Incentivizing Information Acquisition*

Fan Wu†

November 24, 2026

Abstract

I study a principal-agent model in which a principal hires an agent to collect information about an unknown continuous state. The agent acquires a signal whose distribution is centered around the state, controlling the signal's precision at a cost. The principal observes neither the chosen precision nor the signal. Instead, she commits to a transfer rule, subject to limited liability and a hard budget, that can depend on the agent's report and the state realized ex post. A simple transfer rule has a cutoff structure: the agent receives a fixed prize if his report is sufficiently close to the state and receives nothing otherwise. I characterize exactly when cutoff transfers are optimal for all cost functions. The condition is weaker than MLRP and is satisfied by standard signal distributions commonly used in the literature.

Keywords. Information Acquisition; Principal-Agent Model.
JEL Codes. D82, D83, D86

*I am indebted to my committee chair Luciano Pomatto and advisor Omer Tamuz for their generous encouragement, guidance, and support. This paper also benefits from insightful discussions with Alexander Bloedel, Peter Caradonna, Axel Niemeyer, and Kun Zhang. For helpful comments, I also thank S. Nageeb Ali, Pak Hung Au, Daghan Carlos Akkar, Simon Board, Simone Cerreia-Vioglio, Daniel Gottlieb, Simon Grant, Nima Haghpanah, Wei He, Jonathan Libgober, Shuo Liu, Yi Liu, Jay Lu, Yusufcan Masatlioglu, Doron Ravid, Fedor Sandomirskiy, Balazs Szentes, Xin Shan, Andrew J. Sinclair, Jakub Steiner, Rakesh Vohra, Mu Zhang, Mengxi Zhang, Weijie Zhong, seminar participants at Caltech, UCLA, HKU, HKUST, PKU, THU, and conference participants at 2025 CETC, 2024 AMES, 2024 Midwest Theory Meeting.

†Fan Wu: HSBC Business School, Peking University (email: fanwu@phbs.pku.edu.cn).

1 Introduction

In many situations, an agent is tasked with collecting information about a quantity of interest. Examples include a statistician collecting data for the Bureau of Statistics, a meteorologist forecasting weather for the National Weather Service, or a consultant assessing the profitability of a new market. A natural question is how to design a contract to incentivize the agent to collect information and report it accurately.

A difficulty in designing such contracts is that the agent's effort and the information acquired are often unobservable to the principal. For example, a statistician might be able to fabricate part of their data. Similarly, in the case of a meteorologist analyzing weather patterns, the amount of time and effort spent on the analysis might not be easily observable. In the case of the consultant, predicting market profitability may rely on private connections or specialized skills, which are often proprietary. Therefore, the contract between the principal and the agent must take into account moral hazard.

In this paper, I study the design of an optimal contract for a principal (she) that incentivizes the agent (he) to gather information and maximize its precision. The agent conducts a location-scale experiment that is centered around the state and controls the precision (the inverse of the scale) at a cost. Eventually, the state is revealed and the contract can depend on the state and the agent's report. I abstract from the principal's broader decision problem—that is, the way in which the principal uses the information generated by the agent—and focus on the problem where the principal's objective is only to maximize precision subject to some fixed budget constraint on transfers. My analysis and results shall be robust to whatever the broader decision problem may be. The budget reflects the maximum amount of transfer that the principal is able to use.

In practice, a budget that can only be used for a specific task is commonly observed. For example, government agencies typically operate within budget limits imposed by government funding and must return any unused funds to the government.[1] There are numerous reports of unspent government funding being returned to the government, such as Mansouri (2024), Cuccia (2023).[2] In firms, it is also common

[1] Schick (2008) provides an in-depth look at how the federal budget is structured and constrained by funding limits imposed by the government.

[2] Federal regulations also specify where unspent grant money is returned, as seen at the National Archives (2024).

for agents to operate under a budget and to eventually return unused funds to the firm. Anthony et al. (2007) and Horngren (2009) provide comprehensive empirical evidence for this practice.[3] In the consulting example, the transfer could take the form of a promotion, a recommendation, or a grade. Sometimes, budgets are restricted by external funding designated for specific purposes, such as university grants.

In such settings, a simple contract is to allocate the entire budget to the agent if their report is close to the actual state, and withhold payment otherwise. I call these incentive schemes *cutoff transfers*. The cutoff transfer is the simplest transfer rule: It is straightforward to understand and easy to implement. The literature on principal-agent problems has long been interested in when simple contracts are optimal (Carroll, 2015; Herweg et al., 2010; Oyer, 2000; Gottlieb and Moreira, 2022).

My main result is that cutoff transfers are optimal across a large variety of settings. In particular, I identify a sufficient and necessary condition on the agent's information structure such that for every cost function, there exists an optimal transfer that is a cutoff transfer.

Formally, I study a model in which the principal wants to incentivize the agent to acquire information on an underlying state $\theta \in \mathbb{R}$. For simplicity, I assume the state θ admits an improper uniform common prior.[4] The agent can acquire a costly signal $s \in \mathbb{R}$ about the state. The signal takes the form $s = \theta + \frac{1}{\lambda}\varepsilon$ where ε is drawn from some symmetric and single-peaked probability density function, and λ is a measure of precision controlled by the agent. An important example is the case where ε admits a standard Gaussian distribution. The agent chooses precision λ at a cost $c(\lambda)$.

The principal observes neither the agent's signal s nor the agent's choice of precision λ. The agent makes a report $a \in \mathbb{R}$ after observing the signal. The principal wants to incentivize the agent to maximize precision and truthfully report his signal. She does so by a transfer rule t that depends on the report a and the state θ, which is eventually observed by both players.[5] I impose limited liability ($t \geq 0$) and limited budget. The principal has no use for the unspent savings. The agent's payoff is

[3] Anthony et al. (2007) offer an extensive analysis of organizational budget management, highlighting the frequent practice of returning unspent funds to a central or general fund in both private and public sectors. Horngren (2009) explores cost control and budget allocation, noting how unused resources are often returned to a central pool for future use.

[4] I generalize the model to the case where the state is multi-dimensional and the case with Gaussian prior.

[5] I study the case where the state is unobservable but the principal later receives a private signal about the state.

his expected utility from transfers minus the cost. I allow for any increasing utility function, which accommodates arbitrary risk attitude of the agent. I also allow for arbitrary lower-semicontinuous cost function.

I refer to a transfer rule as *optimal* if it induces the maximum precision among all transfers and elicits truthful reports. There are at least two difficulties in this problem. First, the transfer design problem is an infinite-dimensional optimization problem, as the transfer rule itself is a function. Second, there are two layers of incentive compatibility issues. The principal does not observe the precision chosen by the agent, which is a moral hazard problem, and she does not observe the signal either, leading to a elicitation problem.

I say that a transfer rule is a *cutoff* transfer if it pays the entire budget when the distance between the report and the state is below a cutoff d and pays 0 otherwise. For which signal distributions are cutoff transfers optimal?

My main result (Theorem 1) identifies a sufficient and necessary condition for cutoff transfers to be optimal. This condition is weaker than the monotone likelihood ratio property (Milgrom, 1981; Rogerson, 1985; Jewitt, 1988) and is satisfied by most common distributions, including Gaussian, Laplace, logistic, and the uniform distribution. Theorem 1 shows that if this condition holds, then for all cost functions cutoff transfers are optimal. Conversely, if the condition does not hold, then there is a cost function for which no cutoff transfer is optimal. Moreover, this cost function is not pathological and can be taken to be increasing, convex, and differentiable. Furthermore, when the condition holds, I characterize the optimal cutoff (Theorem 2).

To see why only part of the error distribution enters the condition, consider raising the reward at a small forecast error. If that error becomes more likely as precision rises, the extra reward favors the target precision over lower precision choices. One can therefore pay the full prize throughout the range of small errors with this property. Among the remaining errors, increasing elasticity orders outcomes by how strongly they favor higher precision. Moving rewards toward the more accurate forecasts then yields an interval. Notably, the proofs rely on techniques from monotone comparative statics. This allows me to analyze that problem under weak assumptions on the cost function and signal distribution. In particular, I do not need to assume the validity of the first-order approach (Rogerson, 1985; Jewitt, 1988).

All the results generalize to the n-dimensional case. The signal distribution is still

symmetric, single-peaked, and centered around θ. Here the symmetry means that the density of the distribution depends only on the Euclidean distance from the state θ. A cutoff transfer pays 1 when the Euclidean distance between the report and the state is less than a cutoff d. In this setting I show that an analogous result holds (Proposition 1).

I also generalize my results to the case of a proper Gaussian prior and Gaussian signals. I show that for all cost functions, cutoff transfers are optimal (Proposition 2).

In addition, I extend my results to the setting where the state is unobservable. Instead, the principal privately receives a signal about the state. When designing the transfer, she uses her signal instead of the state to discipline the agent. I show that given a Gaussian or uniform prior and Gaussian signal, cutoff transfers are optimal (Proposition 3).

Lastly, I examine what happens when the principal values savings from the unused budget. In Section 6, she maximizes the benefit of precision minus the expected transfer, subject to the same hard budget constraint, with a risk neutral agent. The problem can be decomposed into minimizing the expected transfer for each target precision and then choosing the target. A transfer rule is *optimal* for a given precision target if it minimizes the expected transfer among all transfer rules and induces truthful report. Theorem 3 identifies a sufficient and necessary condition for cutoff transfers to be optimal. This condition is different from the one from precision-maximization and is also weaker than the monotone likelihood property. When this condition holds, for every continuously differentiable, increasing, and convex cost function and every implementable precision target, cutoff transfers are optimal. Conversely, if this condition does not hold, then there is a linear cost function for which no cutoff transfer is optimal. Notably, the intersection of the condition for precision maximization (Theorem 1) and the condition for expected transfer minimization (Theorem 3) is exactly MLRP. I also characterize the implementable targets, optimal cutoff, and the minimum expected transfer (Theorem 4).

1.1 Related Literature

My paper is related to the literature on incentivizing information acquisition (see, e.g., Osband, 1989; Li and Libgober, 2023; Li et al., 2022; Neyman et al., 2021; Zermeno,

2011; Carroll, 2019; Chen et al., 2023; Chade and Kovrijnykh, 2016; Whitmeyer and Zhang, 2023; Sharma et al., 2023; Clark and Reggiani, 2021).[6]

In Li et al. (2022), an agent exerts a binary level of effort to refine a posterior from a prior. The paper studies optimizing proper scoring rules by maximizing the increase in the score with effort. In Li and Libgober (2023), a principal hires an agent to learn about a binary state. The agent acquires information over time through a Poisson information arrival technology. The principal rewards the agent with a fixed-value prize as a function of the agent's sequence of reports and the state. Li and Libgober (2023) identify conditions under which it is without loss to elicit a single report after all the information has been acquired.

Osband (1989) studies a principal with a quadratic prediction cost who incentivizes an expert to collect information. The expert can increase the precision (the inverse of variance) of the prediction at a constant cost. The principal minimizes the sum of the expected error variance and the expected transfer. In this stylized setting, the optimal transfer consists of a quadratic report error term plus a linear term on an initial belief error, with three parameters.

In Whitmeyer and Zhang (2023) and Sharma et al. (2023), a rational inattentive agent can acquire information flexibly subject to a posterior separable cost. The principal wants to minimize the expected monetary cost of implementing a given information structure. Similarly, Clark and Reggiani (2021) study the Pareto optimal contract that maximizes social welfare. In all these papers, both the experiment and the signal are unobservable to the principal. In Rappoport and Somma (2017), a principal hires an agent to acquire costly information to influence a third party's decision. This paper assumes that the realized piece of information is observable and contractible.

Zermeno (2011) and Papireddygari and Waggoner (2022) study menu design with information acquisition. In their setting, the principal first offers a menu of contracts. Then the agent privately acquires costly information. Next, the agent selects a contract from the menu. The selection therefore reveals some information the agent acquired.

In Carroll (2019), the principal is uncertain about the expert's information acqui-

[6]In Neyman et al. (2021), a rational expert aims to predict the probability of a biased coin flip. The information is acquired by choosing the number of flip trials at a fixed cost per flip. Neyman et al. (2021) study the optimal scoring rule that incentivizes precision.

sition technology and only knows some experiments that the agent can choose. The principal evaluates the incentive contract by a worst-case criterion.

Argenziano et al. (2016) and Kreutzkamp (2023) study costly information acquisition and transmission. In both papers, there is no transfer and the agent cares about the principal's action. In Kreutzkamp (2023) setting, the sender publicly chooses an experiment. In Argenziano et al. (2016), the expert acquires information by choosing a number of binary trials to perform.

In the classic principal-agent model, several papers also show the optimality of the cutoff transfer but rely on different assumptions. Oyer (2000) assumes the validity of the first-order approach, the existence of the optimal contract, and the absence of the IR constraint. He shows that cutoff transfers are optimal among monotone contracts for a risk-neutral agent. Herweg et al. (2010) show that cutoff transfers are optimal for expectation-based loss-averse and risk-neutral agent.

The extension in Section 6 also connects this paper to the classic principal-agent literature. Grossman and Hart (1983) decompose the principal's problem into two stages: minimizing the expected cost of implementing each action, subject to global incentive compatibility and participation, and then choosing the action that maximizes expected benefits net of implementation costs. In their finite-outcome model, the implementation problem can be formulated as a convex program. Their analysis shows that MLRP alone need not imply a monotone optimal transfer. Jewitt et al. (2008) study implementation costs and optimal contracts under upper and lower bounds on payments.

2 The Model

I study a principal-agent model, where the principal (she) wants to incentivize the agent (he) to acquire information regarding an underlying state $\theta \in \mathbb{R}$. They share a common prior over θ. It will be convenient to assume the prior to be an improper uniform prior on $\mathbb{R}$. I generalize my results to the case in which the state is multi-dimensional in Section 5.1, and I study the case of a proper Gaussian prior in Section 5.2.

The agent can acquire a costly signal $s \in \mathbb{R}$ of the form

$$s = \theta + \frac{1}{\lambda}\varepsilon,$$

where ε is drawn from a distribution with a symmetric and single-peaked probability density function (PDF) ϕ. I assume ϕ is continuously differentiable. The function ϕ is supported on an interval, which can be bounded or unbounded. The parameter λ is a measure of the signal's precision, a scale parameter that is inversely proportional to the standard deviation of s. The PDF of the signal, given θ and λ, is denoted by $\varphi(\cdot;\theta,\lambda)$. It takes the form

$$\varphi(x;\theta,\lambda) = \lambda\phi(\lambda(x-\theta)).$$

For example, if ϕ is the PDF of the standard Gaussian distribution, then $\varphi(\cdot;\theta,\lambda)$ is the PDF of a Gaussian distribution with mean θ and standard deviation $1/\lambda$. Note that $\phi(x) = \varphi(x;0,1)$.

The agent chooses the precision λ at a cost $c(\lambda)$. For example, suppose the agent's signal is the aggregation of many small independent signals. Then as the number of small signals becomes large, the aggregate signal tends to a Gaussian distribution. If the agent incurs a cost that depends on the number of small signals he gathers, then the agent effectively controls the standard deviation of the aggregate signal at a cost. I assume that the cost function c is lower semicontinuous.[7] The pair (ϕ, c) constitutes the primitive of the model.

The principal observes neither the agent's signal s nor the agent's choice of precision λ. Instead, after observing the signal, the agent sends a report $a \in \mathbb{R}$ to the principal.[8] Eventually both players observe the state θ.[9] The principal wants to incentivize the agent to maximize precision and truthfully report his signal. She does so by means of a transfer. The transfer can depend on the state and the agent's report. The assumption that the realized state is contractible is familiar from the literature on belief elicitation via (proper) scoring rules and prediction markets, ubiquitous in the principal-expert literature, and well-suited to economic applications in which the state is publicly observable ex post (e.g., the temperature tomorrow, the electricity demand next week, the outcome of an election being forecast by a pollster, or the

[7] It would be natural to also assume that the cost function is increasing—that is, the higher the precision, the higher the cost. Imposing this assumption or not does not affect my results.

[8] It is without loss for the agent to report his signal only. Asking the agent to report the precision additionally would only introduce cheap talk. The agent shall report the precision that results in the largest expected transfer.

[9] I relax this assumption in Section 5.3 where instead of eventually observing the state, they eventually observe a signal about the state.

conditions of a new market being analyzed by a consultant).

I assume that the transfer t is a function of the difference $\theta - a$ and it vanishes at infinity. That is, it satisfies

$$\lim_{x\to-\infty} t(x) = \lim_{x\to\infty} t(x) = 0.$$

As I discuss below, the assumption that the transfer depends only on the difference between $a - \theta$ is to ensure the agent's payoff is well-defined. I also assume limited liability, i.e., that the principal cannot take money from the agent, so that $t \geq 0$. In addition, the principal has a limited budget that does not depend on the state. She has no use for the unspent budget.[10] I consider the case where the principal values savings in Section 6. Without loss, I take the budget to be 1.

The timing of moves is as follows:

1. The principal commits to a transfer rule $t\colon \mathbb{R} \to [0, 1]$.
2. The agent chooses to accept or not. The game continues if the agent accepts and terminates otherwise.
3. The agent privately chooses a signal precision λ.
4. The agent privately observes a signal realization s drawn from the distribution $\varphi(\cdot; \theta, \lambda)$.
5. The agent makes a report $a \in \mathbb{R}$.
6. Finally, the state θ is revealed and the agent receives the transfer $t(\theta - a)$.

The agent's payoff is given by the expected utility of the transfer minus the cost of acquiring information:

$$\mathbb{E}u(t) - c(\lambda),$$

where $u\colon [0, 1] \to \mathbb{R}$ is strictly increasing and continuous, and normalized so that $u(0) = 0$, and $u(1) = 1$.[11] To ease the notation, it will be convenient to treat the

[10] Notably, there is a growing literature that adopts the same assumption that the principal cannot take money from the agent and can only reward the agent with a prize for which they have no other uses; see Li and Libgober (2023); Li et al. (2022); Deb et al. (2018, 2023); Dasgupta (2023); Hébert and Zhong (2022); Wong (2023).

[11] Here I can drop the continuity assumption on u and only require u to be weakly increasing. All my results still hold.

transfer rule as paying in utils, rather than money. Under my assumptions on u, this is equivalent, and without loss of generality, to assuming u is the identity function. Finally, I assume the agent has an outside option which gives payoff 0 if he rejects the contract.

After choosing λ, the agent observes a signal. Conditional on a signal realization s, it follows from Bayes' rule (adapted to a uniform prior) that the agent's posterior has PDF $\varphi(\cdot; s, \lambda)$. The agent now chooses a report a to maximize the expected transfer, which is therefore given by[12]

$$E(\lambda; t) = \max_{a \in \mathbb{R}} \int_{\mathbb{R}} t(\theta' - a)\varphi(\theta'; s, \lambda) d\theta'.$$

Maximizing over report a shows that when choosing the precision, the agent anticipates how he optimally reports later. Thus, the model allows for double deviation. The assumption that t vanishes at infinity ensures that the maximum is well defined.[13]

Note that as t depends only on the difference $\theta - a$, the expected transfer computed above does not depend on s.[14] Consequently, the agent's (interim) expected transfer conditional on signal s does not depend on s, which must coincide with his unconditional expected transfer. If t depends arbitrarily on a and θ, the agent's unconditional expected transfer might not be well-defined, since the distribution of s is improper. In Appendix B, I show that when the agent's expected transfer is well-defined, it is without loss to assume that the transfer depends only on the difference $a - \theta$. An interesting fact is that almost all popular key performance indicators in the forecast industry depend only on the prediction error $a - \theta$ (Vandeput, 2021).

The uniform prior assumption is mainly a translation-invariance device: together with the location-scale signal structure, it makes the posterior distribution of the prediction error independent of the realized signal, up to the precision parameter. This invariance allows for a sharp characterization of the optimal contract. With an arbitrary prior, the posterior error distribution would generally depend on the realized signal, so the optimal contract would typically be state-dependent, which is

[12] I set $E(0; t) = 0$, which agrees with the limit as precision tends to zero.

[13] I prove this in Lemma 5 in the appendix.

[14] To see this, note that

$$\max_a \int_{\mathbb{R}} t(\theta' - a)\varphi(\theta'; s, \lambda) d\theta' = \max_a \int_{\mathbb{R}} t(\theta' - a + s)\varphi(\theta'; 0, \lambda) d\theta' = \max_a \int_{\mathbb{R}} t(\theta' - a)\varphi(\theta'; 0, \lambda) d\theta'$$

beyond the scope of this paper. I return to proper Gaussian priors in Section 5.2, where conjugacy restores an effective location-scale representation.

The principal's objective is to maximize λ under the constraint of inducing a truthful report from the agent. (As I show later, this constraint will not be binding.) This objective can be interpreted as capturing, in reduced-form, settings in which the principal uses the information conveyed by the agent's report to solve some (unmodeled) decision problem. Indeed, under broad conditions, it is optimal for the principal to maximize precision regardless of the decision problem that she faces. First, when the density ϕ is strongly unimodal, for a large class of monotone decision problems (Karlin and Rubin, 1956), higher precision is always better for the principal (see Lehmann, 2011, Theorem 5.1 and 5.2).[15][16] Second, when the distribution ϕ is *self-decomposable*, the signal with different precisions are ranked in the Blackwell order; thus the principal is better off with higher precision for all decision problems. The self-decomposable distributions include all stable distributions (such as Gaussian and Cauchy distributions) and some non-stable ones such as Laplace.[17]

3 Preliminary Analysis

First, I define a relaxed problem, in which the principal designs a transfer to maximize the agent's signal precision, without incentivizing truthful reporting.

$$\begin{aligned} \underset{\lambda,\, t}{\text{maximize}} \quad & \lambda \\ \text{subject to} \quad & 0 \leq t \leq 1 \\ & \lambda \in \arg\max E(\cdot; t) - c(\cdot) \qquad \text{IC} \\ & E(\lambda; t) - c(\lambda) \geq 0 \qquad\quad \text{IR} \end{aligned} \tag{1}$$

The first constraint in the relaxed problem is limited liability and limited budget. The second constraint is an incentive compatibility constraint, since the precision λ

[15] ϕ is strongly unimodal if $-\ln \phi$ is convex. This is slightly stronger than global increasing elasticity.

[16] The class of *monotone* decision problems is defined in terms of the action space and the permissible loss functions. For each θ, there is a correct action $A(\theta)$. The function $A(\theta)$ is real-valued and nondecreasing. The range of $A(\theta)$ is the action space. The loss function is minimized at the correct action and is nondecreasing as the action moves away from the correct action on either side.

[17] See Goel and DeGroot (1979) and Lehmann (2011) for a detailed discussion.

is chosen by the agent and is unobserved by the principal and cannot be contracted upon.[18] The last constraint is a participation constraint, also known as individual rationality, and is implied by the outside option available to the agent. A transfer rule t is *optimal* if it solves problem (1) and induces the agent to report truthfully.

A quantity that will be key to the analysis is the elasticity of the standardized signal distribution ϕ.

Definition 1. The *elasticity* η of signal distribution ϕ at $x > 0$ is defined as

$$\eta(x) = -\frac{d\phi(x)/\phi(x)}{dx/x}.$$

For x such that $\phi(x) = 0$, I let $\eta(x) = +\infty$.

Elasticity measures by how much a percentage change in x leads to a percentage change in the density function $\phi(x)$. For example, if ϕ is the PDF of a standard Gaussian, then $\eta(x) = x^2$. If ϕ is the PDF of a standard Laplace, then $\eta(x) = x$. Notice that both examples have weakly increasing elasticity. In fact, weakly increasing elasticity is satisfied by most common nonatomic distributions defined on an interval, including the uniform distribution, the triangular distribution, and the logistic distribution. The next lemma shows that the weakly increasing elasticity condition is equivalent to a monotone likelihood property: for any $0 \leq x_1 \leq x_2$, the ratio $\frac{\varphi(x_1;0,\lambda)}{\varphi(x_2;0,\lambda)}$ increases in λ. That is, a more accurate signal (x_1) is more likely to appear than a less accurate signal (x_2) as precision increases.

Lemma 1.

$$\frac{\partial[\ln \phi(\lambda x)]}{\partial \ln \lambda} = -\eta(\lambda x).$$

In particular, for any $0 \leq x_1 \leq x_2$,

$$\frac{\varphi(x_1;0,\lambda)}{\varphi(x_2;0,\lambda)} \text{ increases in } \lambda \quad \Leftrightarrow \quad \eta(\lambda x_1) \leq \eta(\lambda x_2).$$

Global increasing elasticity is sufficient but not necessary; only the tail beyond the point where elasticity exceeds 1 matters. Formally, I define a condition weaker than MLRP.

[18] The continuity of $E(\cdot;t)$ and the lower semicontinuity of c ensures that the arg max in IC is well defined under very weak assumptions. For instance, the set of λ is compact or $\lim_{\lambda\to+\infty} c(\lambda) > 1$.

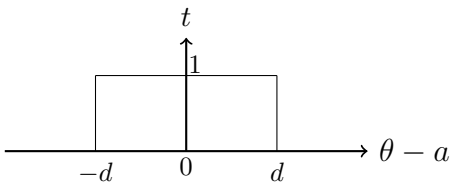

Figure 1: Cutoff Transfer.

Definition 2. The signal distribution satisfies *increasing elasticity above 1* if

$$x \longmapsto \max\{\eta(x), 1\}$$

is weakly increasing.

I define a useful quantity.

$$\eta^{-1}(1) = \inf\{x \in \mathbb{R}_+ | \eta(x) > 1\}.$$

Note that $\eta^{-1}(1)$ is well-defined, as the density function ϕ is integrable.[19] Later on, when I generalize the state and the signal to be n dimensional, $\eta^{-1}(n)$ is defined similarly.

4 Characterization of the Optimal Transfer

A simple transfer rule is the *cutoff transfer* that pays the agent 1 when the distance between the report and the state is less than a cutoff d and pays 0 otherwise (see Figure 1). As cutoff transfers are symmetric and single-peaked, they have the additional desirable property of inducing truthful reports by the agent.

As transfer rules are functions, the principal faces an infinite dimensional optimization problem. In my main result I show that when the standardized distribution ϕ satisfies increasing elasticity above one—which holds for all commonly used signal distributions—the principal chooses a cutoff transfer rule, reducing the problem to a one-dimensional one. Moreover, I show that this property of ϕ is also necessary for cutoff transfer rules to always be optimal.

[19] Notice that function $1/x$ is not integrable over any neighborhood around 0. Thus, $\eta(0^+) < 1$ since ϕ is integrable over any neighborhood around 0. Moreover, function $1/x$ is not integrable over any neighborhood around $+\infty$. Thus, $\eta(x) > 1$ must hold for some $x \in \mathbb{R}_+$.

Theorem 1. *The following are equivalent.*

1. *For every cost function, there exists an optimal transfer that is a cutoff transfer.*
2. *For every increasing, convex, and continuously differentiable cost function, there exists an optimal transfer that is a cutoff transfer.*
3. *Signal density function ϕ satisfies increasing elasticity above* 1.

Note that Theorem 1 holds even if we assume that cost functions are differentiable, convex, and increasing; that is, the result is not driven by considering ill-behaved cost functions.

Once we verify the condition in statement (3), then we do not need to worry about the infinite dimensional optimization problem, as cutoff transfers shall be optimal. Moreover, the condition is very easy to check. We only need to compute the elasticity and check if it is monotone in some region. A direct implication of Theorem 1 is that for every cost function, there exists an optimal transfer that is a cutoff transfer when the signal distribution is Gaussian, since the elasticity of a Gaussian distribution is increasing. The lower bound 1 in the statement is the dimension of the problem. So far both the state and the report are one-dimensional. In Section 5.1, where I generalize this result to n-dimensional signal and states, the dimension n will replace 1.

Note that there is a unique solution—which must be a cutoff transfer—if we additionally impose three mild assumptions.[20]

1. Signal density function ϕ satisfies strictly increasing elasticity above[21] 1.
2. The cost function is continuously differentiable.
3. The optimal precision is an interior solution.

In the following, I shall first illustrate the sufficiency part (i.e., why (3) implies (1)). Second, I optimize over cutoff transfers and characterize the optimal cutoff. Third, I provide intuition for the necessity part (i.e., why (2) implies (3)).

[20] I prove this uniqueness result at the end of the Proof of Theorem 1.

[21] Density function ϕ satisfies *strictly increasing elasticity above 1* if $\eta(\cdot)$ single-crosses 1 from below, $\{x|\eta(x) = 1\}$ is a singleton, and is strictly increasing after the cross.

I define the agent's expected payoff given transfer t as

$$\pi(t) = \max_{\lambda} E(\lambda; t) - c(\lambda)$$

and the agent's choice of precision given t

$$\boldsymbol{\lambda}(t) = \begin{cases} \max[\arg\max_{\lambda} E(\lambda; t) - c(\lambda)] & \text{if } \pi(t) \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

When analyzing cutoff transfers, with slight abuse of notations, I use $E(\lambda; d)$ to represent the expected transfer at precision λ given the cutoff transfer with cutoff d, i.e., $E(\lambda; d) = E(\lambda; \mathbf{1}_{|\theta - a| \leq d})$. Similarly, I denote by $\pi(d)$ and $\boldsymbol{\lambda}(d)$ the expected payoff and choice of precision for the cutoff transfer d, respectively.

4.1 Cutoff Transfers Are Optimal

I explain why increasing elasticity above 1 implies that for all cost functions, cutoff transfers are optimal. Suppose that increasing elasticity above 1 holds. I show that for any transfer rule t, there exists a cutoff transfer that induces a weakly larger precision and truthful report.

Here, for ease of exposition, I assume that η is weakly increasing and consider a transfer rule t that is symmetric in $\theta - a$ and weakly decreases in $|\theta - a|$. Given such a transfer rule, the agent reports truthfully and chooses the precision $\boldsymbol{\lambda}(t)$. I construct a cutoff transfer by choosing a cutoff d such that

$$E(\boldsymbol{\lambda}(t); d) = E(\boldsymbol{\lambda}(t); t). \tag{2}$$

Such a d exists because $E(\boldsymbol{\lambda}(t); d)$ increases in d from 0 to 1 (as the budget is 1). See Figure 2 for an example.

Next, I argue that

$$\begin{aligned} &E(\lambda; d) - E(\lambda; t) \geq 0 \text{ for } \lambda > \boldsymbol{\lambda}(t) \\ &E(\lambda; d) - E(\lambda; t) \leq 0 \text{ for } \lambda < \boldsymbol{\lambda}(t). \end{aligned} \tag{3}$$

This is depicted in Figure 3.

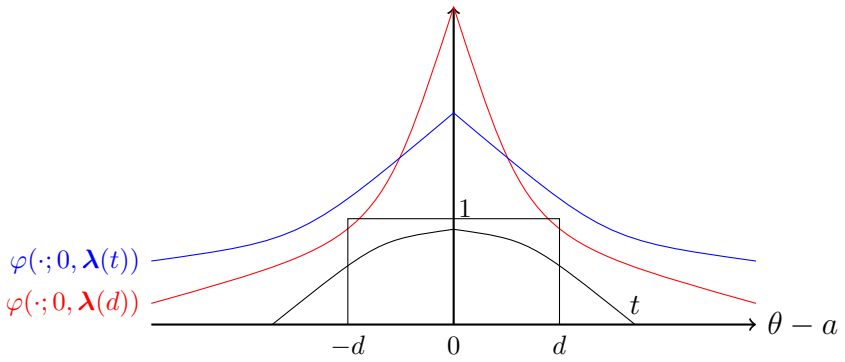

Figure 2: The Transfer t and The Cutoff Transfer d. The given transfer t and the matching cutoff transfer d are shown in black. The wider distribution shown in blue is the signal distribution chosen by the agent under t, while the narrower (more precise) distribution in red is the one chosen under d.

To see this, note that as the cutoff transfer pays the entire budget 1 when $|\theta - a| \leq d$, it is larger than the transfer t within the cutoff region. In addition, Lemma 1 shows that, since the elasticity of ϕ is weakly increasing, for any $0 < x_1 < d < x_2$ the ratio

$$\frac{\varphi(x_1;0,\lambda)}{\varphi(x_2;0,\lambda)}$$

increases in λ. Since $E(\lambda;t)$ is the integral of the transfer with respect to the signal distribution, (3) follows. As a result, we have $\boldsymbol{\lambda}(d) \geq \boldsymbol{\lambda}(t)$ (see Figure 3).

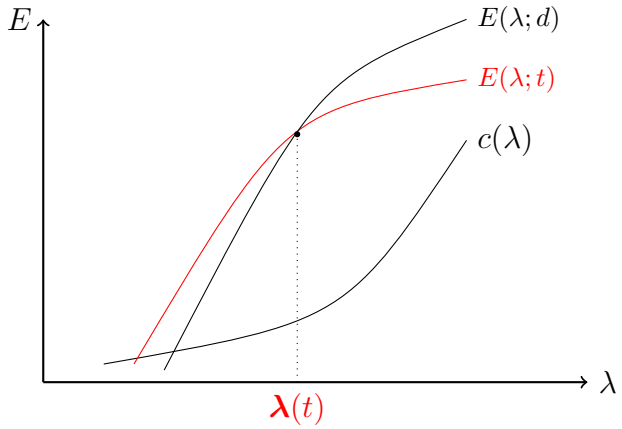

Figure 3: Expected Transfer.

To complete the proof of sufficiency in the general setting, there are two more technical issues. First, for a general transfer t (which is not necessarily symmetric),

the agent's report depends on both the signal and the precision. But in our previous example (Figure 3), the agent's report equals the signal and does not depend on the precision. Second, the proof sketch above assumed globally increasing elasticity, which is stronger than increasing elasticity above 1. To see why only part of the error distribution enters the condition, consider raising the reward at a small forecast error. For a nonzero error in the interior of the support,

$$\frac{\partial \log \varphi(x; 0, \lambda)}{\partial \log \lambda} = 1 - \eta(\lambda|x|).$$

If $|x| < \eta^{-1}(1)/\boldsymbol{\lambda}(t)$, the density at x weakly increases as precision rises to $\boldsymbol{\lambda}(t)$ from any lower level. Maximizing transfer throughout this region therefore adds at least as much expected payment at $\boldsymbol{\lambda}(t)$ as at every lower precision, which strengthens incentives. The proof in the appendix formally deals with both issues.

4.2 Optimization of Cutoff Transfers

When ϕ satisfies increasing elasticity above 1, Theorem 1 reduces the infinite dimensional optimization problem to a one-dimensional problem of finding the value of the optimal cutoff. In this section, I will solve this one-dimensional problem and identify the optimal cutoff.

As the cutoff d increases, the agent's expected transfer $E(\cdot; d)$ increases. Let $\bar{d}$ denote the minimum cutoff such that IR constraint holds, i.e.,

$$\bar{d} = \min\{d \geq 0 | \pi(d) \geq 0\}.$$

Notice that $\bar{d}$ depends on the cost function c. The continuity of $E(\cdot; d)$ for all d and the lower semicontinuity of c ensures that this minimum is well defined.

Theorem 2. *If ϕ satisfies increasing elasticity above 1, the cutoff transfer*

$$d^* = \min\{d \geq \bar{d} | \boldsymbol{\lambda}(d) d \geq \eta^{-1}(1)\}$$

is optimal.

To obtain this result, the following lemma is crucial.

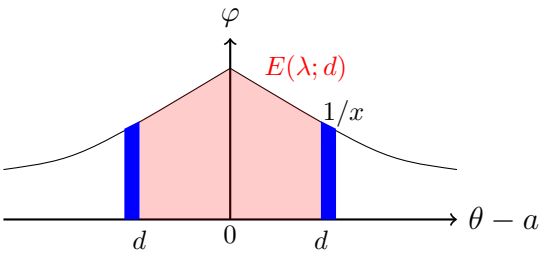

Figure 4: Increment of Expected Transfer When Increasing λ.
The agent slightly increases precision from λ to $\lambda + \Delta\lambda$. The area of the red region is the expected transfer $E(\lambda; d)$. The area of two blue regions is the increment of probability that the signal lies within the cutoff region.

Lemma 2 (Complements or Substitutes)**.** *The expected transfer $E(\lambda; d)$ satisfies:*

$$\frac{\partial^2 E(\lambda; d)}{\partial \lambda \partial d} \geq 0 \quad \Leftrightarrow \quad \eta(\lambda d) \leq 1.$$

This lemma characterizes whether the precision and the cutoff are complements or substitutes. When the cross derivative $\frac{\partial^2 E(\lambda;d)}{\partial \lambda \partial d}$ is positive, increasing the cutoff would increase the marginal return to the precision; thus they are complements. On the other hand, when the cross derivative is negative, the cutoff and the precision are substitutes.

I illustrate the intuition of the boundary case when $\eta(\lambda d) = 1$ and show that

$$\frac{\partial^2 E(\lambda; d)}{\partial \lambda \partial d} = 0.$$

I plot the agent's the expected transfer in Figure 4. Since the agent receives transfer 1 when the signal lies in the cutoff, the expected transfer $E(\lambda; d)$ is the probability that the signal lies in the cutoff, which is the area of the light red region. Suppose that the agent slightly increases precision from λ to $\lambda + \Delta\lambda$. The signal shall lie in the cutoff with a higher probability. The area of dark blue region is the incremental probability that the signal lies in the cutoff. Since $\eta(\lambda d) = 1$, in a neighborhood of d, $\varphi(\cdot; 0, \lambda)$ behaves like x^{-1}. We can compute the incremental expected transfer, which coincides with the incremental probability, by the area of the dark blue region

$$E(\lambda + \Delta\lambda; d) - E(\lambda; d) \approx 2\varphi(d; 0, \lambda) \left(d - d\frac{\lambda}{\lambda + \Delta\lambda} \right)$$

which is independent of d as $\varphi(\cdot; 0, \lambda)$ decays at x^{-1}. As a result, $\frac{\partial^2 E(\lambda;d)}{\partial \lambda \partial d} = 0$. A similar argument shows that $\frac{\partial^2 E(\lambda;d)}{\partial \lambda \partial d}$ and $\eta(\lambda d) - 1$ have the opposite sign. Here 1 appears as this problem is one dimensional, which is also why "increasing elasticity above 1" appears in Theorem 1 and $\eta^{-1}(1)$ appears in Theorem 2. Later on when I generalize this problem to n-dimensional, n shows up in both characterizations.

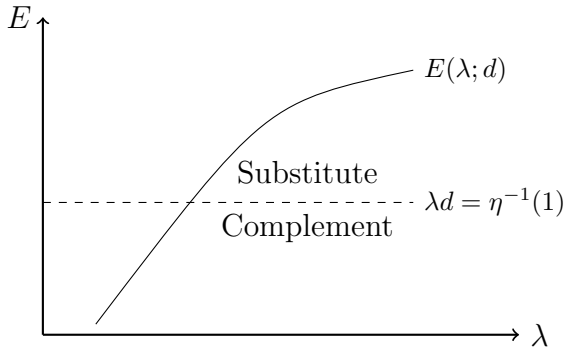

Figure 5: Expected Transfer

As the expected transfer $E(\lambda; d)$ is the probability that a signal lies within the cutoff region, we have $E(\lambda; d) = 2\Phi(\lambda d) - 1$ where Φ is the CDF of the distribution ϕ. Consequently, $E(\lambda; d)$ is increasing in λd. Since $\frac{\partial^2 E(\lambda;d)}{\partial \lambda \partial d}$ and $\eta(\lambda d) - 1$ have the opposite sign, by the definition of $\eta^{-1}(1)$, we have

$$\frac{\partial^2 E(\lambda; d)}{\partial \lambda \partial d} \geq 0 \quad \text{if} \quad \lambda d \leq \eta^{-1}(1)$$

$$\frac{\partial^2 E(\lambda; d)}{\partial \lambda \partial d} \leq 0 \quad \text{if} \quad \lambda d \geq \eta^{-1}(1).$$

Thus, the graph of expected transfer is separated into two parts (see Figure 5). When $\lambda d \geq \eta^{-1}(1)$, $\frac{\partial^2 E(\lambda;d)}{\partial \lambda \partial d} \leq 0$, I refer to this part as the *substitute region*. When $\lambda d < \eta^{-1}(1)$, $\frac{\partial^2 E(\lambda;d)}{\partial \lambda \partial d} > 0$, I refer to this part as the *complement region*.

Recall that $\bar{d}$ denotes the minimum cutoff that the agent accepts the contract. If $(\bar{d}, \boldsymbol{\lambda}(\bar{d}))$ lies in the substitute region, increasing d shall reduce $\boldsymbol{\lambda}(d)$, which is not optimal for the principal. Consequently, she should set the cutoff at the minimum $\bar{d}$. In this case, the IR constraint is binding and the agent's surplus is zero. If $(\bar{d}, \boldsymbol{\lambda}(\bar{d}))$ lies in the complement region, increasing d shall first increase $\boldsymbol{\lambda}(d)$, and $\boldsymbol{\lambda}(d)d$ shall increase until it hits the boundary $\eta^{-1}(1)$. After hitting the boundary, if we further increase d, then $\boldsymbol{\lambda}(d)$ shall decrease as $(d, \boldsymbol{\lambda}(d))$ enters the substitute region. Thus,

the optimal cutoff is the d where $\boldsymbol{\lambda}(d)d$ first hits the boundary $\eta^{-1}(1)$. In this case, the IR constraint is relaxed and the agent enjoys some surplus.

Note that although optimal precision λ^* is unique by design, the optimal cutoff transfer is not necessarily unique. To see this, recall that changing d causes the curve $E(\cdot; d)$ to rotate. If the cost function c has a kink and $\boldsymbol{\lambda}(d)$ is stuck at this kink, rotating $E(\cdot; d)$ slightly might not affect $\boldsymbol{\lambda}(d)$, which gives rise to multiple optimal cutoff transfers. The cutoff d^* in Theorem 2 is the optimal cutoff transfer that provides the strongest local incentive around λ^*, i.e., it has the largest derivative $\frac{\partial E(\lambda;d)}{\partial \lambda}\big|_{\lambda=\lambda^*}$ among all optimal cutoff transfers. But this example is not generic due to the kink in the cost function. This is also why a continuously differentiable cost function helps us to get a unique optimal transfer rule.

In Appendix A, I conduct comparative statics analysis to study how the optimal cutoff changes with the cost function.

4.3 The Necessity of Increasing Elasticity above 1

In this section, I explain the intuition behind the necessity result—(2) implies (3)—in Theorem 1. That is, why increasing elasticity above 1 is necessary for a cutoff transfer to always be optimal.

I first provide an example where ϕ does not satisfy increasing elasticity above 1. In this case, for some increasing and differentiable cost function, all cutoff transfers are suboptimal.

Example 1. The standardized signal distribution is truncated $\exp(1/x)$:

$$\phi(x) = \begin{cases} k\exp(1/\epsilon), & \text{if } |x| \in [0, \epsilon), \\ k\exp(1/|x|), & \text{if } |x| \in [\epsilon, 1], \\ 0, & \text{otherwise,} \end{cases}$$

where k is a normalizing factor (see the left panel of Figure 6).[22] I plot the elasticity of ϕ on the right panel of Figure 6. The signal distribution does not satisfy increasing elasticity above 1 as $\eta(x)$ is decreasing for $x \in [\epsilon, 1]$. Fix a pair (λ^*, d^*) such that $\epsilon < \lambda^* d^* < 1$. Suppose an increasing cost function $c \in C^1$ is tangent to $E(\cdot; d^*)$ at

[22] We can smooth this density function near ϵ and the boundary without changing the local elasticity pattern.

$\lambda = \lambda^*$ and $c(\lambda)$ is strictly above $E(\lambda; d^*)$ for all $\lambda \neq \lambda^*$ (see the right panel of Figure 7).

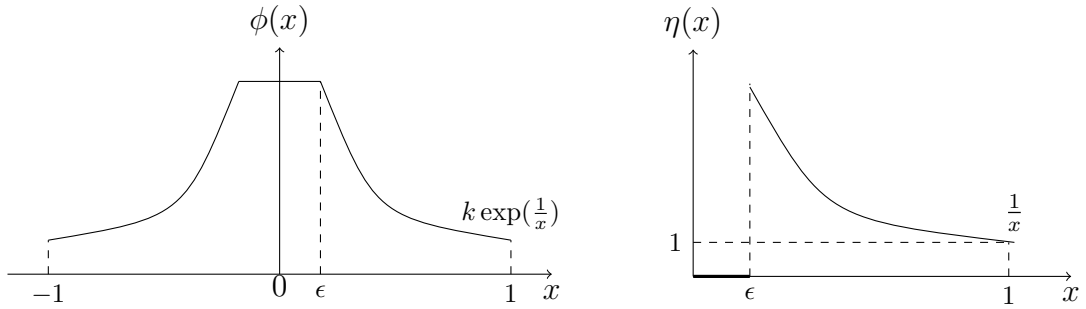

Figure 6: Signal Distribution and Elasticity

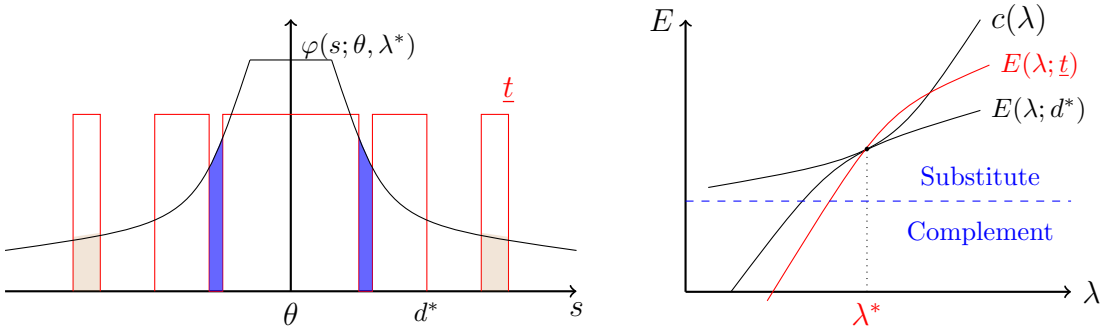

Figure 7: The New Transfer Rule and The Expected Transfer

First, note that d^* is the best cutoff among all cutoff transfers: by design, d^* is the minimum cutoff for the agent to work. By Lemma 2, as (λ^*, d^*) lies in the substitute region, any cutoff transfer with cutoff $d > d^*$ induces a smaller precision. Thus, d^* is the best cutoff transfer.

Second, I construct a new transfer rule $\underline{t}$ which induces a strictly larger precision. Starting with the best cutoff transfer rule d^*, I modify the transfer rule by setting one interior region (dark blue) within the cutoff to 0 and one exterior region (light brown) outside the cutoff to 1 (see the left panel of Figure 7). These two regions are chosen such that the area of the dark blue region and the light brown region are the same. Note that the area of each region is the expected transfer contributed by this region at precision λ^*. Thus, the new transfer rule $\underline{t}$ has the same expected transfer as cutoff transfer d^* at precision λ^*.

As the signal distribution features decreasing elasticity, the density function $\varphi(s;\theta,\lambda^*)$ is steeper at the blue region. Once we increase precision, the blue area shall shrink more than the brown area. This implies that the new transfer rule $\underline{t}$ has higher expected transfer than the cutoff transfer d^* for slightly higher precision. I translate this comparison to the right panel of Figure 7. For higher precision $\lambda > \lambda^*$, $E(\lambda;\underline{t}) > E(\lambda;d^*)$. Thus, $\underline{t}$ shall induce a higher precision.

This example is key to understand more generally the necessity of increasing elasticity above 1. I call a pair (λ^*,d^*) *exposed* if the cutoff transfer d^* is the optimal transfer for some increasing convex cost function $c \in C^1$ and induces precision λ^*. A pair (λ^*,d^*) being exposed implies that I cannot perturb the optimal cutoff transfer d^* to generate a larger incentive for the agent to increase the precision. That is,

$$\left.\frac{\partial E(\lambda;d^*)}{\partial \lambda}\right|_{\lambda=\lambda^*}$$

cannot be increased by any perturbations of cutoff transfer d^*. Otherwise, any small increment shall be detected by the continuously differentiable cost function c and reflected in a larger optimal precision $\arg\max_\lambda E(\lambda;t)-c(\lambda)$. This explains why only smooth cost functions are needed as testing functions.

There are at least two ways to perturb the cutoff transfer d^*

- First, I can slightly increase the cutoff (without affecting IR). This implies that (λ^*,d^*) must be substitutes,

$$\left.\frac{\partial^2 E(\lambda;d)}{\partial\lambda\partial d}\right|_{(\lambda^*,d^*)} \leq 0 \quad \Leftrightarrow \quad \eta(\lambda^* d^*) \geq 1.$$

  Otherwise, increasing the cutoff shall lead to a precision larger than λ^*.

- Second, I can alter the cutoff transfer d^* as in Figure 7. This implies that $\eta(x_1) \leq \eta(x_2)$ for all $0 \leq x_1 \leq \lambda^* d^* \leq x_2$. Otherwise, I can use the construction in Example 1 to construct a new transfer that induces a strictly larger precision.

When for every increasing, convex, continuously differentiable cost function, there exists an optimal transfer that is a cutoff transfer. Exposing any pair (λ^*,d^*) with $\eta(d^*\lambda^*) \geq 1$ is not hard, as I can always find an increasing, convex, continuously differentiable cost function that is tangent to $E(\cdot;d^*)$ at λ^* and strictly above $E(\cdot;d^*)$

everywhere else, like in Figure 7. This implies that all pairs (λ^*, d^*) with $\eta(d^*\lambda^*) \geq 1$ are exposed. Then the second bullet point above implies increasing elasticity above 1. The first bullet point above implies that all pairs in the complement regions are not exposed. Consequently, we do not have monotonicity for η when it is below 1.

5 Extensions

5.1 The Multi-Dimensional Case

My results can generalize to the case with a multi-dimensional state. In this section, suppose the state, the signal, and the report are n-dimensional, i.e., θ, s, $a \in \mathbb{R}^n$. The signal distribution is still symmetric, single-peaked, and centered around θ. Here symmetry means that the density of the distribution depends only on the Euclidean distance from the state θ. A cutoff transfer pays 1 when the Euclidean distance between the report and the state is less than a cutoff d. The next proposition shows that a statement analogous to that of Theorem 1 applies in this case too.

Proposition 1. *Suppose the state is n-dimensional. The following statements are equivalent.*

1. *For every cost function, there exists an optimal transfer that is a cutoff transfer.*
2. *For every increasing, convex, and continuously differentiable cost function, there exists an optimal transfer that is a cutoff transfer.*
3. *Density function ϕ satisfies increasing elasticity above n.*

Moreover, if ϕ satisfies increasing elasticity above n, the cutoff transfer

$$d^* = \min\{d \geq \bar{d} | \boldsymbol{\lambda}(d)d \geq \eta^{-1}(n)\}$$

is optimal.

Note that the increasing elasticity condition of Theorem 1 is generalized in this proposition to increasing elasticity above the dimension n, providing an explanation for why the number 1 appeared in Theorem 1.

5.2 Gaussian Prior and Gaussian Signal

In this section, I assume that both the prior and the signal admit Gaussian distributions. Since the prior is no longer uniform, a slight adjustment is that now the principal wants the agent to truthfully report the posterior mean rather than the signal.

Let the prior distribution be $\mathcal{N}(0, 1/\lambda_0^2)$. Conditional on a signal s, the agent's posterior is also Gaussian

$$\mathcal{N}(\frac{s\lambda^2}{\lambda_0^2+\lambda^2}, \frac{1}{\lambda_0^2+\lambda^2}),$$

where λ is the agent's choice of precision. I let Λ denote the precision of the posterior

$$\Lambda(\lambda) = \sqrt{\lambda_0^2+\lambda^2}.$$

Then, the expected transfer becomes $E(\Lambda(\lambda); t)$ instead of $E(\lambda; t)$. Therefore, $\Lambda(\lambda)$ shall replace λ in the characterization. Other than this difference, the results remain the same as in previous sections, as the posterior is Gaussian whose elasticity is increasing.

Proposition 2. *Suppose the prior and the signal admit Gaussian distribution. There exists an optimal transfer that is a cutoff transfer. Moreover, the cutoff transfer*

$$d^* = \min\{d \geq \bar{d} | \Lambda(\boldsymbol{\lambda}(d))d \geq \eta^{-1}(1)\}$$

is optimal.

5.3 Unobserved State

In this section, I extend my results to the setting where the state is unobservable. Instead, the principal has some private information about the state. When designing the transfer, she uses her private information instead of the state to discipline the agent.[23]

I assume that the prior is uniform or Gaussian $\mathcal{N}(0, 1/\lambda_0^2)$. The principal wants the agent to truthfully report his posterior mean and maximizes λ. The principal and

[23]Note that the principal does not have utility for their budget, and so have no incentive to lie about their private signals. Thus, it is without loss of generality to assume that the principal has commitment power.

the agent receive independent Gaussian signals: $s_p \sim \mathcal{N}(\theta, 1/\lambda_p^2)$ and $s \sim \mathcal{N}(\theta, 1/\lambda^2)$. I show that it is without loss to focus on cutoff transfers that pay 1 when $|s_p - a|$ is less than a cutoff and 0 otherwise.

Proposition 3. *Suppose that the state is unobservable and the prior admits uniform or Gaussian distribution. The principal and the agent receive independent Gaussian signals. There exists an optimal transfer that is a cutoff transfer. Moreover, the cutoff transfer*

$$d^* = \min\{d \geq \bar{d} | \Lambda(\boldsymbol{\lambda}(d))d \geq \eta^{-1}(1)\}$$

is optimal where $\Lambda(\boldsymbol{\lambda}(d)) = (\frac{1}{\lambda_p^2} + \frac{1}{\boldsymbol{\lambda}^2(d)})^{-\frac{1}{2}}$ *for the uniform prior and* $\Lambda(\boldsymbol{\lambda}(d)) = (\frac{1}{\lambda_p^2} + \frac{1}{\boldsymbol{\lambda}^2(d)+\lambda_0^2})^{-\frac{1}{2}}$ *for the Gaussian prior.*

6 When the Principal Values Savings

So far, I have assumed that the principal does not value the unused budget. What happens if the principal also cares about savings? As a robustness check, I return to the baseline model and allow the principal to trade off the benefit of precision against the expected transfer, while retaining the hard budget constraint. The principal thereby maximizes the gross benefit minus the expected transfer and wants to induce truthful report. Let $U(\lambda)$ denote the principal's gross benefit from precision, where U is increasing. The agent is risk neutral, so that transfers are measured in money. The signal distribution and the class of admissible transfers are unchanged. As in Section 3, I first relax the requirement that the transfer induce truthful reporting:

$$\begin{aligned} \underset{\lambda,\, t}{\text{maximize}} \quad & U(\lambda) - E(\lambda; t) \\ \text{subject to} \quad & 0 \leq t \leq 1 \\ & \lambda \in \arg\max E(\cdot; t) - c(\cdot) \quad \text{IC} \\ & E(\lambda; t) - c(\lambda) \geq 0 \quad \text{IR} \end{aligned} \tag{4}$$

This problem can be decomposed into a two-step optimization problem. The first

step is to minimize the expected transfer for each fixed target $\bar{\lambda}$:

$$\begin{aligned} \underset{t}{\text{minimize}} \quad & E(\bar{\lambda}; t) \\ \text{subject to} \quad & 0 \leq t \leq 1 \\ & \bar{\lambda} \in \arg\max E(\cdot; t) - c(\cdot) \quad \text{IC} \\ & E(\bar{\lambda}; t) - c(\bar{\lambda}) \geq 0 \quad \text{IR} \end{aligned} \tag{5}$$

The second step is to choose precision to maximize the gross benefit minus this minimum expected transfer. As before, a transfer is optimal for a given target if it solves the relaxed problem (5) and induces truthful reporting. I call a target *implementable* if this problem has a feasible transfer.[24] Let $C(\bar{\lambda})$ denote the infimum in (5), with $C(\bar{\lambda}) = +\infty$ if the target is not implementable.

I take $\lambda \in [0, \infty)$. Different from the baseline model, I impose the following restriction on the cost function.

Assumption 1. The cost function belongs to

$$\mathcal{C} = \left\{ c \in C^1([0,\infty)) : c(0) = 0, \quad c \text{ is nondecreasing and convex} \right\}.$$

Here, $c(0) = 0$ ensures that participation follows from the costless zero-precision option.

Definition 3. The signal distribution satisfies *increasing elasticity below 1* if

$$x \longmapsto \min\{\eta(x), 1\}$$

is weakly increasing.

Thus, elasticity is weakly increasing until it first reaches 1 and never subsequently falls below 1. The condition imposes no monotonicity restriction on elasticity once it remains above 1. The following theorem is a counterpart of Theorem 1 for a principal who minimizes the cost of implementing a given precision.

Theorem 3. *The following are equivalent.*

[24] Implementation in this section means that the target belongs to the agent's set of best responses, as specified by IC. It need not be the largest best response selected by $\boldsymbol{\lambda}(t)$ in Section 4. The cutoff characterized below uniquely implements a positive target with positive marginal cost if ϕ is strictly decreasing on the positive half of its support, or if c is strictly convex.

1. *For every cost function $c \in \mathcal{C}$ and every implementable target $\bar{\lambda}$, there exists an optimal transfer that is a cutoff transfer.*

2. *For every increasing linear cost function $c(\lambda) = k\lambda$ with $k > 0$ and every implementable target $\bar{\lambda}$, there exists an optimal transfer that is a cutoff transfer.*

3. *Signal density function ϕ satisfies increasing elasticity below* 1.

The economic force behind this result is the marginal incentive provided by each expected dollar of payment. At an error x in the positive density support, the ratio of the change in density to the density itself is

$$\left.\frac{\partial\varphi(x;0,\lambda)/\partial\lambda}{\varphi(x;0,\lambda)}\right|_{\lambda=\bar{\lambda}} = \frac{1-\eta(\bar{\lambda}|x|)}{\bar{\lambda}}. \tag{6}$$

A small increase in the payment near x raises the expected transfer in proportion to the denominator and raises the marginal return to precision in proportion to the numerator. Their ratio therefore measures how efficiently a payment rewards precision. Increasing elasticity below 1 says that the positive part of this ratio is weakly decreasing in the absolute error. Payments on errors closer to zero provide at least as much positive marginal incentive per expected dollar. With a unit payment cap, the principal can minimize expenditure by paying the entire prize on a central interval.

This also explains why only elasticity below 1 matters. When $\eta(\bar{\lambda}|x|) \geq 1$, payment near x does not provide a positive marginal incentive for precision. How these nonpositive contributions are ordered is immaterial for the least expensive way to provide a positive incentive. Conversely, if increasing elasticity below 1 fails, a more distant error can provide a strictly higher positive marginal incentive per expected dollar than a closer error. A small reallocation of payments can then preserve the incentive at the target while reducing expenditure. The proof shows that such a reallocation preserves global feasibility for a suitable linear cost function. Thus, the necessity result already holds within the simplest class of convex costs.

The sufficiency proof proceeds in two steps. First, a necessary marginal first-order condition yields a lower bound on the expected payment of *any* transfer implementing the target. Increasing elasticity below 1 orders positive marginal incentives per expected dollar. Then I verify a cutoff transfer can attain this bound by concentrating payments on errors closest to zero. Second, I verify that this cutoff satisfies the global

incentive compatibility constraint. Thus, the proof combines a payment lower bound for arbitrary feasible transfers with a separate verification of global feasibility for the candidate cutoff.

Compared with Theorem 1, this characterization uses three additional restrictions. First, risk neutrality is now substantive. In the precision-maximization problem, I can express transfers in utils without changing the principal's objective. When she values monetary savings, a nonlinear utility function u would instead make the monetary cost of a utility transfer $u^{-1}(t)$. The same normalization would therefore change her objective.

Second, convexity of costs ensures that the marginal condition for the proposed cutoff is sufficient for global incentive compatibility. For any cutoff d, function $E(\cdot; d)$ is concave. Consequently, $E(\cdot; d) - c(\cdot)$ is concave under Assumption 1. Thus, I do not need to assume the validity of the first-order approach (Rogerson, 1985; Jewitt, 1988). Unlike Theorem 1, this result does not cover all lower-semicontinuous cost functions.

Third, $c(0) = 0$ ensures that participation follows from the costless zero-precision option. In Theorem 1, the cost function space is strictly larger and does not impose $c(0) = 0$. Alternatively, we could also impose $c(0) = 0$ while allowing for arbitrary outside option utility $\bar{u} \geq 0$ in Theorem 1. Then the restriction in Theorem 3 is that the outside option utility is zero.

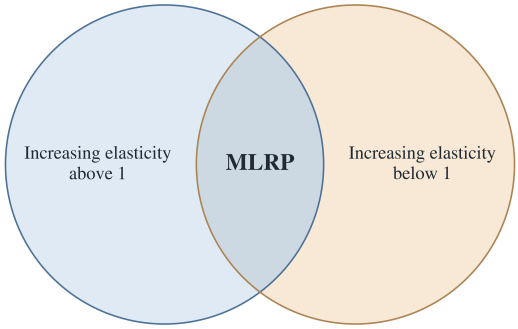

Figure 8: Relationship between the elasticity conditions. The intersection of increasing elasticity above 1 and increasing elasticity below 1 consists exactly of signal distributions satisfying MLRP.

Finally, the two objectives, maximizing precision and minimizing expected transfer, impose complementary restrictions on the signal distribution. Increasing elasticity

above 1 is equivalent to weak monotonicity of $\max\{\eta, 1\}$, whereas increasing elasticity below 1 is weak monotonicity of $\min\{\eta, 1\}$. Their intersection is exactly global weak monotonicity of η, since

$$\eta = \min\{\eta, 1\} + \max\{\eta, 1\} - 1.$$

By Lemma 1, global increasing elasticity is equivalent to MLRP. Hence MLRP is precisely the intersection of the distributional conditions for universal cutoff optimality in the two problems (Figure 8).

6.1 The Optimal Cutoff

Once the condition in Theorem 3 holds, the remaining problem is to determine the optimal cutoff. Define

$$G(z) = 2\Phi(z) - 1, \qquad H(z) = 2z\phi(z), \qquad z \geq 0. \tag{7}$$

The function G gives the expected transfer under a cutoff with standardized radius z, while H measures its marginal incentive for precision. In particular,

$$E(\lambda; d) = G(\lambda d), \qquad \frac{\partial E(\lambda; d)}{\partial \lambda} = \frac{H(\lambda d)}{\lambda}, \qquad \lambda > 0.$$

Define $z_{\bar{\lambda}}$ as

$$z_{\bar{\lambda}} = \inf\{z \geq 0 | H(z) = \bar{\lambda} c'(\bar{\lambda})\} \tag{8}$$

Theorem 4. *Suppose that ϕ satisfies increasing elasticity below 1. For any cost function $c \in \mathcal{C}$, a target $\bar{\lambda}$ is implementable if and only if*

$$\bar{\lambda} c'(\bar{\lambda}) \leq \max_{z \geq 0} H(z). \tag{9}$$

When this condition holds, the optimal cutoff and the minimum expected transfer are

$$d^*(\bar{\lambda}) = \frac{z_{\bar{\lambda}}}{\bar{\lambda}}, \qquad C(\bar{\lambda}) = G(z_{\bar{\lambda}}). \tag{10}$$

The feasibility condition compares the marginal incentive required at the target with the largest marginal incentive that the budget can provide. Under increasing

elasticity below 1,

$$H'(z) = 2\phi(z)[1 - \eta(z)]$$

on the positive density support. Hence H is strictly increasing until elasticity first reaches 1 and weakly decreasing afterward. The quantity $\max H$ is the resulting limit on scaled marginal incentives. The proof shows that no admissible transfer can exceed this limit, so the feasibility condition applies to all transfers.

Several cutoffs may provide the same marginal incentive. The smallest solution in (8) minimizes the probability of payment, since G is increasing on the positive support. Theorem 4 shows that the resulting cutoff minimizes expected payments among all admissible transfers, not just among cutoff transfers.

The second step of the principal's problem is therefore

$$\sup_{\lambda \geq 0:\, \lambda c'(\lambda) \leq \max_{z \geq 0} H(z)} \{U(\lambda) - C(\lambda)\}.$$

Thus, whenever increasing elasticity below 1 holds, the original contract design problem reduces to a scalar choice of precision.

References

Anthony, R. N., Govindarajan, V., Hartmann, F. G., Kraus, K., and Nilsson, G. (2007). *Management control systems*, volume 12. McGraw-Hill Boston.

Argenziano, R., Severinov, S., and Squintani, F. (2016). Strategic information acquisition and transmission. *American Economic Journal: Microeconomics*, 8(3):119–155.

Carroll, G. (2015). Robustness and linear contracts. *American Economic Review*, 105(2):536–563.

Carroll, G. (2019). Robust incentives for information acquisition. *Journal of Economic Theory*, 181:382–420.

Chade, H. and Kovrijnykh, N. (2016). Delegated information acquisition with moral hazard. *Journal of Economic Theory*, 162:55–92.

Chen, S., Wu, J., Wu, Y., and Yang, Z. (2023). Learning to incentivize information acquisition: Proper scoring rules meet principal-agent model. *arXiv preprint arXiv:2303.08613*.

Clark, A. and Reggiani, G. (2021). Contracts for acquiring information. *arXiv preprint arXiv:2103.03911*.

Cuccia, A. (2023). DC left millions unspent from federal grants to end homelessness. https://thedcline.org/2023/06/08/dc-left-millions-unspent-from-federal-grants-to-end-homelessness/.

Dasgupta, S. (2023). Optimal test design for knowledge-based screening. *Available at SSRN 4403119*.

Deb, R., Pai, M. M., and Said, M. (2018). Evaluating strategic forecasters. *American Economic Review*, 108(10):3057–3103.

Deb, R., Pai, M. M., and Said, M. (2023). *Indirect Persuasion*. Centre for Economic Policy Research.

Edlin, A. S. and Shannon, C. (1998). Strict monotonicity in comparative statics. *Journal of Economic Theory*, 81(1):201–219.

Goel, P. K. and DeGroot, M. H. (1979). Comparison of experiments and information measures. *The Annals of Statistics*, 7(5):1066–1077.

Gottlieb, D. and Moreira, H. (2022). Simple contracts with adverse selection and moral hazard. *Theoretical Economics*, 17(3):1357–1401.

Grossman, S. J. and Hart, O. D. (1983). An analysis of the principal-agent problem. *Econometrica*, 51(1):7–45.

Hébert, B. and Zhong, W. (2022). Engagement maximization. *arXiv preprint arXiv:2207.00685*.

Herweg, F., Müller, D., and Weinschenk, P. (2010). Binary payment schemes: Moral hazard and loss aversion. *American Economic Review*, 100(5):2451–77.

Horngren, C. T. (2009). *Cost accounting: a managerial emphasis*. Pearson Education India.

Jewitt, I. (1988). Justifying the first-order approach to principal-agent problems. *Econometrica: Journal of the Econometric Society*, pages 1177–1190.

Jewitt, I., Kadan, O., and Swinkels, J. M. (2008). Moral hazard with bounded payments. *Journal of Economic Theory*, 143(1):59–82.

Karlin, S. and Rubin, H. (1956). The theory of decision procedures for distributions with monotone likelihood ratio. *The Annals of Mathematical Statistics*, pages 272–299.

Kreutzkamp, S. (2023). Endogenous information acquisition in cheap-talk games. Technical report, Working paper.

Lehmann, E. L. (2011). Comparing location experiments. In *Selected Works of EL Lehmann*, pages 779–791. Springer.

Li, Y., Hartline, J. D., Shan, L., and Wu, Y. (2022). Optimization of scoring rules. In *Proceedings of the 23rd ACM Conference on Economics and Computation*, pages 988–989.

Li, Y. and Libgober, J. (2023). Optimal scoring for dynamic information acquisition. *Working Paper*.

Mansouri, K. (2024). Millions remain unspent in federal funds for homeless students — and time is running out. https://www.lpm.org/news/2024-07-16/millions-remain-unspent-in-federal-funds-for-homeless-students-and-time-is-running

Milgrom, P. R. (1981). Good news and bad news: Representation theorems and applications. *The Bell Journal of Economics*, pages 380–391.

National Archives (2024). Code of Federal Regulations: 886.20 What happens to unused funds from my grant? https://www.ecfr.gov/current/title-30/chapter-VII/subchapter-R/part-886/section-886.20.

Neyman, E., Noarov, G., and Weinberg, S. M. (2021). Binary scoring rules that incentivize precision. In *Proceedings of the 22nd ACM Conference on Economics and Computation*, pages 718–733.

Osband, K. (1989). Optimal forecasting incentives. *Journal of Political Economy*, 97(5):1091–1112.

Oyer, P. (2000). A theory of sales quotas with limited liability and rent sharing. *Journal of labor Economics*, 18(3):405–426.

Papireddygari, M. and Waggoner, B. (2022). Contracts with information acquisition, via scoring rules. In *Proceedings of the 23rd ACM Conference on Economics and Computation*, pages 703–704.

Rappoport, D. and Somma, V. (2017). Incentivizing information design. *Available at SSRN 3001416*.

Rogerson, W. P. (1985). The first-order approach to principal-agent problems. *Econometrica: Journal of the Econometric Society*, pages 1357–1367.

Schick, A. (2008). *The federal budget: Politics, policy, process*. Brookings Institution Press.

Sharma, S., Tsakas, E., and Voorneveld, M. (2023). Procuring unverifiable information. Technical report, Mimeo.

Vandeput, N. (2021). *Data science for supply chain forecasting*. de Gruyter.

Whitmeyer, M. and Zhang, K. (2023). Buying opinions. *arXiv preprint arXiv:2202.05249*.

Wong, Y. F. (2023). Dynamic monitoring design. *Available at SSRN 4466562*.

Zermeno, L. (2011). A principal-expert model and the value of menus. *Working Paper*.

A Comparative Statics

In this section, I study how the optimal cutoff changes with the cost function. Without loss of generality, I assume that the cost function is weakly increasing.[25] I express dependence on the cost function. Let $d^*(c)$ and $\lambda^*(c)$ denote the optimal cutoff and precision given cost function c. Consider two cost functions $c_1 \leq c_2$ where c_1 is less costly. Intuitively, if the agent's cost is higher, the optimal cutoff increases. The next result formalizes this intuition.

Proposition 4 (Comparative Statics)**.** *Suppose that ϕ satisfies increasing elasticity above 1 and $c_1 \leq c_2$. If $c_2(\lambda) - c_1(\lambda)$ is weakly increasing in λ, then $d^*(c_1) \leq d^*(c_2)$ and $\lambda^*(c_2) \leq \lambda^*(c_1)$. In particular, if $c_2 = kc_1$ for some constant $k > 1$, then $d^*(c_1) \leq d^*(c_2)$ and $\lambda^*(c_2) \leq \lambda^*(c_1)$.*

The cost difference $c_2(\lambda) - c_1(\lambda)$ being weakly increasing occurs if c_2 is more convex than c_1. Another interesting comparative statics is to change the budget from 1 to $1/k$. This is equivalent to keeping the budget at 1 and change the cost function from c_1 to kc_1. Thus, lowering the budget would lead to a larger optimal cutoff.

As a direct corollary, I can study what happens if the agent's signal is more precise. Suppose that $\varepsilon_2 = k\varepsilon_1$ with some constant $k > 1$. Recall that the agent's signal is $s = \theta + \frac{1}{\lambda}\varepsilon$. Thus, a smaller noise ε_1 corresponds to a more precise signal, which leads to a smaller cutoff by the next result.

Corollary 1. *Suppose that the cost function is weakly convex. If $\varepsilon_2 = k\varepsilon_1$ with some constant $k > 1$, $d_1^* \leq d_2^*$.*

What if the cost difference $c_2 - c_1$ is not increasing? Then the answer is more involved. Denote by $\boldsymbol{\lambda}(d; c)$ the induced precision given cutoff transfer d and the cost function c. Given the cost function c_1, the minimum cutoff $\bar{d}(c_1)$ and the induced precision $\boldsymbol{\lambda}(\bar{d}(c_1); c_1)$ pair lies either in the substitute region $\boldsymbol{\lambda}(\bar{d}(c_1); c_1)\bar{d}(c_1) \geq \eta^{-1}(1)$ or not. If this pair lies in the substitute region, then the optimal cutoff coincides with the minimum cutoff, $d^*(c_1) = \bar{d}(c_1)$, by Theorem 2. Then given a larger cost function c_2, the minimum cutoff must be larger, $\bar{d}(c_2) \geq \bar{d}(c_1)$, which entails a larger optimal cutoff, $d^*(c_2) \geq d^*(c_1)$.

[25] Given any cost function $\tilde{c}$ that is not weakly increasing, I can define a new cost function c to be the largest weakly increasing function that is below $\tilde{c}$. The agent's precision choice problem under cost function c is the same as under cost $\tilde{c}$, because $E(\lambda; d)$ is weakly increasing in λ.

However, if the $\boldsymbol{\lambda}(\bar{d}(c_1); c_1)$, $\bar{d}(c_1)$ pair lies in the complement region, the comparison of d^* can go both directions. If $c_2 - c_1$ is decreasing, it may render higher precision relatively more attractive to the agent under cost c_2 than c_1. This may induce a higher precision for each cutoff transfer. As the optimal precision and optimal cutoff are substitutes by Theorem 2, a higher induced precision leads to a lower optimal cutoff.

Readers may wonder what happens to the optimal precision λ^* given $c_2 \geq c_1$. It turns out that increasing the cost function can shift λ^* in both directions, depending on the shape of the cost difference $c_2 - c_1$.

B The Assumption That Transfer Depends Only on the Difference Between the State and Report

In this section, I show that given the uniform prior, it is without loss to assume that the transfer depends only on the difference $a - \theta$. To make the expected transfer well-defined for the uniform prior, I take two approaches. In the first approach, I define the agent's expected transfer to be conditional on the event that the signal lies in some set of finite Lebesgue measure. Then I show that the conclusion holds for any such set. The conditioning on some set is merely an artifact to deal with the improperness of unbounded uniform prior. In the second approach, I study an alternative setting where the state space is a circle in $\mathbb{R}^2$ and the prior is proper and uniform. This unit circle is similar to the one-point compactification of $\mathbb{R}$.

Throughout this section, I allow the transfer to depend on both the state and the report. To ensure that the agent's report is well-defined, I assume that the family of maps $a \mapsto t(y + a, a)$ parametrized by y is equicontinuous. This is trivially satisfied when the transfer depends only on $\theta - a$, as $t(y+a, a)$ stays constant as a varies. When the prior is improper uniform on $\mathbb{R}$, I additionally assume that $t(\theta, \theta + x)$ vanishes uniformly (in θ) as $|x|$ tends to infinity.

First, I show that the agent's report is well-defined.

Lemma 3. *The maximum*

$$\max_{a \in \mathbb{R}} \int_{\mathbb{R}} t(\theta', a) \varphi(\theta'; s, \lambda) d\theta'$$

is well-defined.

Proof of Lemma 3. Fix a signal s and precision λ. Let $T(a)$ be the agent's expected transfer after reporting a,

$$T(a) = \int t(\theta', a)\varphi(\theta'; s, \lambda)d\theta'.$$

If $T(a) = 0$ for all $a \in \mathbb{R}$, then the maximum is well-defined. Suppose that there exists $a_1 \in \mathbb{R}$ such that $T(a_1) > 0$. As $t(\theta, \theta + x)$ vanishes uniformly (in θ) as x tends to infinity, there exists a $K_1 > 0$ such that for all $|\theta - a| > K_1$, $t(\theta, a) < \frac{T(a_1)}{2}$. Moreover, there exists a $K_2 > 0$ such that $\Phi(\lambda K_2) > 1 - \frac{T(a_1)}{4}$. Let $K = \max(K_1, K_2)$. If $|a - s| > 2K$,

$$\begin{aligned}
T(a) =& \int t(\theta', a)\varphi(\theta'; s, \lambda)d\theta' \\
=& \int_{\theta':|\theta'-s|<K} t(\theta', a)\varphi(\theta'; s, \lambda)d\theta' + \int_{\theta':|\theta'-s|\geq K} t(\theta', a)\varphi(\theta'; s, \lambda)d\theta' \\
\leq& \frac{T(a_1)}{2} \int_{\theta':|\theta'-s|<K} \varphi(\theta'; s, \lambda)d\theta' + 1 \cdot \int_{\theta':|\theta'-s|\geq K} \varphi(\theta'; s, \lambda)d\theta' \\
\leq& \frac{T(a_1)}{2} + 2[1 - \Phi(\lambda K)] \\
=& T(a_1).
\end{aligned}$$

Thus,

$$\max_{a\in\mathbb{R}} T(a) = \max_{|a-s|\leq 2K} T(a).$$

Next, I show that $T(\cdot)$ is continuous.

$$\begin{aligned}
T(a+\epsilon) - T(a) =& \int t(\theta', a+\epsilon)\varphi(\theta'; s, \lambda)d\theta' - \int t(\theta', a)\varphi(\theta'; s, \lambda)d\theta' \\
=& \int [t(\theta', a+\epsilon) - t(\theta' - \epsilon, a)]\varphi(\theta'; s, \lambda)d\theta' + \int [t(\theta' - \epsilon, a) - t(\theta', a)]\varphi(\theta'; s, \lambda)d\theta' \\
=& \int [t(\theta', a+\epsilon) - t(\theta' - \epsilon, a)]\varphi(\theta'; s, \lambda)d\theta' + \int t(\theta', a)[\varphi(\theta' + \epsilon; s, \lambda) - \varphi(\theta'; s, \lambda)]d\theta'
\end{aligned}$$

$$
\begin{aligned}
|T(a+\epsilon)-T(a)| &\leq \sup_{\theta'} |t(\theta', a+\epsilon) - t(\theta'-\epsilon, a)| + \left| \int t(\theta', a)[\varphi(\theta'; s-\epsilon, \lambda) - \varphi(\theta'; s, \lambda)] d\theta' \right| \\
&\leq \sup_{\theta'} |t(\theta', a+\epsilon) - t(\theta'-\epsilon, a)| + \int |\varphi(\theta'; s-\epsilon, \lambda) - \varphi(\theta'; s, \lambda)| d\theta' \\
&\leq \sup_{\theta'} |t(\theta', a+\epsilon) - t(\theta'-\epsilon, a)| + 2 \int_{\theta' \leq s - \frac{\epsilon}{2}} \varphi(\theta'; s-\epsilon, \lambda) - \varphi(\theta'; s, \lambda) d\theta' \\
&\leq \sup_{\theta'} |t(\theta', a+\epsilon) - t(\theta'-\epsilon, a)| + 2[\Phi(\frac{\epsilon}{2}\lambda) - \Phi(-\frac{\epsilon}{2}\lambda)]
\end{aligned}
$$

Note that $\sup_{\theta'} |t(\theta', a+\epsilon) - t(\theta'-\epsilon, a)|$ tends to 0 as ϵ tends to 0 by the equicontinuity of t. Moreover, the CDF Φ is continuous. Thus, T is continuous. A continuous function achieves a maximum on a compact set. So the maximum is well-defined. $\square$

Next, I show that it is without loss to assume that the transfer depends only on the difference $\theta - a$. For a transfer rule $t(\theta - a)$ that depends only on the difference $\theta - a$, let $\hat{E}(\lambda; t)$ denote the agent's expected transfer when he has to report truthfully,

$$
\hat{E}(\lambda; t) = \int_{\mathbb{R}} t(\theta') \varphi(\theta'; 0, \lambda) d\theta'.
$$

Let $\hat{\boldsymbol{\lambda}}(t)$ denote the corresponding induced precision,

$$
\hat{\boldsymbol{\lambda}}(t) = \begin{cases} \max[\arg\max \hat{E}(\cdot; t) - c(\cdot)], & \text{if } \max \hat{E}(\cdot; t) - c(\cdot) \geq 0, \\ 0, & \text{otherwise.} \end{cases}
$$

For a transfer rule $t(\theta, a)$ that depends on both the state and the report, we need to integrate the signal when computing the expected transfer. Note that the signal distribution is also uniform. Fix any measurable set A with $0 < |A| < \infty$. I define the agent's expected transfer conditional on the signal lies in set A, i.e., $s \in A$

$$
E(\lambda; t) = \int_{s \in A} \frac{1}{|A|} ds \left(\max_a \int t(\theta', a) \varphi(\theta'; s, \lambda) d\theta' \right)
$$

$$
\boldsymbol{\lambda}(t) = \begin{cases} \max[\arg\max E(\cdot; t) - c(\cdot)], & \text{if } \max E(\cdot; t) - c(\cdot) \geq 0, \\ 0, & \text{otherwise.} \end{cases}
$$

Lemma 4. *For a transfer rule $t(\theta, a)$, there exists a transfer rule $\tilde{t}(\theta - a)$ that depends only on the difference $\theta - a$; given transfer rule $\tilde{t}$, the agent chooses precision $\boldsymbol{\lambda}(t)$*

when he is forced to reveal the signal. That is,

$$\boldsymbol{\lambda}(t) = \hat{\boldsymbol{\lambda}}(\tilde{t}).$$

Proof. Choose a measurable optimal reporting rule $\boldsymbol{a}(s;\lambda)$ after observing signal s at precision λ:

$$\boldsymbol{a}(s;\lambda) \in \arg\max_{a} \int_{\mathbb{R}} t(\theta', a)\varphi(\theta'; s, \lambda)d\theta'.$$

$$E(\lambda; t) = \int_{s\in A} \frac{1}{|A|} ds \int t(\theta', \boldsymbol{a}(s;\lambda))\varphi(\theta'; s, \lambda)d\theta'$$

Since the reporting rule is measurable, the integrand is bounded, and $0 < |A| < \infty$, this integral is well-defined. Let $\theta'' = \theta' - s$,

$$E(\lambda; t) = \int_{s\in A} \frac{1}{|A|} ds \int t(\theta'' + s, \boldsymbol{a}(s;\lambda))\varphi(\theta''; 0, \lambda)d\theta''.$$

Let $\tilde{t}(\theta'') = \int_{s\in A} \frac{1}{|A|} t(\theta'' + s, \boldsymbol{a}(s; \boldsymbol{\lambda}(t)))ds$. The transfer $\tilde{t}$ is measurable and takes values in $[0, 1]$. For each fixed s, the prediction error $x+s-\boldsymbol{a}(s; \boldsymbol{\lambda}(t))$ tends to infinity in absolute value as $|x| \to \infty$. The uniform vanishing assumption and dominated convergence therefore imply that $\tilde{t}(x) \to 0$. By the Fubini–Tonelli theorem,

$$\begin{aligned} E(\boldsymbol{\lambda}(t); t) &= \int_{s\in A} \frac{1}{|A|} ds \int t(\theta'' + s, \boldsymbol{a}(s; \boldsymbol{\lambda}(t)))\varphi(\theta''; 0, \boldsymbol{\lambda}(t))d\theta'' \\ &= \int_{\mathbb{R}} \tilde{t}(\theta'')\varphi(\theta''; 0, \boldsymbol{\lambda}(t))d\theta'' \\ &= \hat{E}(\boldsymbol{\lambda}(t); \tilde{t}). \end{aligned}$$

Moreover, at $\lambda \neq \boldsymbol{\lambda}(t)$,

$$\begin{aligned} E(\lambda; t) &= \int_{s\in A} \frac{1}{|A|} ds \int_{\mathbb{R}} t(\theta', \boldsymbol{a}(s;\lambda))\varphi(\theta'; s, \lambda)d\theta' \\ &\geq \int_{s\in A} \frac{1}{|A|} ds \int_{\mathbb{R}} t(\theta', \boldsymbol{a}(s; \boldsymbol{\lambda}(t)))\varphi(\theta'; s, \lambda)d\theta' \\ &= \int_{s\in A} \frac{1}{|A|} ds \int t(\theta'' + s, \boldsymbol{a}(s; \boldsymbol{\lambda}(t)))\varphi(\theta''; 0, \lambda)d\theta'' \\ &= \int_{\mathbb{R}} \tilde{t}(\theta'')\varphi(\theta''; 0, \lambda)d\theta'' \\ &= \hat{E}(\lambda; \tilde{t}) \end{aligned}$$

where the inequality follows by the definition of $\boldsymbol{a}(s;\lambda)$. By the definition of $\boldsymbol{\lambda}(t)$, $\hat{\boldsymbol{\lambda}}(\tilde{t}) = \boldsymbol{\lambda}(t)$. □

Given this lemma, I can use the constructed $\tilde{t}$ in the proof of Theorem 1, where I show that there exists a cutoff transfer that induces a weakly higher precision.

The same construction also applies to the payment-minimization problem in Section 6. Fix a positive target $\bar{\lambda}$ implemented by a transfer rule $t(\theta, a)$, and replace $\boldsymbol{\lambda}(t)$ by $\bar{\lambda}$ in the construction above. Denote the resulting transfer by $\tilde{t}$. Then

$$\hat{E}(\bar{\lambda};\tilde{t}) = E(\bar{\lambda};t), \qquad \hat{E}(\mu;\tilde{t}) \leq E(\mu;t) \quad \text{for all } \mu \geq 0.$$

Consequently, $\bar{\lambda}$ maximizes $\hat{E}(\mu;\tilde{t}) - c(\mu)$, and the expected payment at the target is unchanged. The transfer $\tilde{t}$ can therefore be used in the marginal-incentive and payment-bound arguments in the sufficiency proof of Theorem 3. Under increasing elasticity below 1, the cutoff constructed in that proof implements the same target, induces truthful reporting, and has an expected payment no greater than $E(\bar{\lambda};t)$. Thus, allowing transfers to depend separately on the state and the report does not reduce the minimum implementation cost.

In the second approach, the state space is a circle in $\mathbb{R}^2$ with a circumference of 1, and the prior distribution is uniform. This setup corresponds to cases where the state is periodic, such as the unit vector in $\mathbb{R}^2$ which corresponds to a direction in $\mathbb{R}^2$ (In Machina triangle, each direction corresponds to an indifference curve.), the time of day (e.g., 15:20) or the day of the year (e.g., October 19). The model remains almost the same. The only exception is that ϕ admits a compact support $[-M, M]$ and the precision $\lambda \in [2M, +\infty]$. The proof that it is without loss to assume that transfer depends only on the difference $\theta - a$ is almost the same as above. The exception is that the integral of s in Lemma 4 can be taken over the entire circle.

C Omitted Proofs of Section 3-5

Proof of Lemma 1.

$$\frac{\partial \ln \phi(\lambda x)}{\partial \ln \lambda} = \frac{\phi'(\lambda x)}{\phi(\lambda x)} \lambda x = -\eta(\lambda x).$$

Thus, we have

$$\frac{\partial[\ln \frac{\phi(\lambda x_1)}{\phi(\lambda x_2)}]}{\partial \ln \lambda} = \eta(\lambda x_2) - \eta(\lambda x_1).$$

The second part follows by

$$\frac{\partial[\frac{\varphi(x_1;0,\lambda)}{\varphi(x_2;0,\lambda)}]}{\partial \lambda} = \frac{\partial[\frac{\phi(\lambda x_1)}{\phi(\lambda x_2)}]}{\partial \lambda}.$$

□

Proof of Lemma 2. Let Φ denote the anti-derivative of ϕ, i.e., $\Phi(x) = \int_{-\infty}^{x} \phi(y)dy$. We compute the expected transfer given the cutoff transfer d

$$E(\lambda; d) = 2\Phi(\lambda d) - 1.$$

The derivative of the expected transfer with respect to λ is

$$\frac{\partial E(\lambda; d)}{\partial \lambda} = 2d\phi(\lambda d).$$

Let us see how the derivative changes when varying d,

$$\frac{\partial \frac{\partial E(\lambda;d)}{\partial \lambda}}{\partial d} = 2\phi(\lambda d)[1 - \eta(\lambda d)].$$

□

Lemma 5. *The agent's expected transfer*

$$E(\lambda; t) = \max_{a \in \mathbb{R}} \int_{\mathbb{R}} t(\theta' - a)\varphi(\theta'; s, \lambda)d\theta'$$

is well-defined.

Proof of Lemma 5. I show that the maximum is well-defined, given that the transfer rule t vanishes at infinity. Fix a signal s and precision λ. Let $T(a)$ be the agent's expected transfer after reporting a,

$$T(a) = \int t(\theta' - a)\varphi(\theta'; s, \lambda)d\theta'.$$

If $T(a) = 0$ for all $a \in \mathbb{R}$, the $\arg\max_a T(a)$ is well-defined. Suppose that there

exists $a_1 \in \mathbb{R}$ such that $T(a_1) > 0$. As t vanished at infinity, there exists a $K_1 > 0$ such that for all $|x| > K_1$, $t(x) < \frac{T(a_1)}{2}$. Moreover, there exists a $K_2 > 0$ such that $\Phi(\lambda K_2) > 1 - \frac{T(a_1)}{4}$. Let $K = \max(K_1, K_2)$. If $|a - s| > 2K$,

$$\begin{aligned}
T(a) =& \int t(\theta' - a)\varphi(\theta'; s, \lambda)d\theta' \\
=& \int_{\theta':|\theta'-s|<K} t(\theta' - a)\varphi(\theta'; s, \lambda)d\theta' + \int_{\theta':|\theta'-s|\geq K} t(\theta' - a)\varphi(\theta'; s, \lambda)d\theta' \\
\leq& \frac{T(a_1)}{2} \int_{\theta':|\theta'-s|<K} \varphi(\theta'; s, \lambda)d\theta' + 1 \cdot \int_{\theta':|\theta'-s|\geq K} \varphi(\theta'; s, \lambda)d\theta' \\
\leq& \frac{T(a_1)}{2} + 2[1 - \Phi(\lambda K)] \\
=& T(a_1).
\end{aligned}$$

Thus,

$$\arg\max_{a\in\mathbb{R}} T(a) = \arg\max_{|a-s|\leq 2K} T(a).$$

Next, I show that $T(\cdot)$ is continuous.

$$\begin{aligned}
|T(a+\epsilon) - T(a)| =& \left| \int t(\theta' - a - \epsilon)\varphi(\theta'; s, \lambda)d\theta' - \int t(\theta' - a)\varphi(\theta'; s, \lambda)d\theta' \right| \\
=& \left| \int t(\theta' - a)\varphi(\theta' + \epsilon; s, \lambda)d\theta' - \int t(\theta' - a)\varphi(\theta'; s, \lambda)d\theta' \right| \\
=& \left| \int t(\theta' - a)\varphi(\theta'; s - \epsilon, \lambda)d\theta' - \int t(\theta' - a)\varphi(\theta'; s, \lambda)d\theta' \right| \\
=& \left| \int t(\theta' - a)[\varphi(\theta'; s - \epsilon, \lambda) - \varphi(\theta'; s, \lambda)]d\theta' \right| \\
\leq& \int \left| \varphi(\theta'; s - \epsilon, \lambda) - \varphi(\theta'; s, \lambda) \right| d\theta' \\
=& 2\int_{\theta'\leq s-\frac{\epsilon}{2}} \varphi(\theta'; s - \epsilon, \lambda) - \varphi(\theta'; s, \lambda)d\theta' \\
=& 2[\Phi(\frac{\epsilon}{2}\lambda) - \Phi(-\frac{\epsilon}{2}\lambda)]
\end{aligned}$$

Thus, T is continuous as the CDF Φ is continuous. A continuous function achieves a maximum on a compact set. So the maximum in $E(\lambda; t)$ is well-defined. □

Let $\hat{E}(\lambda; t)$ denote the agent's expected transfer when he has to report truthfully

$$\hat{E}(\lambda; t) = \int_{\mathbb{R}} t(\theta')\varphi(\theta'; 0, \lambda)d\theta'.$$

Let $\hat{\boldsymbol{\lambda}}(t)$ denote the corresponding induced precision

$$\hat{\boldsymbol{\lambda}}(t) = \begin{cases} \max[\arg\max \hat{E}(\cdot; t) - c(\cdot)], & \text{if } \max \hat{E}(\cdot; t) - c(\cdot) \geq 0, \\ 0, & \text{otherwise.} \end{cases}$$

Lemma 6. *Suppose ϕ satisfies increasing elasticity above 1. Suppose the agent has to report truthfully. Given a symmetric transfer rule $\bar{t}$ and induced $\lambda^* = \hat{\boldsymbol{\lambda}}(\bar{t})$, the new transfer*

$$\bar{t}'(x) = \begin{cases} 1, & \text{if } \lambda^*|x| < \eta^{-1}(1), \\ \bar{t}(x), & \text{otherwise} \end{cases}$$

can induce a weakly higher precision $\hat{\boldsymbol{\lambda}}(\bar{t}') \geq \hat{\boldsymbol{\lambda}}(\bar{t})$.

Proof of Lemma 6. Let $\Delta t = \bar{t}' - \bar{t}$. I shall show that

$$\hat{E}(\lambda^*; \Delta t) \geq \hat{E}(\lambda; \Delta t) \quad \text{for all } \lambda \leq \lambda^*.$$

This implies that $\hat{\boldsymbol{\lambda}}(\bar{t}') \geq \hat{\boldsymbol{\lambda}}(\bar{t})$ since $\hat{\boldsymbol{\lambda}}(t) = \max[\arg\max \hat{E}(\cdot; t) - c(\cdot)]$.

Consider two cutoff transfers $0 < d_1 < d_2 \leq \eta^{-1}(1)/\lambda^*$. For all $\lambda \leq \lambda^*$, we have $\eta(\lambda d_1) \leq 1$, $\eta(\lambda d_2) \leq 1$. By Lemma 2,

$$\frac{\partial^2 E(\lambda; d)}{\partial \lambda \partial d} \geq 0 \quad \Leftrightarrow \quad \eta(\lambda d) \leq 1.$$

This implies $\frac{\partial E(\lambda; d)}{\partial \lambda}$ increases in $d \in [d_1, d_2]$ for all $\lambda \leq \lambda^*$, which implies $E(\lambda; d_2) - E(\lambda; d_1)$ increases in $\lambda \in [0, \lambda^*]$. As Δt is an integral of such difference of cutoff transfers and $\hat{E}(\lambda; t)$ is linear in t, $\hat{E}(\lambda; \Delta t)$ increases in $\lambda \in [0, \lambda^*]$.

□

Lemma 7. *Suppose ϕ satisfies increasing elasticity above 1. Let $\eta^{-1}(1) < x_1 < x_2$. Then*

$$\frac{\phi(\lambda' x_2)}{\phi(\lambda' x_1)} \geq \frac{\phi(x_2)}{\phi(x_1)} \quad \text{for all} \quad \lambda' < 1.$$

Proof of Lemma 7. By Lemma 1,

$$\ln \frac{\phi(\lambda' x)}{\phi(x)} = \int_{\lambda'}^{1} \eta(\lambda x) d \ln \lambda.$$

Then,

$$\frac{\phi(\lambda' x_1)}{\phi(x_1)} \leq \frac{\phi(\lambda' x_2)}{\phi(x_2)} \quad \text{for all} \quad \lambda' < 1$$

is equivalent to

$$\int_{\lambda'}^{1} \eta(\lambda x_1) d \ln \lambda \leq \int_{\lambda'}^{1} \eta(\lambda x_2) d \ln \lambda \quad \text{for all} \quad \lambda' < 1.$$

Let $y = \ln \lambda$, then it suffices to show

$$\int_{\ln \lambda'}^{0} \eta(x_1 \exp(y)) dy \leq \int_{\ln \lambda'}^{0} \eta(x_2 \exp(y)) dy \quad \text{for all} \quad \lambda' < 1.$$

Let $\Delta y = \ln(x_2/x_1) > 0$. Then it suffices to show

$$\int_{\ln \lambda'}^{0} \eta(x_2 \exp(y - \Delta y)) dy \leq \int_{\ln \lambda'}^{0} \eta(x_2 \exp(y)) dy \quad \text{for all} \quad \lambda' < 1.$$

Let $g(y) = \eta(x_2 \exp(y))$. Then it suffices to show

$$\int_{\ln \lambda'}^{0} g(y - \Delta y) dy \leq \int_{\ln \lambda'}^{0} g(y) dy \quad \text{for all} \quad \lambda' < 1. \tag{11}$$

Note that function $g(y)$ is increasing in $y \in [\ln \frac{\eta^{-1}(1)}{x_2}, 0]$ and single-crosses 1 from below at $\ln \frac{\eta^{-1}(1)}{x_2}$.

If $\ln \lambda' \geq -\Delta y$, then for all $y \in [\ln \lambda', 0]$,

$$g(y) \geq g(\ln \lambda') \geq g(-\Delta y) \geq g(y - \Delta y).$$

Suppose $\ln \lambda' < -\Delta y$. Equation (11) is

$$\int_{\ln \lambda' - \Delta y}^{-\Delta y} g(y) dy \leq \int_{\ln \lambda'}^{0} g(y) dy \quad \text{for all} \quad \lambda' < 1.$$

I can subtract $\int_{\ln\lambda'}^{-\Delta y} g(y)dy$ from both sides. Then the inequality becomes

$$\int_{\ln\lambda'-\Delta y}^{\ln\lambda'} g(y)dy \leq \int_{-\Delta y}^{0} g(y)dy \quad \text{for all} \quad \lambda' < 1.$$

It holds since for all $y \in [-\Delta y, 0]$, $y' \in [\ln\lambda' - \Delta y, \ln\lambda']$,

$$g(y) \geq g(-\Delta y) \geq g(\ln\lambda') \geq g(y').$$

□

Proof of Theorem 1 (3) implies (1) and Proof of Theorem 2. Suppose that ϕ satisfies increasing elasticity above 1. If there is no t such that $\boldsymbol{\lambda}(t) > 0$, then the problem is trivial.[26] Now suppose there exists a t such that $\boldsymbol{\lambda}(t) > 0$. In step 1, I shall show that for all transfer t with $\boldsymbol{\lambda}(t) > 0$, I can construct a new cutoff transfer d such that $\boldsymbol{\lambda}(d) \geq \boldsymbol{\lambda}(t)$. That is, cutoff transfer d weakly improves over t. In step 2, I shall optimize over cutoff transfers and prove Theorem 2. In step 3, I show uniqueness under stronger conditions.

Step 1: Optimality of Cutoff Transfers

I first provide a roadmap for this proof. First, I construct an auxiliary transfer $\tilde{t}$ that preserves the induced precision when the agent has to report truthfully. Second, I symmetrize $\tilde{t}$ to obtain $\bar{t}$. Third, I fill the low prediction-error region up to $\frac{\eta^{-1}(1)}{\boldsymbol{\lambda}(t)}$ to obtain $\bar{t}'$. Fourth, I replace $\bar{t}'$ by a cutoff with the same payoff at $\boldsymbol{\lambda}(t)$. Finally, I use Lemma 7 to compare incentives below $\boldsymbol{\lambda}(t)$.

Fix a transfer rule t. First, I show that there exists a transfer rule $\tilde{t}(\theta - a)$ such that given $\tilde{t}$, the agent chooses precision $\boldsymbol{\lambda}(t)$ when he has to report truthfully. That is,

$$\hat{\boldsymbol{\lambda}}(\tilde{t}) = \boldsymbol{\lambda}(t).$$

For the transfer rule t that depends on the state and the report, Lemma 4 gives us the desired $\tilde{t}$. Now assume that t depends only on the difference between the state and the report, $\theta - a$. The construction of $\tilde{t}$ is simpler and is as follows. Let $\boldsymbol{a}(s; \boldsymbol{\lambda}(t))$ denote the agent's report after observing signal s at the precision $\boldsymbol{\lambda}(t)$.

$$\boldsymbol{a}(s; \boldsymbol{\lambda}(t)) = \arg\max_{a} \int_{\mathbb{R}} t(\theta' - a)\varphi(\theta'; s, \boldsymbol{\lambda}(t))d\theta'.$$

[26] In this case, we adopt the convention that every transfer is optimal.

As both t and φ are translation invariant, $\boldsymbol{a}(s;\boldsymbol{\lambda}(t)) - s$ is a constant. I define a new transfer rule $\tilde{t}$ by

$$\tilde{t}(x) = t(x + s - \boldsymbol{a}(s;\boldsymbol{\lambda}(t))).$$

Given transfer $\tilde{t}$, the agent reports truthfully at $\boldsymbol{\lambda}(t)$. Note that replacing t by $\tilde{t}$ only changes the agent's report by a constant. Two transfers t and $\tilde{t}$ provide the same incentive for the agent to choose λ, i.e.,

$$E(\cdot;t) = E(\cdot;\tilde{t}), \quad \boldsymbol{\lambda}(\tilde{t}) = \boldsymbol{\lambda}(t).$$

I define an auxiliary problem: the agent has to reveal his signal. The agent's expected transfer is

$$\hat{E}(\lambda;\tilde{t}) = \int_{\mathbb{R}} \tilde{t}(\theta')\varphi(\theta';0,\lambda)d\theta'.$$

It must be true that for all $\lambda \in \mathbb{R}_+$,

$$\hat{E}(\lambda;\tilde{t}) \leq E(\lambda;\tilde{t}) \tag{12}$$

and

$$\hat{E}(\boldsymbol{\lambda}(t);\tilde{t}) = E(\boldsymbol{\lambda}(t);\tilde{t}). \tag{13}$$

The inequality holds as the agent loses the flexibility to misreport in the auxiliary problem. The equality holds as the agent reports truthfully at $\lambda = \boldsymbol{\lambda}(t)$. By Equation (12) and (13),

$$\hat{\boldsymbol{\lambda}}(\tilde{t}) = \boldsymbol{\lambda}(\tilde{t}) = \boldsymbol{\lambda}(t).$$

Then, I symmetrify the transfer rule $\tilde{t}$ to obtain

$$\bar{t}(x) = \frac{1}{2}[\tilde{t}(x) + \tilde{t}(-x)].$$

For all $\lambda \in \mathbb{R}_+$, we have

$$\hat{E}(\lambda;\bar{t}) = \hat{E}(\lambda;\tilde{t}).$$

$$\hat{\boldsymbol{\lambda}}(\bar{t}) = \hat{\boldsymbol{\lambda}}(\tilde{t}) = \boldsymbol{\lambda}(t).$$

By Lemma 6, I can augment transfer $\bar{t}$ to a new transfer $\bar{t}'$

$$\bar{t}'(x) = \begin{cases} 1, & \text{if } \boldsymbol{\lambda}(t)|x| < \eta^{-1}(1), \\ \bar{t}(x), & \text{otherwise} \end{cases}$$

and induces higher precision

$$\hat{\boldsymbol{\lambda}}(\bar{t}') \geq \hat{\boldsymbol{\lambda}}(\bar{t}) = \boldsymbol{\lambda}(t).$$

Then I can construct a cutoff transfer d. I pin down the cutoff by requiring that cutoff transfer d and $\bar{t}'$ offers the same expected transfer at $\boldsymbol{\lambda}(t)$, i.e.,

$$E(\boldsymbol{\lambda}(t); d) = \hat{E}(\boldsymbol{\lambda}(t); \bar{t}'). \tag{14}$$

Let $\Delta t(x) = 1_{|x| \leq d} - \bar{t}'(x)$ denote the transfer difference. The transfer difference must take the form

$$\Delta t(x) \begin{cases} = 0, & \text{if } |x| \leq \eta^{-1}(1)/\boldsymbol{\lambda}(t), \\ \geq 0, & \text{if } \eta^{-1}(1)/\boldsymbol{\lambda}(t) < |x| \leq d, \\ \leq 0, & \text{otherwise.} \end{cases}$$

Compare the expected transfer given cutoff transfer d and the expected transfer $\bar{t}'$ in the auxiliary problem.

$$E(\lambda; d) - \hat{E}(\lambda; \bar{t}') = \int_{\mathbb{R}} \Delta t(\theta') \varphi(\theta'; 0, \lambda) d\theta'.$$

Consider y_1, y_2 such that $\eta^{-1}(1)/\boldsymbol{\lambda}(t) < y_1 < y_2$. Let $x_1 = \boldsymbol{\lambda}(t) y_1$, $x_2 = \boldsymbol{\lambda}(t) y_2$. Then $\eta^{-1}(1) < x_1 < x_2$. By Lemma 7, for all $\lambda < \boldsymbol{\lambda}(t)$, we have

$$\frac{\phi(x_2 \lambda / \boldsymbol{\lambda}(t))}{\phi(x_1 \lambda / \boldsymbol{\lambda}(t))} \geq \frac{\phi(x_2)}{\phi(x_1)}.$$

$$\frac{\phi(\lambda y_2)}{\phi(\lambda y_1)} \geq \frac{\phi(\boldsymbol{\lambda}(t) y_2)}{\phi(\boldsymbol{\lambda}(t) y_1)}.$$

$$\frac{\varphi(y_2; 0, \lambda)}{\varphi(y_1; 0, \lambda)} \geq \frac{\varphi(y_2; 0, \boldsymbol{\lambda}(t))}{\varphi(y_1; 0, \boldsymbol{\lambda}(t))}.$$

This implies

$$E(\lambda; d) - \hat{E}(\lambda; \bar{t}') \leq 0, \text{ if } \lambda < \boldsymbol{\lambda}(t).$$

Given $\hat{\boldsymbol{\lambda}}(\bar{t}') \geq \boldsymbol{\lambda}(t)$ and the definition $\boldsymbol{\lambda}(t) = \max[\arg\max E(\cdot; t) - c(\cdot)]$, we have $\boldsymbol{\lambda}(d) \geq \boldsymbol{\lambda}(t)$ (see Figure 9).

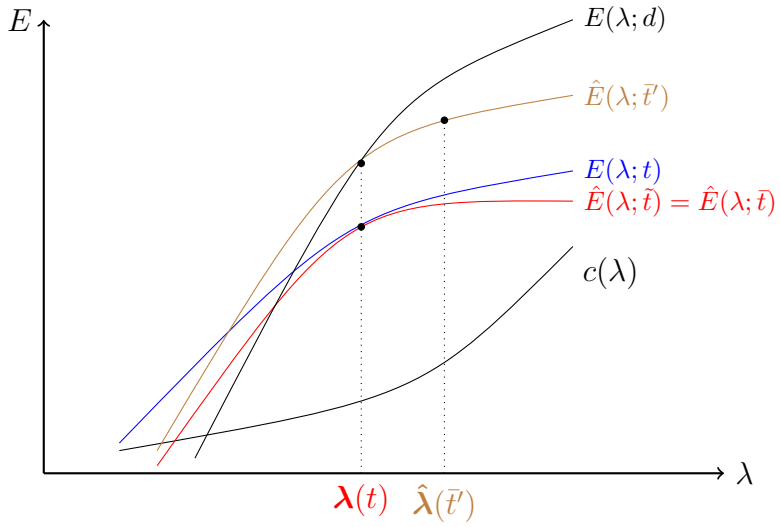

Figure 9: Expected Transfer

Step 2: Proof of Theorem 2

Next, I optimize over cutoff transfers to prove Theorem 2 and show the existence of the optimal cutoff transfer. If we set $d < \bar{d}$, the agent never chooses to work. So consider $d \geq \bar{d}$.

Case 1: $\boldsymbol{\lambda}(\bar{d})\bar{d} \geq \eta^{-1}(1)$.

Let $\tilde{\lambda} = \eta^{-1}(1)/\bar{d} \leq \boldsymbol{\lambda}(\bar{d})$. For any $d > \bar{d}$, if $\boldsymbol{\lambda}(d) \leq \tilde{\lambda}$, then $\boldsymbol{\lambda}(d) \leq \boldsymbol{\lambda}(\bar{d})$. Consider $\boldsymbol{\lambda}(d) > \tilde{\lambda}$. For all $d \geq \bar{d}$ and $\lambda \geq \tilde{\lambda}$,

$$\eta(\lambda d) \geq \eta(\bar{d}\tilde{\lambda}) \geq 1 \quad \Rightarrow \quad \frac{\partial^2 E(\lambda; d)}{\partial \lambda \partial d} \leq 0.$$

Since $\boldsymbol{\lambda}(d) = \max[\arg\max E(\cdot; d) - c(\cdot)]$. By the Topkis' monotone comparative statics theorem, $\boldsymbol{\lambda}(d) \leq \boldsymbol{\lambda}(\bar{d})$ for all $d \geq \bar{d}$. Thus, $\bar{d}$ is the optimal cutoff.

Case 2: $\boldsymbol{\lambda}(\bar{d})\bar{d} < \eta^{-1}(1)$**.** For all $\bar{d} \leq d \leq d^*$ and $\lambda \leq \frac{\eta^{-1}(1)}{d}$,

$$\eta(\lambda d) \leq 1 \quad \Rightarrow \quad \frac{\partial^2 E(\lambda; d)}{\partial \lambda \partial d} \geq 0.$$

This implies $\boldsymbol{\lambda}(d)$ is increasing in d for $\bar{d} \leq d \leq d^*$. Now suppose $d > d^*$. If $\boldsymbol{\lambda}(d)d \leq \eta^{-1}(1)$, then

$$\boldsymbol{\lambda}(d) \leq \eta^{-1}(1)/d \leq \eta^{-1}(1)/d^* = \boldsymbol{\lambda}(d^*).$$

If $\boldsymbol{\lambda}(d)d > \eta^{-1}(1)$, for all $d \geq d^*$ and $\lambda \geq \eta^{-1}(1)/d$,

$$\frac{\partial^2 E(\lambda; d)}{\partial \lambda \partial d} \leq 0 \quad \Rightarrow \quad \boldsymbol{\lambda}(d) \leq \boldsymbol{\lambda}(d^*)$$

by the Topkis' monotone comparative statics theorem.

Now, I prove existence. In case 1, $\bar{d}$ is the optimal cutoff transfer. In case 2, I can increase d starting from $\bar{d}$. Before hitting d^*, $\boldsymbol{\lambda}(d)$ increases in d. As $\eta^{-1}(1)$ is well-defined and finite given that ϕ satisfies increasing elasticity above 1. Eventually, increasing $\boldsymbol{\lambda}(d)d$ can hit $\eta^{-1}(1)$. Thus, d^* is well defined.

Step 3: Uniqueness

Finally, suppose that ϕ satisfies strictly increasing elasticity above 1, the cost function is continuously differentiable, and the optimal precision is an interior solution (not 0 or $+\infty$).[27] I show that the optimal transfer rule is unique. First, I show that for any transfer rule t that is not a cutoff transfer, the cutoff transfer constructed above induces a strictly larger precision: $\boldsymbol{\lambda}(d) > \boldsymbol{\lambda}(t)$. Since t is not a cutoff transfer, at least one of $\bar{t}' - \bar{t}$ and Δt is a non-zero function. If $\bar{t}' - \bar{t}$ is non-zero,

$$\left.\frac{\partial[\hat{E}(\lambda; \bar{t}') - \hat{E}(\lambda; \bar{t})]}{\partial \lambda}\right|_{\lambda = \boldsymbol{\lambda}(t)} > 0. \tag{15}$$

This follows by a strict version of Lemma 6 in which strictly increasing elasticity above 1 implies

$$\eta(\lambda d) < 1 \quad \Rightarrow \quad \frac{\partial^2 E(\lambda; d)}{\partial \lambda \partial d} > 0.$$

As $\hat{E}(\cdot; \bar{t}) - c(\cdot)$ is continuously differentiable,[28] $\hat{\boldsymbol{\lambda}}(\bar{t})$ is interior, Equation (15) implies $\hat{\boldsymbol{\lambda}}(\bar{t}') > \hat{\boldsymbol{\lambda}}(\bar{t})$ by Edlin and Shannon (1998). This implies $\boldsymbol{\lambda}(d) > \boldsymbol{\lambda}(t)$.

[27] Distribution ϕ satisfies *strictly increasing elasticity above 1* if $\eta(\cdot)$ single-crosses 1 from below, $\{x | \eta(x) = 1\}$ is a singleton, and is strictly increasing after the cross.

[28] $\hat{E}(\cdot; \bar{t})$ is continuously differentiable as ϕ is continuously differentiable.

If Δt is non-zero, next I show

$$\left.\frac{\partial[E(\lambda;d)-\hat{E}(\lambda;\bar{t}')]}{\partial\lambda}\right|_{\lambda=\boldsymbol{\lambda}(t)}>0. \tag{16}$$

By the construction of Δt,

$$\int \Delta t(x)\varphi(x;0,\boldsymbol{\lambda}(t))dx=0 \tag{17}$$

$$\int \Delta t(x)\phi(\boldsymbol{\lambda}(t)x)dx=0.$$

Take some small $\epsilon>0$. Let $\lambda'=\boldsymbol{\lambda}(t)\exp(\epsilon)$. Consider

$$E(\lambda';d)-\hat{E}(\lambda';\bar{t}')=\int \Delta t(x)\lambda'\phi(\lambda'x)dx.$$

By $\ln(\lambda')-\ln(\boldsymbol{\lambda}(t))=\epsilon$ and Lemma 1,

$$\ln\phi(\lambda'x)-\ln\phi(\boldsymbol{\lambda}(t)x)=-\eta(\boldsymbol{\lambda}(t)|x|))\epsilon+\mathcal{o}(\epsilon).$$

Thus,

$$\begin{aligned}E(\lambda';d)-\hat{E}(\lambda';\bar{t}')&=\int \Delta t(x)\lambda'\phi(\lambda'x)dx\\&=\lambda'\int \Delta t(x)\phi(\boldsymbol{\lambda}(t)x)\exp(-\epsilon\eta(\boldsymbol{\lambda}(t)|x|))dx+\mathcal{o}(\epsilon)\\&=-\lambda'\int \Delta t(x)\phi(\boldsymbol{\lambda}(t)x)\epsilon\eta(\boldsymbol{\lambda}(t)|x|)dx+\mathcal{o}(\epsilon),\end{aligned}$$

where the last equality follows by Equation (17).

$$\begin{aligned}\lim_{\epsilon\to0}\frac{E(\lambda';d)-\hat{E}(\lambda';\bar{t}')}{\epsilon}&=-\boldsymbol{\lambda}(t)\int \Delta t(x)\phi(\boldsymbol{\lambda}(t)x)\eta(\boldsymbol{\lambda}(t)|x|)dx\\&=-\int \Delta t(x)\varphi(x;0,\boldsymbol{\lambda}(t))\eta(\boldsymbol{\lambda}(t)|x|)dx\end{aligned}$$

$$\left.\frac{\partial[E(\lambda;d)-\hat{E}(\lambda;\bar{t}')]}{\partial\lambda}\right|_{\lambda=\boldsymbol{\lambda}(t)}=-\frac{1}{\boldsymbol{\lambda}(t)}\int \Delta t(x)\varphi(x;0,\boldsymbol{\lambda}(t))\eta(\boldsymbol{\lambda}(t)|x|)dx$$

which is strictly positive by Equation (17), $\eta(y)$ being strictly increasing for $y>$

$\eta^{-1}(1)$, and $\Delta t(x)$ being supported on $|x| \geq \frac{\eta^{-1}(1)}{\boldsymbol{\lambda}(t)}$. Again by $\hat{E}(\cdot;\bar{t}') - c(\cdot)$ being continuously differentiable, $\hat{\boldsymbol{\lambda}}(\bar{t}')$ is interior, Equation (16) implies $\boldsymbol{\lambda}(d) > \hat{\boldsymbol{\lambda}}(\bar{t}') \geq \boldsymbol{\lambda}(t)$ by Edlin and Shannon (1998).

I have shown that for any transfer rule t that is not a cutoff transfer, $\boldsymbol{\lambda}(d) > \boldsymbol{\lambda}(t)$. Thus, any optimal transfer must be a cutoff transfer. Given strictly increasing elasticity above 1, cost function being continuously differentiable, optimal precision being interior, all comparative statics analyses in the proof of Theorem 2 are strict. Thus, there exists a unique optimal transfer rule, which is a cutoff transfer.

□

Next, I prove a lemma which shall be critical for proving the necessity of Theorem 1, statement (2) implies (3).

Lemma 8. *Suppose that for every increasing, convex, and continuously differentiable cost function, there exists an optimal transfer that is a cutoff transfer. For an increasing and convex cost function $c \in C^1$, let d^* be the optimal cutoff and λ^* be the induced precision with $\eta(\lambda^* d^*) \geq 1$. Then for all $x_1 \in [0, \lambda^* d^*)$ and $x_2 \in [\lambda^* d^*, +\infty)$,*

$$\eta(x_1) \leq \eta(x_2).$$

Proof of Lemma 8. Without loss I can assume $c(\lambda^*) = E(\lambda^*; d^*)$ since I can increase c by a constant without affecting d^* and λ^*. Similarly, without loss of generality, I assume that λ^* uniquely maximizes $E(\lambda; d^*) - c(\lambda)$.[29]

Towards a contradiction, suppose $\exists\, x_1 \in [0, \lambda^* d^*)$, $x_2 \in [\lambda^* d^*, +\infty)$ such that $\eta(x_1) > \eta(x_2)$. As $\eta(\cdot)$ is continuous, I can pick $\delta_1 > 0$, $\delta_2 > 0$ small enough such that for all $y_1 \in [\frac{x_1}{\lambda^*} - \delta_1, \frac{x_1}{\lambda^*} + \delta_1]$ and $y_2 \in [\frac{x_2}{\lambda^*} - \delta_2, \frac{x_2}{\lambda^*} + \delta_2]$,

$$\eta(y_1 \lambda^*) > \eta(y_2 \lambda^*) \tag{18}$$

and

$$\int_{\frac{x_1}{\lambda^*} - \delta_1}^{\frac{x_1}{\lambda^*} + \delta_1} \varphi(x; 0, \lambda^*) dx = \int_{\frac{x_2}{\lambda^*} - \delta_2}^{\frac{x_2}{\lambda^*} + \delta_2} \varphi(x; 0, \lambda^*) dx.$$

I define

$$\Delta t(x) = -1_{|x| \in [\frac{x_1}{\lambda^*} - \delta_1, \frac{x_1}{\lambda^*} + \delta_1]} + 1_{|x| \in [\frac{x_2}{\lambda^*} - \delta_2, \frac{x_2}{\lambda^*} + \delta_2]}$$

[29] Otherwise, I can always increase $c(\lambda)$ for $\lambda \neq \lambda^*$, without affecting d^* and λ^*.

and a new transfer $\underline{t}(x) = 1_{|x|\leq d^*} + \Delta t(x)$.

I can set δ_1 and δ_2 to be small enough such that the agent reports truthfully given the transfer $\underline{t}$, since $\varphi'(d^*; 0, \lambda^*) < 0$ is bounded from above due to $\eta(\lambda^* d^*) \geq 1$. To see this, it suffices to consider the case $\lambda^* = 1$. If the agent reports truthfully, the expected transfer is $2\Phi(d^*) - 1$. Now suppose the agent misreports by $\epsilon' > 0$. Then his expected transfer is

$$\begin{aligned}
&\Phi(d^* - \epsilon') + \Phi(d^* + \epsilon') - 1 \\
&+\Phi(x_2 - \epsilon' + \delta_2) - \Phi(x_2 - \epsilon' - \delta_2) + \Phi(x_2 + \epsilon' + \delta_2) - \Phi(x_2 + \epsilon' - \delta_2) \\
&-[\Phi(x_1 - \epsilon' + \delta_1) - \Phi(x_1 - \epsilon' - \delta_1)] - [\Phi(x_1 + \epsilon' + \delta_1) - \Phi(x_1 + \epsilon' - \delta_1)],
\end{aligned}$$

where the second and the third line can be made arbitrarily close to 0 as δ_1 and δ_2 tend to 0, since Φ is continuous. Moreover, the difference

$$\Phi(d^* - \epsilon') + \Phi(d^* + \epsilon') - 2\Phi(d^*)$$

is strictly negative and is strictly decreasing in ϵ', due to ϕ being single-peaked and $\eta(d^*) = -\frac{\phi'(d^*)}{\phi(d^*)} d^* \geq 1$. Therefore, the expected transfer under misreport is lower than $2\Phi(d^*) - 1$ when δ_1 and δ_2 are small enough.

Then we have

$$E(\lambda^*; \underline{t}) = E(\lambda^*; d^*).$$

By Lemma 1 and Equation (18),

$$\left.\frac{\partial[E(\lambda; \underline{t}) - E(\lambda; d^*)]}{\partial \lambda}\right|_{\lambda=\lambda^*} > 0.$$

Since previously λ^* uniquely maximizes $E(\lambda; d^*) - c(\lambda)$, we can pick δ_1 and δ_2 small enough such that $\boldsymbol{\lambda}(\underline{t}) \in (\lambda^* - \epsilon, \lambda^* + \epsilon)$, for some small $\epsilon > 0$. As c is continuously differentiable, $E(\lambda^*; \underline{t}) = E(\lambda^*; d^*)$,

$$\left.\frac{\partial[E(\lambda; \underline{t}) - E(\lambda; d^*)]}{\partial \lambda}\right|_{\lambda=\lambda^*} > 0,$$

by Edlin and Shannon (1998) we have $\boldsymbol{\lambda}(\underline{t}) > \lambda^*$, contradicting that (d^*, λ^*) is optimal (Figure 10).

□

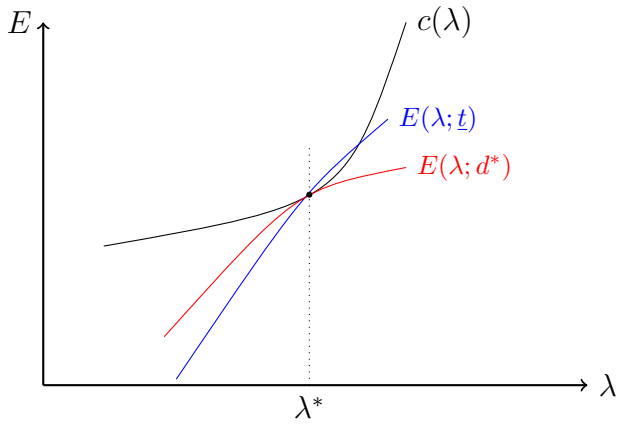

Figure 10: Expected Transfer

Proof of Theorem 1 (2) implies (3). Suppose that for every increasing, convex, and continuously differentiable cost function, there exists an optimal transfer that is a cutoff transfer. Take a increasing, convex cost function $c_0 \in C^1$, with d_0 and λ_0 being the corresponding optimal cutoff and induced precision. I can pick c_0 such that λ_0 is the unique maximizer of[30] $E(\cdot; d_0) - c_0(\cdot)$. As $E(\cdot; d_0) - c(\cdot)$ is continuously differentiable, I can use the first-order approach. As

$$\frac{\partial^2 E(\lambda; d)}{\partial \lambda \partial d} \leq 0 \quad \Leftrightarrow \quad \eta(\lambda d) \geq 1,$$

we have $\eta(\lambda_0 d_0) \geq 1$. If not, I can slightly increase d_0, which leads to a larger

$$\left.\frac{\partial E(\lambda; d)}{\partial \lambda}\right|_{\lambda=\lambda_0}.$$

This induces a larger $\lambda > \lambda_0$, a contradiction.[31]

Since the cutoff transfer d_0 is the optimal transfer and $c_0 \in C^1$, by Lemma 8 and $\eta(\cdot)$ being continuous,

$$\eta(x) \leq \eta(\lambda_0 d_0) \quad \text{if } x \leq \lambda_0 d_0$$

$$\eta(x) \geq \eta(\lambda_0 d_0) \quad \text{if } x \geq \lambda_0 d_0.$$

[30] Given any increasing and convex cost function $c_0 \in C^1$, I can always construct a new increasing, convex, and continuously differentiable cost function by increasing $c_0(\lambda)$ for all $\lambda \neq \lambda_0$ while keeping $c_0(\lambda_0)$ and $\frac{dc_0}{d\lambda}|_{\lambda=\lambda_0}$ unchanged.

[31] As λ_0 is the unique maximizer of $E(\cdot; d_0) - c_0(\cdot)$, I can always make the increment on d_0 small enough such that the new maximizer never lies in $[0, \lambda_0)$.

Note that $\eta(0^+) < 1$ (as ϕ is integrable around 0), $\eta(\lambda_0 d_0) \geq 1$, $\eta(\cdot)$ is continuous, I can find the largest x where $\eta(\cdot)$ crosses 1 from below

$$x_1 = \min\{x | \eta(x') \geq 1, \text{ if } x' \geq x\} \leq \lambda_0 d_0.$$

$\eta(x) \geq 1$ when $x \geq x_1$.

Now pick d_1 and λ_1 such that $\lambda_1 d_1 = x_1$. Construct an increasing, convex cost function $c_1 \in C^1$ such that

$$c_1(\lambda) > E(\lambda; d_1) \quad \text{if } \lambda \neq \lambda_1$$

$$c_1(\lambda_1) = E(\lambda_1; d_1)$$
$$\left.\frac{\partial E(\lambda; d_1)}{\partial \lambda}\right|_{\lambda=\lambda_1} = \left.\frac{dc_1(\lambda)}{d\lambda}\right|_{\lambda=\lambda_1}.$$

As

$$\frac{\partial^2 E(\lambda; d)}{\partial \lambda \partial d} \leq 0 \quad \Leftrightarrow \quad \eta(\lambda d) \geq 1$$

and $\eta(x) \geq 1$ for all $x \geq x_1$, the cutoff transfer d_1 is the best among all cutoff transfers for cost function c_1. Since there exists an optimal transfer that is a cutoff transfer, cutoff transfer d_1 is the optimal transfer, and λ_1 is the maximum precision. By Lemma 8,

$$\eta(x) \leq 1 \quad \text{if } x \leq x_1$$

$$\eta(x) \geq 1 \quad \text{if } x \geq x_1.$$

Thus, $\eta(\cdot)$ single-crosses 1 at x_1.

Similarly, for all $x_2 > x_1$, I can pick d_2 and λ_2 such that $\lambda_2 d_2 = x_2$. Pick a increasing, convex cost function $c_2 \in C^1$ such that

$$c_2(\lambda) > E(\lambda; d_2) \quad \text{if } \lambda \neq \lambda_2$$

$$c_2(\lambda_2) = E(\lambda_2; d_2)$$
$$\left.\frac{\partial E(\lambda; d_2)}{\partial \lambda}\right|_{\lambda=\lambda_2} = \left.\frac{dc_2(\lambda)}{d\lambda}\right|_{\lambda=\lambda_2}.$$

By the same argument, we have

$$\eta(x) \leq \eta(x_2) \quad \text{if } x \leq x_2$$

$$\eta(x) \geq \eta(x_2) \quad \text{if } x \geq x_2.$$

Thus, $\eta(\cdot)$ is increasing once it goes above 1 at x_1.

□

Proof of Proposition 4. First suppose for all λ, $c_2(\lambda) \geq E(\lambda; d^*(c_1))$. Then $\bar{d}(c_2) \geq d^*(c_1)$, implying $d^*(c_2) \geq d^*(c_1)$. Moreover, since

$$\text{for all } \lambda \geq \lambda^*(c_1) \text{ and } d \geq d^*(c_1), \quad \frac{\partial^2 E(\lambda; d)}{\partial \lambda \partial d} \leq 0$$

and $c_2 - c_1$ is increasing,

$$\max[\arg\max_{\lambda} E(\lambda; d^*(c_2)) - c_2(\lambda)] \leq \max[\arg\max_{\lambda} E(\lambda; d^*(c_1)) - c_1(\lambda)].$$

Thus, $\lambda^*(c_2) \leq \lambda^*(c_1)$. Now suppose for some λ, $c_2(\lambda) \leq E(\lambda; d^*(c_1))$. Note that

$$\begin{aligned}
&\max[\arg\max_{\lambda} E(\lambda; d^*(c_1)) - c_2(\lambda)] \\
=&\max[\arg\max_{\lambda} E(\lambda; d^*(c_1)) - c_1(\lambda) - (c_2(\lambda) - c_1(\lambda))] \\
\leq&\max[\arg\max_{\lambda} E(\lambda; d^*(c_1)) - c_1(\lambda)] = \lambda^*(c_1),
\end{aligned}$$

where the inequality follows by $c_2 - c_1$ being increasing. Thus,

$$\boldsymbol{\lambda}(d^*(c_1); c_2) d^*(c_1) \leq \lambda^*(c_1) d^*(c_1) = \eta^{-1}(1).$$

By the proof of Theorem 2, $d^*(c_2) \geq d^*(c_1)$ and $\lambda^*(c_2) \leq \lambda^*(c_1)$.

□

Proof of Corollary 1. Changing the noise from ε_1 to $k\varepsilon_1$ is equivalent to keeping the noise at ε_1 and changing the cost function from $c_1(\lambda)$ to $c_2(\lambda) = c_1(k\lambda)$. Notice that

$$c_2(\lambda) - c_1(\lambda) = c_1(k\lambda) - c_1(\lambda)$$

which is increasing in λ due to the convexity of c_1. The conclusion follows by Propo-

sition 4. □

Proof of Proposition 1. The proof of the first part remains the same as Theorem 1. For proving the second part, the only difference is to note that the expected transfer is

$$E(\lambda; d) = \int_0^{\lambda d} \phi(r) \frac{\pi^{n/2}}{\Gamma(n/2+1)} dr^n,$$

where Γ is the gamma function and $\frac{\pi^{n/2}}{\Gamma(n/2+1)} r^n$ is the volume of the n-dimensional ball.

$$\frac{\partial^2 E(\lambda; d)}{\partial \lambda \partial d} = \frac{n\pi^{n/2}}{\Gamma(n/2+1)} \phi(\lambda d)(\lambda d)^{n-1}[n - \eta(\lambda d)].$$

Thus,

$$\frac{\partial^2 E(\lambda; d)}{\partial \lambda \partial d} \geq 0 \quad \Leftrightarrow \quad \eta(\lambda d) \leq n.$$

The rest of the argument remains the same.

□

Proof of Proposition 2. The optimality of cutoff transfers follows by the Proof of Theorem 1 (3) implies (1), as Gaussian distributions satisfy increasing elasticity.

For proving the counterpart of Theorem 2, it suffices to notice that

$$\frac{\partial^2 E(\Lambda(\lambda); d)}{\partial \lambda \partial d} = 2\phi(\Lambda(\lambda) d) \frac{d\Lambda(\lambda)}{d\lambda} [1 - \eta(\Lambda(\lambda) d)].$$

□

Proof of Proposition 3. I first prove the case of uniform prior. Given the agent's signal s, his posterior belief about the state is $\mathcal{N}(s, 1/\lambda^2)$. Since the principal's signal is Gaussian centered at θ with variance $\frac{1}{\lambda_p^2}$, the agent's posterior belief about s_p is

$$\mathcal{N}(s, \frac{1}{\lambda^2} + \frac{1}{\lambda_p^2}).$$

Now the agent's payoff maximization problem is same as in Section 4 except that s_p replaces θ and his precision is $(\frac{1}{\lambda^2} + \frac{1}{\lambda_p^2})^{-\frac{1}{2}}$. Thus, Proposition 3 follows from Theorem 1 and 2.

The proof for the Gaussian prior is similar. The only difference is that conditional

on a signal s, the agent's posterior about s_p is

$$\mathcal{N}\left(\frac{s\lambda^2}{\lambda_0^2+\lambda^2}, \frac{1}{\lambda_0^2+\lambda^2}+\frac{1}{\lambda_p^2}\right).$$

Thus, the precision is $\Lambda = \left[\frac{1}{\lambda_0^2+\lambda^2}+\frac{1}{\lambda_p^2}\right]^{-\frac{1}{2}}$. □

D Omitted Proofs of Section 6

I first establish the regularity needed to compare arbitrary transfer rules. Recall that $\hat{E}(\lambda;t)$ denotes the expected transfer when the agent reports truthfully. The maximum over reports defining $E(\lambda;t)$ is attained by Lemma 5.

Lemma 9. *The map $\lambda \mapsto \varphi(\cdot;0,\lambda)$ is continuously differentiable as an $L^1(\mathbb{R})$-valued map on $(0,+\infty)$, with*

$$\frac{\partial \varphi(x;0,\lambda)}{\partial \lambda} = \phi(\lambda x) + \lambda x \phi'(\lambda x).$$

Consequently, for every bounded measurable transfer t,

$$\frac{\partial \hat{E}(\lambda;t)}{\partial \lambda} = \int_{\mathbb{R}} t(x) \frac{\partial \varphi(x;0,\lambda)}{\partial \lambda}\, dx.$$

For every admissible transfer rule, $E(\lambda;t)$ tends to zero as λ tends to zero.

Proof. Single-peakedness and integrability imply that $x\phi(x)$ tends to zero as x tends to $+\infty$, since

$$x\phi(x) \leq 2\int_{x/2}^{x} \phi(y)\, dy \longrightarrow 0.$$

Integration by parts, symmetry, and $\phi'(x) \leq 0$ for $x > 0$ give

$$\int_{\mathbb{R}} |x\phi'(x)|\, dx = 2\int_0^{+\infty} -x\phi'(x)\, dx = 1.$$

Thus $g(x) = \phi(x) + x\phi'(x)$ belongs to $L^1(\mathbb{R})$. The dilation map $\lambda \mapsto g(\lambda\cdot)$ is continuous in L^1 for $\lambda > 0$. This follows first for continuous functions of compact support and then by approximation in L^1.

Pointwise differentiation gives $\partial_\lambda \varphi(x;0,\lambda) = g(\lambda x)$. For sufficiently small nonzero h, the fundamental theorem of calculus implies

$$\left\| \frac{\varphi(\cdot;0,\lambda+h) - \varphi(\cdot;0,\lambda)}{h} - g(\lambda\cdot) \right\|_1 \leq \frac{1}{|h|} \int_{\min\{\lambda,\lambda+h\}}^{\max\{\lambda,\lambda+h\}} \|g(u\cdot) - g(\lambda\cdot)\|_1 \, du.$$

The right-hand side tends to zero. This proves L^1 differentiability and continuity of the derivative. Integrating against a bounded t proves the formula for $\hat{E}$.

Finally, fix $\epsilon > 0$ and choose $A > 0$ such that $t(x) \leq \epsilon$ whenever $|x| > A$. Since $\varphi(x;0,\lambda) \leq \lambda\phi(0)$, for every report displacement b,

$$\int_{\mathbb{R}} t(x-b)\varphi(x;0,\lambda)\, dx \leq 2A\lambda\phi(0) + \epsilon.$$

The bound is uniform in b, so $E(\lambda;t) \to 0$ as $\lambda \downarrow 0$. □

Proof of Theorem 3, (3) ⇒ (1)*, and Theorem 4.* Suppose that ϕ satisfies increasing elasticity below 1. Fix a cost function satisfying Assumption 1. If $\bar{\lambda}c'(\bar{\lambda}) = 0$, convexity and $c(0) = 0$ imply $c(\bar{\lambda}) = 0$. The zero transfer therefore implements $\bar{\lambda}$ and satisfies participation. Since transfers are nonnegative, it is optimal and $C(\bar{\lambda}) = 0$. Moreover, $H(0) = G(0) = 0$, so the formulas in Theorem 4 hold with $z_{\bar{\lambda}} = d^*(\bar{\lambda}) = 0$. Henceforth, suppose that $\bar{\lambda} > 0$ and $c'(\bar{\lambda}) > 0$.

Step 1: a bound on marginal incentives. Extend η to zero by $\eta(0) = 0$, and write

$$z_0 = \inf\{z > 0 : \eta(z) \geq 1\}.$$

The function H vanishes at zero and tends to zero at the positive support endpoint, or at infinity when the support is unbounded. It is strictly positive in the interior of the positive half-support. It therefore attains a positive maximum at an interior point. Moreover,

$$H'(z) = 2\phi(z)[1 - \eta(z)]$$

where $\phi(z) > 0$. Increasing elasticity below 1 implies that z_0 is finite, lies in the positive support, and satisfies

$$\eta(z_0) = 1, \qquad \eta(z) < 1 \quad (0 \leq z < z_0), \qquad \eta(z) \geq 1 \quad (z > z_0).$$

Hence H is strictly increasing on $[0, z_0]$, weakly decreasing thereafter, and

$$H(z_0) = \max_{z \geq 0} H(z).$$

The first point where elasticity reaches 1 is used here, so this argument also covers an interval on which $\eta = 1$.

On the positive support of $\varphi(\cdot; 0, \bar{\lambda})$, define

$$q_{\bar{\lambda}}(x) = \frac{1 - \eta(\bar{\lambda}|x|)}{\bar{\lambda}},$$

and set $q_{\bar{\lambda}}(x) = 0$ outside that support. Then

$$\left.\frac{\partial \varphi(x; 0, \lambda)}{\partial \lambda}\right|_{\lambda = \bar{\lambda}} = q_{\bar{\lambda}}(x)\varphi(x; 0, \bar{\lambda})$$

almost everywhere. The convention outside the support avoids multiplying an infinite elasticity by a zero density.

Consider any transfer t feasible in the relaxed problem (5). If t depends separately on the state and the report, use the auxiliary transfer $\tilde{t}$ constructed in Lemma 4, with $\bar{\lambda}$ in place of $\boldsymbol{\lambda}(t)$ as explained after that lemma. If t depends only on the prediction error, choose an optimal report displacement b^* at the target and let $\tilde{t}(x) = t(x - b^*)$. In either case,

$$\hat{E}(\bar{\lambda}; \tilde{t}) = E(\bar{\lambda}; t), \qquad \hat{E}(\mu; \tilde{t}) \leq E(\mu; t) \quad \text{for all } \mu \geq 0.$$

The IC constraint therefore implies that $\bar{\lambda}$ maximizes $\hat{E}(\mu; \tilde{t}) - c(\mu)$. Lemma 9 and the interiority of $\bar{\lambda}$ give the necessary condition

$$\int_{\mathbb{R}} \tilde{t}(x) q_{\bar{\lambda}}(x) \varphi(x; 0, \bar{\lambda})\, dx = c'(\bar{\lambda}). \tag{19}$$

Thus the bound below applies without assuming that t induces truthful reporting. It also does not require differentiating the report-optimized value E.

Because $0 \le \tilde{t} \le 1$, condition (19) implies

$$\begin{aligned} c'(\bar{\lambda}) &\le \int_{\mathbb{R}} (q_{\bar{\lambda}}(x))_+ \varphi(x; 0, \bar{\lambda})\, dx \\ &= \frac{2}{\bar{\lambda}} \int_0^{z_0} \phi(z)[1 - \eta(z)]\, dz = \frac{H(z_0)}{\bar{\lambda}}. \end{aligned} \tag{20}$$

The last equality follows from $(z\phi(z))' = \phi(z)[1 - \eta(z)]$. This proves the necessity of (9).

Step 2: construction and global incentive compatibility. Suppose that (9) holds. Strict increase of H on $[0, z_0]$ gives a unique solution $z_{\bar{\lambda}} \in (0, z_0]$ to (8); this is also the smallest nonnegative solution. Set $d^* = z_{\bar{\lambda}}/\bar{\lambda}$ and let t^* be the cutoff transfer with cutoff d^*.

By symmetry and single-peakedness, a centered interval has weakly greater probability than any other interval of the same length. Thus the cutoff induces truthful reporting at every precision, and

$$E(\mu; t^*) = G(\mu d^*) \quad \text{for all } \mu \ge 0.$$

Moreover,

$$\frac{d}{d\mu} G(\mu d^*) = 2d^* \phi(\mu d^*), \qquad \frac{d^2}{d\mu^2} G(\mu d^*) = 2(d^*)^2 \phi'(\mu d^*) \le 0.$$

Consequently, $G(\mu d^*) - c(\mu)$ is concave. The equation defining $z_{\bar{\lambda}}$ gives

$$2d^* \phi(\bar{\lambda} d^*) = c'(\bar{\lambda}),$$

so $\bar{\lambda}$ is a global maximizer. Since truthful reporting is optimal at every precision, this verifies IC including changes in reporting after a precision deviation. Comparing with $\mu = 0$ also verifies IR, because $G(0) - c(0) = 0$.

Step 3: minimum payment below the boundary. First suppose that $\bar{\lambda} c'(\bar{\lambda}) < H(z_0)$. Then $z_{\bar{\lambda}} < z_0$, and

$$\rho = \frac{1 - \eta(z_{\bar{\lambda}})}{\bar{\lambda}} > 0.$$

Increasing elasticity below 1 implies, on the positive density support,

$$q_{\bar{\lambda}}(x) \geq \rho \quad \text{if } |x| \leq d^*, \qquad q_{\bar{\lambda}}(x) \leq \rho \quad \text{if } |x| > d^*.$$

For any feasible t, take the auxiliary transfer $\tilde{t}$ constructed in Step 1. The payment bounds imply the pointwise inequality

$$[q_{\bar{\lambda}}(x) - \rho][\tilde{t}(x) - t^*(x)] \leq 0$$

almost everywhere under $\varphi(\cdot; 0, \bar{\lambda})$. Both $\tilde{t}$ and t^* satisfy the same marginal condition (19). Integrating therefore gives

$$-\rho \int_{\mathbb{R}} [\tilde{t}(x) - t^*(x)]\varphi(x; 0, \bar{\lambda})\, dx \leq 0.$$

The integral equals $E(\bar{\lambda}; t) - E(\bar{\lambda}; t^*)$. Since $\rho > 0$,

$$E(\bar{\lambda}; t) \geq E(\bar{\lambda}; t^*) = G(z_{\bar{\lambda}}).$$

Thus t^* solves the relaxed problem and induces truthful reporting.

Step 4: minimum payment at the boundary. Now suppose that $\bar{\lambda}c'(\bar{\lambda}) = H(z_0)$, so that $z_{\bar{\lambda}} = z_0$. Equality must hold in (20). For the auxiliary transfer $\tilde{t}$ constructed in Step 1 from any feasible contract, this requires

$$\int_{\mathbb{R}} (1 - \tilde{t})(q_{\bar{\lambda}})_+ \varphi(x; 0, \bar{\lambda})\, dx + \int_{\mathbb{R}} \tilde{t}(q_{\bar{\lambda}})_- \varphi(x; 0, \bar{\lambda})\, dx = 0,$$

where $q_- = \max\{-q, 0\}$. Both terms are nonnegative. Consequently, $\tilde{t} = 1$ almost everywhere on positive-score events and $\tilde{t} = 0$ almost everywhere on negative-score events. The positive-score region is $|x| < z_0/\bar{\lambda}$, so the expected payment is at least $G(z_0)$. The proposed cutoff attains this bound by also paying zero on zero-score events, apart from irrelevant boundary points. This establishes optimality even if $\eta = 1$ on an interval.

Together, these steps prove (3) $\Rightarrow$ (1) in Theorem 3, and all the claims in Theorem 4. □

Proof of Theorem 3, (1) $\Rightarrow$ (2) $\Rightarrow$ (3). The implication (1) $\Rightarrow$ (2) is immediate. I

prove (2) $\Rightarrow$ (3) by contraposition.

Step 1: a cutoff on the initial increasing branch. Suppose that increasing elasticity below 1 fails. There exist $0 < x_1 < y$ in the positive support of ϕ such that

$$\eta(y) < 1, \qquad \eta(y) < \min\{\eta(x_1), 1\}.$$

Choose v such that $\eta(y) < v < \min\{\eta(x_1), 1\}$, and let r be the first point where $\eta(r) = v$. Continuity of elasticity on the positive support and $\eta(0) = 0$ imply

$$0 < r < x_1 < y, \qquad 0 < v < 1, \qquad \eta(x) < v \quad (0 \le x < r).$$

In particular,

$$\phi'(r) = -\frac{v}{r}\phi(r) < 0, \tag{21}$$

and H is strictly increasing on $[0, r]$.

Take the target $\bar{\lambda} = 1$ and the linear cost

$$c(\mu) = k\mu, \qquad k = H(r) = 2r\phi(r) > 0.$$

Let t_0 be the cutoff transfer with cutoff r.

Step 2: the baseline has a unique joint optimum. For a precision $\mu > 0$ and report displacement $b = a - s$, let

$$V_0(\mu, b) = \int_{-r}^{r} \mu\phi(\mu(x + b))\, dx - k\mu.$$

Set $V_0(0, b) = 0$. Symmetry and single-peakedness imply $V_0(\mu, b) \le V_0(\mu, 0)$, while

$$\frac{\partial V_0(\mu, 0)}{\partial \mu} = 2r\phi(\mu r) - k.$$

Weak monotonicity of ϕ, together with its strict local decrease at r, makes this derivative strictly positive for $\mu < 1$ and strictly negative for $\mu > 1$. Thus precision one is the unique optimum. At $\mu = 1$, the centered interval probability is nonincreasing in $|b|$, and its second derivative at $b = 0$ is $2\phi'(r) < 0$. It is therefore strictly smaller for

every $b \neq 0$. Hence $(1,0)$ is the unique global maximizer, with strictly positive value

$$V_0(1,0) = G(r) - H(r) = 2\int_0^r [\phi(x) - \phi(r)]\,dx > 0. \tag{22}$$

Step 3: reduce expected payment while preserving the marginal incentive. Choose symmetric sets A and B, each a finite union of bounded intervals, with A just inside $\pm r$ and B around $\pm y$. By continuity, these sets can be chosen disjoint and inside the positive density support so that every value of $1 - \eta(|x|)$ on B is strictly greater than every such value on A, and all these values are positive. Write

$$\begin{aligned} p_A &= \int_A \phi(x)\,dx, \quad m_A = \int_A [1-\eta(|x|)]\phi(x)\,dx, \\ p_B &= \int_B \phi(x)\,dx, \quad m_B = \int_B [1-\eta(|x|)]\phi(x)\,dx. \end{aligned}$$

These four quantities are positive, and $m_A/p_A < m_B/p_B$. Let $\kappa = m_A/m_B$ and define

$$t_\epsilon = t_0 - \epsilon \mathbf{1}_A + \epsilon\kappa \mathbf{1}_B. \tag{23}$$

For $0 < \epsilon < \min\{1, 1/\kappa\}$, the transfer takes values in $[0,1]$. It is symmetric and compactly supported. Lemma 9 gives

$$\left.\frac{\partial \hat{E}(\mu; t_\epsilon)}{\partial \mu}\right|_{\mu=1} = k - \epsilon m_A + \epsilon\kappa m_B = k. \tag{24}$$

At the same time, the truthful expected payment strictly decreases:

$$\begin{aligned} \hat{E}(1; t_\epsilon) &= G(r) - \epsilon p_A + \epsilon\kappa p_B \\ &= G(r) - \epsilon p_A \left(1 - \frac{m_A/p_A}{m_B/p_B}\right) < G(r). \end{aligned} \tag{25}$$

Step 4: preserve the exact local joint optimum. Define the joint net payoff

$$V_\epsilon(\mu, b) = \int_{\mathbb{R}} t_\epsilon(x)\mu\phi(\mu(x+b))\,dx - k\mu$$

for $\mu > 0$, and set $V_\epsilon(0, b) = 0$. These functions are C^2 near $(1, 0)$. Indeed, for any bounded interval $[\ell, u]$,

$$\int_\ell^u \mu\phi(\mu(x + b))\, dx = \Phi(\mu(u + b)) - \Phi(\mu(\ell + b)),$$

which is C^2 because ϕ is C^1. Every transfer in (23) is a finite linear combination of interval indicators. At the baseline optimum,

$$\nabla^2 V_0(1, 0) = \begin{pmatrix} 2r^2\phi'(r) & 0 \\ 0 & 2\phi'(r) \end{pmatrix}$$

is negative definite. Thus there is a convex neighborhood N of $(1, 0)$, with compact closure in $(0, +\infty) \times \mathbb{R}$, on which the baseline Hessian is uniformly negative definite. Since $V_\epsilon = V_0 + \epsilon W$ for a fixed C^2 function W, the same is true for every sufficiently small ϵ. Equation (24) keeps the precision derivative zero at $(1, 0)$, and symmetry keeps the reporting derivative zero there. Therefore $(1, 0)$ remains the unique maximizer of V_ϵ within N.

Step 5: rule out distant joint deviations. All sufficiently small perturbations have support in one common interval $[-D, D]$ and take values in $[0, 1]$. For all report displacements b,

$$V_\epsilon(\mu, b) \le 2D\phi(0)\mu \quad \text{as } \mu \downarrow 0, \qquad V_\epsilon(\mu, b) \le 1 - k\mu \quad \text{as } \mu \to +\infty.$$

Moreover, for $\mu \in [\delta, L]$, where $0 < \delta < L < +\infty$, and for $|b| > D$, symmetry and monotonicity of ϕ give

$$V_\epsilon(\mu, b) \le 2DL\phi(\delta(|b| - D)) \longrightarrow 0 \quad \text{as } |b| \to +\infty.$$

These bounds are uniform over sufficiently small ϵ. The positive baseline payoff in (22) therefore implies that all choices capable of matching the payoff at $(1, 0)$ lie in one common compact rectangle K. Choose K to contain $\overline{N}$ in its interior. Uniqueness of the baseline maximum and continuity give a strictly positive gap

$$\Delta = V_0(1, 0) - \max_{(\mu,b)\in K\setminus N} V_0(\mu, b) > 0.$$

Uniformly over the entire choice set,

$$|V_\epsilon(\mu, b) - V_0(\mu, b)| \leq \|t_\epsilon - t_0\|_\infty \leq \epsilon \max\{1, \kappa\}.$$

Choose ϵ small enough to preserve this gap and keep $V_\epsilon(1, 0) > 0$. Step 4 rules out deviations inside N; the gap rules out those in $K \setminus N$; and the uniform bounds rule out those outside K. Thus $(1, 0)$ is the unique global maximizer of V_ϵ.

The perturbed transfer is therefore feasible, induces truthful reporting at the target, and satisfies strict IR. This argument allows the agent to change his report after any precision deviation. In particular, $E(1; t_\epsilon) = \hat{E}(1; t_\epsilon) < G(r)$.

Step 6: every feasible cutoff is more expensive. Any cutoff transfer with cutoff d that implements precision one under cost $k\mu$ must satisfy

$$2d\phi(d) = k = H(r).$$

Since H is strictly increasing on $[0, r]$, no such cutoff has $d < r$. Its expected payment is therefore at least $G(r)$. By (25), the globally feasible transfer t_ϵ costs strictly less. Hence no cutoff transfer is optimal for this linear cost and this implementable target. This contradicts statement (2), proving (2) $\Rightarrow$ (3). $\square$