquire better knowledge, we derive a closed-form expression for the expected welfare gain that arises from the optimal allocations under better knowledge. This gain is given as the difference

$$E_\mu(\triangle W):=E_\mu\left(W^{\text{Kno}}\right)-E_\mu\left(W^{\text{Unc}}\right) \tag{20}$$

such that $E_\mu(W^{\text{Unc}})$ is defined by (11) and $E_\mu(W^{\text{Kno}})$ is defined by (19).

Of course, the expected welfare achievable under better knowledge must be at least as good as the expected welfare under complete uncertainty. Beyond this general insight that more information cannot be worse than less information for a standard maximization problem, we would like to come up with a closed form characterization of this informational advantage in terms of our model's 'marginal deadweight losses from taxes' parameter. To make the analysis tractable, we impose the following technically convenient assumptions.

Parameter assumptions.

(A1). All states of the world have equal probability, i.e.

$$\mu(\beta_i^{\star}, \beta_j) = \mu(\beta_i, \beta_j^{\star}) = \frac{1}{6} \text{ for all } i < j.$$

(A2). The cost-type difference satisfies $\Delta\beta \le \frac{1}{4}$.

The crucial assumption is A1 which greatly simplifies the optimal effort levels of Propositions 1 and 2. Moreover, A1 ensures that condition (9) of Proposition 1 is satisfied. Assumption A1 together with Assumption A2, which concerns the maximal difference between cost-types, thus guarantee that Propositions 1 and 2 indeed characterize the optimal allocations achievable under both scenarios for all possible values $\lambda \ge 0$ of the 'marginal deadweight losses from taxes' parameter.

Proposition 3. *Suppose that the Assumptions A1 and A2 hold. Then the expected welfare gain (20) from better knowledge becomes*

$$E_\mu(\Delta W) = \frac{1}{3}\lambda\Delta\beta\left(1 - \frac{\lambda}{1+\lambda}\Delta\beta\right). \tag{21}$$

Let us interpret the expected welfare gain (21) as a function in the 'marginal deadweight losses from taxes' parameter. Observe that the first-order derivative

$$\frac{d}{d\lambda}E_\mu(\Delta W) = \frac{1}{3}\Delta\beta\left(1 - \frac{2\lambda}{1+2\lambda+\lambda^2}\Delta\beta\right)$$

is always strictly greater zero for $\lambda \ge 0$. Moreover, taking the second order derivative shows that the expected welfare gain (21) is a strictly concave function in the 'marginal deadweight losses from taxes' parameter. In words: The expected welfare gain from knowledge acquisition strictly increases in the 'marginal deadweight losses from taxes' parameter but with diminishing returns.

Suppose now that the acquisition of better knowledge comes at the fixed costs $L \geq 0$, which translates into the social costs $(1 + \lambda)L$. The acquisition of better knowledge thus strictly improves the overall social welfare if and only if

$$\begin{gathered}(1+\lambda)L < E_\mu(\Delta W) \\ \Leftrightarrow \\ L < T(\lambda;\Delta\beta)\end{gathered} \tag{22}$$

whereby the threshold function is given as

$$T(\lambda;\Delta\beta) = \frac{1}{3}\,\frac{\lambda}{1+\lambda}\,\Delta\beta\left(1 - \frac{\lambda}{1+\lambda}\,\Delta\beta\right). \tag{23}$$

That is, knowledge acquisition strictly increases the overall welfare for fixed parameter values λ and $\Delta\beta$ if and only if the costs L fall below the threshold (23). Under Assumption A2, the threshold function (23) strictly increases in both parameters λ and $\Delta\beta$, respectively. Fix $\Delta\beta$ and observe that there exists a least upper bound given by

$$T(\lambda;\Delta\beta) = \frac{1}{3}\left(\Delta\beta - (\Delta\beta)^2\right).$$

By continuity of $T(\lambda; \Delta\beta)$, there exists a finite $\lambda^* \geq 0$ such that

$$L = T(\lambda^*;\Delta\beta) \tag{24}$$

if and only if the following strict inequality holds

$$L < \frac{1}{3}\left(\Delta\beta - (\Delta\beta)^2\right). \tag{25}$$

As $T(\lambda; \Delta\beta)$ is strictly increasing in λ, the inequality (22) is then satisfied for all $\lambda \in (\lambda^*, \infty)$. In words: Knowledge acquisition strictly increases the overall welfare if and only if the value of λ is strictly greater than the critical value λ^*.

Figure 1 plots the graphs of the functions $T(\lambda; \Delta\beta)$ in $\lambda \in [0, 1]$ for the three different cost-type differences $\Delta\beta \in \{0.08, 0.16, 0.24\}$. For the two different values of the fixed costs $L' = 0.005$ and $L'' = 0.01$ the critical values λ^* correspond to the λ-values at the intersections of the L' and L''-lines, respectively, with the graphs of the $T(\lambda; \Delta\beta)$ functions. Analytically, λ^* is, by (24), pinned down as the positive root[13] of the quadratic equation

13 The other root of this quadratic equation defined for $\lambda \in \mathbb{R}$ must be negative under our parameter assumptions. To see this, note that $T(\cdot;\Delta\beta) : \mathbb{R} \to \mathbb{R}$ has a unique extreme point at $\lambda_\star = -\frac{1}{1-2\Delta\beta} < 0$. As $T(\lambda_\star; \Delta\beta)$ is the global minimum, the other root is the intersection of $T(\lambda; \Delta\beta)$ with the line L at some value strictly smaller than $\lambda_\star < 0$.

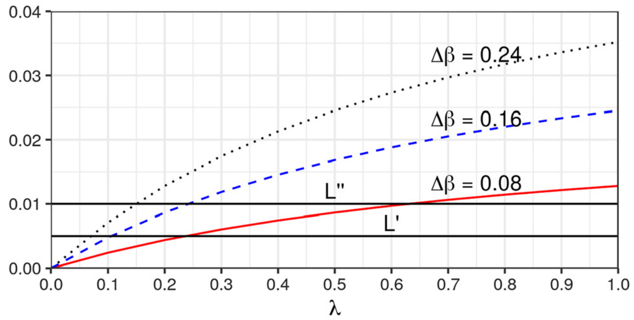

Figure 1: Threshold functions $T(\lambda;\Delta\beta)$ for $\Delta\beta \in \{0.08, 0.16, 0.24\}$.

$$L = \frac{1}{3}\frac{\lambda^*}{1+\lambda^*}\Delta\beta\left(1-\frac{\lambda^*}{1+\lambda^*}\Delta\beta\right)$$
$$\Leftrightarrow$$
$$0 = \left((\Delta\beta - (\Delta\beta)^2) - 3L\right)(\lambda^*)^2 + (\Delta\beta - 6L)\lambda^* - 3L.$$

The following proposition summarizes the above argument.

Proposition 4. *Suppose that the Assumptions A1 and A2 hold.*

(i) *If inequality (25) is violated, then knowledge acquisition strictly decreases the overall welfare for any value $\lambda > 0$ of the 'marginal deadweight losses from taxes' parameter.*

(ii) *If, instead, inequality (25) holds, then knowledge acquisition strictly increases the overall welfare if and only if the value λ of the 'marginal deadweight losses from taxes' parameter satisfies $\lambda \in (\lambda^*, \infty)$ whereby the critical value is given as*

$$\lambda^* = \frac{\sqrt{(\Delta\beta - 6L)^2 + 4\left((\Delta\beta - (\Delta\beta)^2) - 3L\right)3L} - (\Delta\beta - 6L)}{2\left((\Delta\beta - (\Delta\beta)^2) - 3L\right)}. \qquad (26)$$

Observe that the critical value (26) for the 'marginal deadweight losses from taxes' parameter equals zero if and only the costs L are zero. That is, strictly positive costs $L > 0$ imply $\lambda^* > 0$ so that there will always exist sufficiently small values for the 'marginal deadweight losses from taxes' parameter $\lambda < \lambda^*$ for which knowledge acquisition would strictly decrease the overall social welfare.

7 Concluding Remarks

In the Laffont–Tirole regulatory framework both the welfare gains from knowledge acquisition but also the associated social costs of such acquisition strictly increase in the 'marginal deadweight losses from taxes' parameter. This paper constructs an analytically tractable model which allows us to derive an analytical expression for the expected welfare gain that results from the implementation of optimal allocations under better knowledge. Taking into account knowledge acquisition costs, we show that the regulator should only acquire better knowledge about the firm's cost-type if the value of the 'marginal deadweight losses from taxes' parameter is above a critical threshold, which is strictly greater zero for positive acquisition costs. In other words, knowledge acquisition within the Laffont–Tirole regulatory framework would actually decrease the overall social welfare for economies in which deadweight losses from taxation are rather low.

Appendix Formal Proofs

Proof of Proposition 1. Step 1

We show that the optimal allocation of Proposition 1 obtains if (i) the participation constraint of the least efficient type and (ii) the local downward incentive constraints (LDICs) are binding.

Assume that the participation constraint

$$U(t_3, e_3; \beta_3) = 0 \tag{27}$$

and the LDICs

$$t_i - (\beta_i - e_i) - \frac{1}{2}(e_i)^2 = t_{i+1} - (\beta_{i+1} - e_{i+1}) - \frac{1}{2}(e_{i+1} - (\beta_{i+1} - \beta_i))^2 \text{ for all } i = 1, 2$$

are binding in an optimum. Rewriting the optimal transfers as functions in $\mathbf{e} = (e_1, e_2, e_3)$ (on some open neighborhood around $\mathbf{e}^*$) gives

$$t_3^\star(\mathbf{e}) = \beta_3 - e_3 + \frac{1}{2} e_3^2,$$

$$t_2^\star(\mathbf{e}) = \beta_2 - e_2 + \frac{1}{2} e_2^2 + (\beta_3 - \beta_2)e_3 - \frac{1}{2}(\beta_3 - \beta_2)^2,$$

$$t_1^\star(\mathbf{e}) = \beta_1 - e_1 + \frac{1}{2} e_1^2 + (\beta_2 - \beta_1)e_2 - \frac{1}{2}(\beta_2 - \beta_1)^2 + (\beta_3 - \beta_2)e_3 - \frac{1}{2}(\beta_3 - \beta_2)^2.$$

Substitute these transfers into the expected welfare function to obtain

$$E_\mu(W(\mathbf{t}^\star(\mathbf{e}), \mathbf{e})) = R - \sum_{i=1}^{3} (\lambda t_i^\star + (\beta_i - e_i) + \varphi(e_i))\mu(\beta_i).$$

Maximizing this expected welfare function through first-order conditions gives, by strict concavity, the optimal (interior) effort levels $e_i^\star \in [0,1]$ of Proposition 1. □

Step 2

Note that the LDICs only pin down an interior solution in Step 1 if $e_{i+1}^\star \geq \beta_{i+1} - \beta_i$ for $i \in \{1, 2\}$. We show that these inequalities are ensured by (10) together with the likelihood ratio condition (9).

At first observe that

$$e_3^\star \geq \beta_3 - \beta_2$$

$$\Leftrightarrow$$

$$1 - \frac{\lambda}{1+\lambda}\,\frac{1-\mu(\beta_3)}{\mu(\beta_3)}\,\Delta\beta \geq \Delta\beta$$

$$\Leftrightarrow$$

$$\frac{\mu(\beta_3)+\mu(\beta_3)\lambda}{\mu(\beta_3)+\lambda} \geq \Delta\beta.$$

By the likelihood ratio condition (9), $e_2^\star \geq e_3^\star$, which implies

$$e_2^\star \geq e_3^\star \geq \Delta\beta$$

$$\Rightarrow$$

$$e_2^\star \geq \beta_2 - \beta_1$$

whenever (10) holds. □

The following three steps show that the remaining participation constraints and ICCs are satisfied under the conditions of Proposition 1.

Step 3

A standard argument shows that the participations constraints for $i = 1, 2$ hold when the LDICs and the participation constraint (27) hold.

Because of the binding LDIC for cost-type β_2 with

$$e_2' = e_3^\star - \Delta\beta$$

we have

$$U(t_2^\star, e_2^\star; \beta_2) = t_2^\star - C_2^\star - \frac{1}{2}(e_2^\star)^2 = t_3 - C_3^\star - \frac{1}{2}(e_3^\star - \Delta\beta)^2$$

$$> t_3 - C_3^\star - \frac{1}{2}(e_3^\star)^2$$

$$= U(t_3^\star, e_3^\star; \beta_3) = 0.$$

Analogously, by the binding LDIC for cost-type β_1 with

$$e_1' = e_2^\star - \Delta\beta,$$

we obtain

$$\begin{aligned} U(t_1^\star, e_1^\star; \beta_1) &= t_2 - C_2^\star - \frac{1}{2}(e_2^\star - \Delta\beta)^2 \\ &> U(t_2^\star, e_2^\star; \beta_2). \end{aligned}$$

□

Step 4

We show that we can simply ignore the UICs for the optimal allocation.

We can ignore any upward incentive constraint

$$t_j^\star - (\beta_j - e_j^\star) - \frac{1}{2}(e_j^\star)^2 \geq t_i^\star - (\beta_i - e_i^\star) - \frac{1}{2}(e_j')^2 \text{ for all } j > i$$

because of

$$e_j' = e_i^\star + (\beta_j - \beta_i) > 1.$$

That is, a less efficient type cannot mimic the contracted costs of an efficient type as this would require greater effort levels than physically available. To see this for $i = 2$, $j = 3$, note that

$$\begin{aligned} e_2^\star + (\beta_3 - \beta_2) &= 1 - \frac{\lambda}{1+\lambda}\frac{\mu(\beta_1)}{\mu(\beta_2)}\Delta\beta + \Delta\beta \\ &= 1 + \left(1 - \frac{\lambda}{1+\lambda}\frac{\mu(\beta_1)}{\mu(\beta_2)}\right)\Delta\beta \\ &> 1. \end{aligned}$$

For $i = 1$, $j = 3$ we have that

$$e_1^\star + (\beta_3 - \beta_1) = 1 + 2\Delta\beta > 1.$$

□

Step 5

We show that the likelihood ratio condition (9) ensures that cost-type β_1 does not want to mimic the contracted costs $C_3^\star$ of type β_3.

Transforming the corresponding DIC in several steps gives

$$t_1^\star - (\beta_1 - e_1^\star) - \frac{1}{2}(e_1^\star)^2 \geq t_3^\star - (\beta_3 - e_3^\star) - \frac{1}{2}(e_3^\star - (\beta_3 - \beta_1))^2$$

$$\Leftrightarrow$$

$$(e_2^\star - e_3^\star)(\beta_2 - \beta_1) \geq \frac{1}{2}(\beta_2 - \beta_1)^2 + \frac{1}{2}(\beta_3 - \beta_2)^2 - \frac{1}{2}(\beta_3 - \beta_1)^2.$$

Next observe that the likelihood ratio condition (9) implies $(e_2^\star - e_3^\star) \geq 0$. The above inequality thus always holds if

$$0 \geq \frac{1}{2}(\beta_2 - \beta_1)^2 + \frac{1}{2}(\beta_3 - \beta_2)^2 - \frac{1}{2}(\beta_3 - \beta_1)^2$$

$$\Leftrightarrow$$

$$(\beta_3 - \beta_1)^2 \geq (\beta_2 - \beta_1)^2 + (\beta_3 - \beta_2)^2.$$

Observe that this inequality is equivalent to the superadditivity of a function f such that

$$f(x + y) \geq f(x) + f(y)$$

where

$$x = \beta_3 - \beta_2$$
$$y = \beta_2 - \beta_1$$

Recall that a convex function f which is increasing and satisfies $f(0) = 0$ must be superadditive. Setting $f(x) = x^2$ thus proves that the above DIC always holds for $e_2^\star \geq e_3^\star$. □

Proof of Proposition 3. Step 1

We subsequently derive for all six states of the world $\omega \in \Omega$ the difference

$$\Delta W(\omega) := W^{\text{Kno}}(\omega) - W^{\text{Unc}}(\omega).$$

Before we proceed, note that the respective effort levels in $W^{\text{Kno}} - W^{\text{Unc}}$ simplify, by Assumption A1, to

$$e_2^\star = 1 - \frac{\lambda}{1+\lambda}\Delta\beta,$$

$$e_3^\star = 1 - \frac{\lambda}{1+\lambda}2\Delta\beta,$$

and

$$e_2^\star[1,2] = e_3^\star[2,3] = 1 - \frac{\lambda}{1+\lambda}\Delta\beta,$$

$$e_3^\star[1,3] = 1 - \frac{\lambda}{1+\lambda}2\Delta\beta.$$

1. State $\omega = (\beta_1^\star, \beta_2)$.

$$W^{\text{Kno}}(\omega) - W^{\text{Unc}}(\omega) = \left((1+\lambda)\left(\beta_1 - \frac{1}{2}\right) + \lambda\left(\Delta\beta\left(e_2^\star + e_3^\star\right) - (\Delta\beta)^2\right)\right)$$
$$-\left((1+\lambda)\left(\beta_1 - \frac{1}{2}\right) + \lambda\left(\Delta\beta e_2^\star[1,2] - \frac{1}{2}(\Delta\beta)^2\right)\right)$$
$$= \lambda\left(\Delta\beta\left(e_2^\star + e_3^\star - e_2^\star[1,2]\right) - \frac{1}{2}(\Delta\beta)^2\right).$$

Substituting

$$e_2^\star + e_3^\star - e_2^\star[1,2] = \left(1 - \frac{\lambda}{1+\lambda}\,\Delta\beta\right) + \left(1 - \frac{\lambda}{1+\lambda}\,2\Delta\beta\right) - \left(1 - \frac{\lambda}{1+\lambda}\,\Delta\beta\right)$$
$$= \left(1 - \frac{\lambda}{1+\lambda}\,2\Delta\beta\right)$$

gives

$$\Delta W(\beta_1^\star, \beta_2) = \lambda\left(\Delta\beta\left(1 - \frac{\lambda}{1+\lambda}\,2\Delta\beta\right) - \frac{1}{2}(\Delta\beta)^2\right).$$

2. State $\omega = (\beta_1^\star, \beta_3)$.

$$W^{\text{Kno}}(\omega) - W^{\text{Unc}}(\omega) =$$
$$\left((1+\lambda)\left(\beta_1 - \frac{1}{2}\right) + \lambda\left(\Delta\beta\left(e_2^\star + e_3^\star\right) - (\Delta\beta)^2\right)\right)$$
$$-\left((1+\lambda)\left(\beta_1 - \frac{1}{2}\right) + \lambda\left(2\Delta\beta e_3^\star[1,3] - \frac{1}{2}(2\Delta\beta)^2\right)\right)$$
$$= \lambda\left(\Delta\beta\left(\left(e_2^\star + e_3^\star\right) - 2e_3^\star[1,3]\right) + (\Delta\beta)^2\right)$$
$$= \lambda\left(\Delta\beta\left(\left(1 - \frac{\lambda}{1+\lambda}\Delta\beta\right) - \left(1 - \frac{\lambda}{1+\lambda}2\Delta\beta\right)\right) + (\Delta\beta)^2\right)$$

which becomes after substituting

$$\left(e_2^\star + e_3^\star\right) - 2e_3^\star[1,3] = \left(1 - \frac{\lambda}{1+\lambda}\,\Delta\beta\right) - \left(1 - \frac{\lambda}{1+\lambda}\,2\Delta\beta\right)$$
$$\Delta W(\beta_1^\star, \beta_3) = \lambda\left(\Delta\beta\left(\left(1 - \frac{\lambda}{1+\lambda}\,\Delta\beta\right) - \left(1 - \frac{\lambda}{1+\lambda}\,2\Delta\beta\right)\right) + (\Delta\beta)^2\right).$$

3. State $\omega = (\beta_1, \beta_2^{\star})$.

$$W^{\text{Kno}}(\omega) - W^{\text{Unc}}(\omega) = \left((1+\lambda)\left(\beta_2 - e_2^{\star} + \frac{1}{2}(e_2^{\star})^2\right) + \lambda\left(\Delta\beta e_3^{\star} - \frac{1}{2}(\Delta\beta)^2\right)\right)$$

$$-(1+\lambda)\left(\beta_2 - e_2^{\star}[1,2] + \frac{1}{2}(e_2^{\star}[1,2])^2\right)$$

$$= \lambda\left(\Delta\beta e_3^{\star} - \frac{1}{2}(\Delta\beta)^2\right)$$

because of $e_2^{\star}[1,2] = e_2^{\star}$. Substituting for $e_3^{\star}$ gives

$$\Delta W(\beta_1, \beta_2^{\star}) = \lambda\left(\Delta\beta\left(1 - \frac{\lambda}{1+\lambda}2\Delta\beta\right) - \frac{1}{2}(\Delta\beta)^2\right).$$

4. State $\omega = (\beta_2^{\star}, \beta_3)$.

$$W^{\text{Kno}}(\omega) - W^{\text{Unc}}(\omega) = \left((1+\lambda)\left(\beta_2 - e_2^{\star} + \frac{1}{2}(e_2^{\star})^2\right) + \lambda\left(\Delta\beta e_3^{\star} - \frac{1}{2}(\Delta\beta)^2\right)\right)$$

$$-\left((1+\lambda)\left(\beta_2 - \frac{1}{2}\right) + \lambda\left(\Delta\beta e_3^{\star}[2,3] - \frac{1}{2}(\Delta\beta)^2\right)\right)$$

$$= \left((1+\lambda)\left(\frac{1}{2} - e_2^{\star} + \frac{1}{2}(e_2^{\star})^2\right) + \lambda(\Delta\beta(e_3^{\star} - e_3^{\star}[2,3]))\right).$$

Note that

$$\frac{1}{2} - e_2^{\star} + \frac{1}{2}(e_2^{\star})^2 = \frac{1}{2} - \left(1 - \frac{\lambda}{1+\lambda}\Delta\beta\right) + \frac{1}{2}\left(1 - \frac{\lambda}{1+\lambda}\Delta\beta\right)^2$$

$$= \frac{1}{2}\left(\frac{\lambda}{1+\lambda}\Delta\beta\right)^2$$

as well as

$$e_3^{\star} - e_3^{\star}[2,3] = \left(1 - \frac{\lambda}{1+\lambda}2\Delta\beta\right) - \left(1 - \frac{\lambda}{1+\lambda}\Delta\beta\right)$$

$$= -\frac{\lambda}{1+\lambda}\Delta\beta$$

so that substitution gives

$$\Delta W(\beta_2^{\star}, \beta_3) = (1+\lambda)\left(\frac{1}{2}\left(\frac{\lambda}{1+\lambda}\Delta\beta\right)^2\right) - \lambda\left(\Delta\beta\left(\frac{\lambda}{1+\lambda}\Delta\beta\right)\right)$$

$$= -\frac{1}{2}\frac{\lambda^2}{1+\lambda}(\Delta\beta)^2.$$

5. **State** $\omega = (\beta_1, \beta_3^\star)$.

$$W^{\mathrm{Kno}}(\omega) - W^{\mathrm{Unc}}(\omega) = (1+\lambda)\left(\beta_3 - e_3^\star + \frac{1}{2}(e_3^\star)^2\right)$$
$$-(1+\lambda)\left(\beta_3 - e_3^\star[1,3] + \frac{1}{2}(e_3^\star[1,3])^2\right)$$

implying

$$\Delta W(\beta_1, \beta_3^\star) = 0$$

because of $e_3^\star = e_3^\star[1,3]$.

6. **State** $\omega = (\beta_2, \beta_3^\star)$.

$$W^{\mathrm{Kno}}(\omega) - W^{\mathrm{Unc}}(\omega) = (1+\lambda)\left(\beta_3 - e_3^\star + \frac{1}{2}(e_3^\star)^2\right)$$
$$-(1+\lambda)\left(\beta_3 - e_3^\star[2,3] + \frac{1}{2}(e_3^\star[2,3])^2\right)$$
$$= (1+\lambda)\left(e_3^\star[2,3] - e_3^\star + \frac{1}{2}(e_3^\star)^2 - \frac{1}{2}(e_3^\star[2,3])^2\right).$$

Because of

$$e_3^\star[2,3] - e_3^\star = \frac{\lambda}{1+\lambda}\Delta\beta$$

and

$$\frac{1}{2}\left((e_3^\star)^2 - (e_3^\star[2,3])^2\right) = \frac{1}{2}\left(\left(1 - \frac{\lambda}{1+\lambda}2\Delta\beta\right)^2 - \left(1 - \frac{\lambda}{1+\lambda}\Delta\beta\right)^2\right)$$
$$= -\frac{\lambda}{1+\lambda}\Delta\beta + \frac{3}{2}\left(\frac{\lambda}{1+\lambda}\Delta\beta\right)^2$$

we have that

$$\Delta W(\beta_2, \beta_3^\star) = (1+\lambda)\left(\frac{\lambda}{1+\lambda}\triangle\beta - \frac{\lambda}{1+\lambda}\Delta\beta + \frac{3}{2}\left(\frac{\lambda}{1+\lambda}\Delta\beta\right)^2\right)$$
$$= \frac{3}{2}\frac{\lambda^2}{1+\lambda}(\Delta\beta)^2.$$

□

Step 2

Taking the expected value of ΔW over all equally likely states gives the desired expression $E_\mu(\Delta W)$ of Proposition 3.

Note that

$$
\begin{aligned}
E_\mu(\Delta W) &:= E_\mu\left(W^{\mathrm{Kno}}\right) - E_\mu\left(W^{\mathrm{Unc}}\right) = \\
&= \Bigg[\lambda\left(\Delta\beta\left(1-\frac{\lambda}{1+\lambda}\,2\Delta\beta\right)-\frac{1}{2}(\Delta\beta)^2\right) \\
&+\lambda\left(\Delta\beta\left(\left(1-\frac{\lambda}{1+\lambda}\,\Delta\beta\right)-\left(1-\frac{\lambda}{1+\lambda}\,2\Delta\beta\right)\right)+(\Delta\beta)^2\right) \\
&+\lambda\left(\Delta\beta\left(1-\frac{\lambda}{1+\lambda}\,2\Delta\beta\right)-\frac{1}{2}(\Delta\beta)^2\right) \\
&+\left(-\frac{1}{2}\,\frac{\lambda^2}{1+\lambda}(\Delta\beta)^2\right) \\
&+0 \\
&+\frac{3}{2}\,\frac{\lambda^2}{1+\lambda}(\Delta\beta)^2\Bigg]\frac{1}{6}
\end{aligned}
$$

which reduces to

$$
\begin{aligned}
E_\mu(\Delta W) &= \left[\lambda\Delta\beta\left(2-\frac{\lambda}{1+\lambda}\,3\Delta\beta\right)+\frac{\lambda^2}{1+\lambda}(\Delta\beta)^2\right]\frac{1}{6} \\
&= \frac{1}{3}\lambda\Delta\beta\left(1-\frac{\lambda}{1+\lambda}\,\Delta\beta\right).
\end{aligned}
$$

□

References

Angeletos, G.-M., and A. Pavan. 2007. "Efficient Use of Information and Social Value of Information." *Econometrica* 75: 1103–42.

Baron, D., and R. Myerson. 1982. "Regulating a Monopolist with Unknown Costs." *Econometrica* 50: 911–30.

Battigalli, P., and G. Bonanno. 1999. "Recent Results on Belief, Knowledge and the Epistemic Foundations of Game Theory." *Research in Economics* 53: 149–225.

Becker, G. S. 1985. "Public Policies, Pressure Groups and Dead Weight Costs." *Journal of Public Economics* 28: 329–47.

Blackwell, D. 1953. "Equivalent Comparisons of Experiments." *The Annals of Mathematical Statistics* 24: 265–72.

Blackwell, D., and M. A. Girshick. 1954. *Theory of Games and Statistical Decisions*. New York: John Wiley & Sons.

Blomquist, S., and L. Simula. 2019. "Marginal Deadweight Loss when the Income Tax Is Nonlinear." *Journal of Econometrics* 211: 47–60.

Bolton, P., and M. Dewatripont. 2005. *Contract Theory*. Cambridge, Massachusetts: The MIT Press.

Campbell, C. M. 2004. "Blackwell's Ordering and Public Information." *Journal of Economic Theory* 114: 179–97.

Colombo, L., G. Femminis, and A. Pavan. 2014. "Information Acquisition and Welfare." *The Review of Economic Studies* 81: 1438–83.
Crémer, J., and F. Khalil. 1992. "Gathering Information before Signing a Contract." *The American Economic Review* 82: 566–78.
Crémer, J., F. Khalil, and J.-C. Rochet. 1998a. "Contracts and Productive Information Gathering." *Games and Economic Behavior* 25: 174–93.
Crémer, J., F. Khalil, and J.-C. Rochet. 1998b. "Strategic Information Gathering before a Contract Is Offered." *Journal of Economic Theory* 81: 163–200.
Estache, A., and L. Wren-Lewis. 2009. "Toward a Theory of Regulation for Developing Countries: Following Jean-Jacques Laffont's Lead." *Journal of Economic Literature* 47: 729–70.
Green, J. R., and N. L. Stokey. 1978. *Two Representations of Information Structures and Their Comparisons*. IMSSS Technical Report No. 271. Stanford University. https://home.uchicago.edu/nstokey/papers/Green%20&%20Stokey%20-%20Two%20representations%20....pdf.
Harberger, A. 1962. "The Incidence of the Corporation Income Tax." *Journal of Political Economy* 70: 215–40.
Harberger, A. 1964. "The Measurement of Waste." *The American Economic Review* 54: 58–76.
Heckman, J. J., S. H. Moon, R. Pinto, P. A. Savelyev, and A. Yavitz. 2010. "The Rate of Return to the HighScope Perry Preschool Program." *Journal of Public Economics* 94: 114–28.
Jacobs, B. 2018. "The Marginal Cost of Public Funds Is One at the Optimal Tax System." *International Tax and Public Finance* 25: 883–912.
Laffont, J-J. 2003. "Enforcement, Regulation and Development." *Journal of African Economies* 12: 193–211.
Laffont, J-J. 2005. *Regulation and Development*. Cambridge: Cambridge University Press.
Laffont, J.-J., and J. Tirole. 1986. "Using Cost Observation to Regulate Firms." *Journal of Political Economy* 94: 614–41.
Laffont, J.-J., and J. Tirole. 1990a. "The Politics of Government Decision Making: Regulatory Institutions." *Journal of Law, Economics, and Organization* 6: 1–31.
Laffont, J.-J., and J. Tirole. 1990b. "Adverse Selection and Renegotiation in Procurement." *The Review of Economic Studies* 57: 597–625.
Laffont, J.-J., and J. Tirole. 1991. "The Politics of Government Decision Making: A Theory of Regulatory Capture." *Quarterly Journal of Economics* 106: 1089–127.
Laffont, J.-J., and J. Tirole. 1993. *A Theory of Incentives in Procurement and Regulation*. Cambridge, Massachusetts: The MIT Press.
Lehrer, E., and E. Shmaya. 2008. "Two Remarks on Blackwell's Theorem." *Journal of Applied Probability* 45: 580–6.
Leshno, M., and Y. Spector. 1992. "An Elementray Proof of Blackwell's Theorem." *Mathematical Social Sciences* 25: 95–8.
Lewis, T. 1996. "Protecting the Environment when Costs and Benefits Are Privately Known." *The RAND Journal of Economics* 27: 819–47.
Magesan, A., and M. Turner. 2010. "The Value of Information in Regulation." *The B.E. Journal of Economic Analysis & Policy* 10: 1–30.
Osborne, M., and A. Rubinstein. 1994. *A Course in Game Theory*. Cambridge, Massachusetts: The MIT Press.
Ottaviani, M., and P. N. Sørensen. 2006a. "Professional Advice." *Journal of Economic Theory* 126: 120–42.
Ottaviani, M., and P. N. Sørensen. 2006b. "Reputational Cheap Talk." *The RAND Journal of Economics* 37: 155–75.

Rogerson, W. P. 2003. “Simple Menus of Contracts in Cost-Based Procurement and Regulation.” *The American Economic Review* 93: 919–26.
Salanié, B. 2005. *The Economics of Contracts: A Primer*, 2nd ed. Cambridge, Massachusetts: The MIT Press.
Serrato, J. C. S., and O. Zidar. 2016. “Who Benefits from State Corporate Tax Cuts? A Local Labor Markets Approach with Heterogeneous Firms.” *The American Economic Review* 106: 2582–624.
Stigler, G. J. 1971. “The Theory of Economic Regulation.” *Bell Journal of Economics and Management Science* 2: 3–21.
Szalay, D. 2009. “Contracts with Endogenous Information.” *Games and Economic Behavior* 65: 586–625.
Tirole, J. 2015. “Market Failures and Public Policy.” *The American Economic Review* 105: 1665–82.
Yagan, D. 2015. “Capital Tax Reform and the Real Economy: The Effects of the 2003 Dividend Tax Cut.” *The American Economic Review* 105: 3531–63.