Data and incentives

Annie Liang
Department of Economics, Northwestern University

Erik Madsen
Department of Economics, New York University

"Big data" gives markets access to previously unmeasured characteristics of individual agents. Policymakers must decide whether and how to regulate the use of this data. We study how new data affects incentives for agents to exert effort in settings such as the labor market, where an agent's quality is initially unknown but is forecast from an observable outcome. We show that measurement of a new covariate has a systematic effect on the average effort exerted by agents, with the direction of the effect determined by whether the covariate is informative about long-run quality versus a shock to short-run outcomes. For a class of covariates satisfying a statistical property that we call *strong homoskedasticity*, this effect is uniform across agents. More generally, new measurements can impact agents unequally, and we show that these distributional effects have a first-order impact on social welfare.

Keywords. Big data, forecasting, effort incentives, career concerns.

JEL classification. C72, D83, L51.

1. Introduction

Online platforms and data brokers extensively track, record, and aggregate consumer activities, producing measurements of everything from the size of an individual's social network,[1] to how often they move residences,[2] to the amount of time they spend playing video games.[3] These new measurements are increasingly available to firms and organizations, who may find them useful as predictors of economic outcomes, such as

[1]The finance startup Lenddo evaluated borrowers on the basis of factors such as "how many friends or followers they have on their social networks" (Gage, 2012).

[2]The alternative credit scoring company ZestFinance used borrowers' frequency of residence changes to predict their creditworthiness (Lippert, 2014).

[3]China's widely-publicized social credit scoring system reportedly plans to incorporate data on how many video games a consumer purchases and how much time they spend playing them (Canales and Mok, 2022).

a worker's future productivity in a new job.[4] Regulating such uses of personal data has emerged as an important policy issue,[5] but our understanding of when and how to do so remains preliminary. The use of new data in such settings may have far-ranging social impacts beyond direct privacy concerns, reshaping the creation and distribution of economic surplus.

In this paper, we study the impact of new data on markets in which moral hazard is an important concern.[6] A motivating application is the labor market, where wages and job opportunities are commonly tied to market forecasts of a worker's productivity based on past output. A well-recognized consequence of this practice is that workers are incentivized to work hard to improve the market's forecast. Since output is typically socially valuable, new data which impacts workers' effort incentives may negatively affect labor market productivity absent regulation.

We propose a simple model of reputational incentives that isolates the effect of new data on moral hazard. Our model builds on the classic "career concerns" framework of Holmström (1999), in which an agent exerts effort to improve an outcome used by a market to forecast his type. Different from Holmström (1999), we suppose that the market additionally bases its forecast on auxiliary data consisting of covariates describing the agent, which are observed prior to his choice of effort.

We separate covariates into two categories: Some covariates, which we call *attributes*, describe the agent's type; while others, which we call *circumstances*, are informative about a transient shock to his outcome. For example, a worker's creativity is an attribute, while an illness or injury is a circumstance. We model the acquisition of new data as an expansion of the set of covariates that are measured and may be used for forecasting. Measurement of new covariates updates the market's beliefs about a given agent's type and shock, reshaping incentives for effort.

Our main results characterize how measurement of a new covariate impacts the population distribution of effort and aggregate welfare. Our basic positive result is that incorporating a new covariate into the market's type forecast leads to both a systematic *reduction in uncertainty* across the population and a *redistribution of uncertainty* between agents. While the systematic effect moves the effort of all agents in the same direction, the redistributionary effect leads to heterogeneous effort responses, which may differ even directionally across agents. Each of these effects has a first-order effect on aggregate welfare, and we find that the redistributionary effect can oppose and even overturn the welfare impact of reducing uncertainty.

---

[4]Employers already widely use similar information collected from internet searches to screen potential hires on the basis of factors such as social media activity (CareerBuilder, 2018).

[5]For instance, proposed European Union rules for artificial intelligence have flagged automated employment screening systems as "high risk" applications subject to strict regulation, in particular regarding the data sets they rely on (European Commission, 2021).

[6]Our approach complements recent work focusing on how data collection impacts markets shaped by asymmetric information. See, for instance, Bergemann, Bonatti, and Gan (2022), Elliott, Galeotti, and Koh (2022), Yang (2022) for price discrimination; Ichihashi (2019), Hidir and Vellodi (2021), Gomes and Pavan (2022) for matching on platforms; and Braverman and Chassang (2022), Brunnermeier, Lamba, and Segura-Rodriguez (2021) for insurance pricing.

We formalize our positive findings through a pair of theorems. Theorem 1 shows that quite generally, measurement of a new attribute reduces average effort in the population, while measurement of a new circumstance increases average effort. Importantly, this result does not guarantee that all agents change their effort by the same amount, or even in the same direction. Theorem 2 shows that any heterogeneity in the effort responses of different agents is entirely attributable to a redistribution of uncertainty across the population.

Our normative findings are summarized in Theorem 3. It establishes that, in the absence of redistribution of uncertainty, the directional effect on welfare of a newly measured covariate is jointly determined by its classification as an attribute or circumstance along with the weight that agents place on their future reputations. It further shows that greater redistribution of uncertainty leads to reduced welfare gains from measurement of a covariate, and that this reduction can be so extreme that the measurement of some covariates is *never* welfare-enhancing, regardless of the magnitude of agents' reputational concerns.

Our work contributes to an emerging literature studying the use of personal data for forecasting. Existing work has highlighted incentives for agents to game forecasts by distorting (Ball (2022), Bonatti and Cisternas (2020), Frankel and Kartik (2022), Haghtalab, Immorlica, Lucier, and Wang (2020), Hu, Immorlica, and Vaughan (2019)) or misreporting (Eliaz and Spiegler (2019, 2022)) their covariates. Such incentives are especially important when a small number of covariates shape forecasts in a well-understood way. We study the complementary question of how data usage impacts incentives for agents to directly improve outcomes. These incentives are particularly relevant when outcomes are a primary forecasting input, making them a natural target for manipulation; or when algorithms used to incorporate additional covariates into forecasts are opaque, obscuring effective strategies for gaming them.

Additionally, our model builds on the career concerns literature. Compared to the original Holmström (1999) model, we focus on a two-period model which incorporates auxiliary signals and non-Gaussian information structures. We share these modeling features with the closely related work of Dewatripont, Jewitt, and Tirole (1999) and Rodina (2018).[7] These papers show that under certain conditions, signals about a single agent's type lower his effort while signals about the shock raise it.[8] In our model with a population of heterogeneous agents, a similar result holds on average across agents

---

[7]An adjacent literature on relative performance comparisons, e.g., Meyer and Vickers (1997), considers settings in which the additional signal is not exogenous, but is instead generated by the outcome of another agent with correlated unobservables. See also Tirole (2021), in which the additional signal is an outcome in another domain in which the agent exerts effort.

[8]Rodina (2018) additionally studies whether garbling the outcome signal can improve agent incentives, a question also examined by Hörner and Lambert (2021) and Smolin (2021) in related contexts. By contrast, we assume that the agent's outcome is perfectly observable and focus on regulation of access to additional covariates.

(Theorem 1). But additionally, introduction of new data potentially changes the cross-sectional dispersion of effort in the population. As we show in Section 3.3, this distributional effect has a first-order effect on welfare, and in some cases it can override the welfare effect of an average change in effort. Our results therefore highlight the importance of explicitly modeling the heterogeneity across agents that is present in applications.

The remainder of this paper proceeds as follows. Section 2 describes our model; Section 3 establishes our main results about the impact of new measurements on effort and welfare; Section 4 discusses extensions; and Section 5 concludes. Supporting analyses and all proofs are collected in the Appendix.

2. Model

In Section 2.1, we describe our basic model of reputational incentives for effort, which is a 2-period version of the Holmström (1999) career concerns model with general information structures. In Section 2.2, we augment the model by introducing auxiliary data.

2.1 *Effort and welfare*

An agent participates in a market across two periods $t = 1, 2$. He possesses a quality type $\theta \sim F_\theta$, which is persistent across time and unknown to himself and the market.

In **period 1**, the agent privately chooses an effort level $e \in \mathbb{R}_+$ at cost $C(e) = \frac{1}{2}e^2$. (We extend our results to general cost functions in Section 4.3.) The agent's effort choice, along with his quality θ and a transient shock $\varepsilon \sim F_\varepsilon$, determine the realization of an observable outcome

$$Y = e + \theta + \varepsilon.$$

We assume that $\mathbb{E}(\theta) = \mu > 0$ while $\mathbb{E}(\varepsilon) = 0$. The agent's reward from his period-1 interaction is independent of Y and normalized to be 0.[9] His period-1 payoff is therefore

$$U_1 = -\frac{1}{2}e^2.$$

In **period 2**, the agent receives a reputational payoff standing in for returns from future participation in the market. This payoff is equal to the market's expectation of his quality conditional on the outcome variable Y.[10] Since the agent's effort choice is private, the market's forecast is based on a conjectured level of effort $\hat{e}$. Letting $Y^{\hat{e}} \equiv \hat{e} + \theta + \varepsilon$ be the outcome supposing that the market's effort conjecture is correct, the agent's second-period payoff conditional on the realized outcome $Y = y$ is

$$U_2 = \mathbb{E}^{\hat{e}}(\theta \mid Y = y),$$

---

[9]This normalization does not rule out wage payments which depend on the market's forecast of period-1 effort, as in Holmström (1999). Such effects do not impact equilibrium effort or social surplus, and so we do not explicitly model them.

[10]None of our results would change if the agent's reputational payoff were instead the market's expectation of any strictly increasing function of θ. Our model therefore accommodates a variety of interpretations for the source of reputational returns from effort.

where $\mathbb{E}^{\hat{e}}(\theta \mid Y = y)$ denotes the market's (potentially misspecified) expectation of θ, updated based on the realized outcome assuming that $Y = Y^{\hat{e}}$.

The agent's ex post payoff from participating in the market in both periods is a weighted sum of payoffs across the two periods:

$$U = (1 - \beta) \cdot U_1 + \beta \cdot U_2,$$

where $\beta \in (0, 1)$ is the *reputation weight*, which denotes the importance to the agent of future reputational rewards versus current effort costs. The agent's expected payoff under effort level e is therefore

$$\mathbb{E}^e(U) = \beta \cdot \mathbb{E}^e\big(\mathbb{E}^{\hat{e}}(\theta \mid Y)\big) - (1 - \beta) \cdot \frac{e^2}{2},$$

where $\mathbb{E}^e$ denotes the expectation operator given the true effort level e.

In equilibrium, the agent must have no incentive to deviate from the market's conjectured level of effort. Let e^* denote equilibrium effort. Then in equilibrium the marginal value of effort (i.e., the equilibrium marginal impact of effort on the expected reputational reward), discounted by its relative weight $\beta/(1 - \beta)$, equals the marginal cost of effort:

$$\frac{\beta}{1 - \beta} \cdot \frac{\partial}{\partial e}\mathbb{E}^e\big(\mathbb{E}^{e^*}(\theta \mid Y)\big)\bigg|_{e=e^*} = e^*.$$

Because effort impacts the outcome additively, the marginal value of effort appearing in this first-order condition is independent of e^* and may be written as

$$MV \equiv \mathbb{E}^0\left(\frac{\partial}{\partial Y}\mathbb{E}^0(\theta \mid Y)\right), \tag{1}$$

where $\mathbb{E}^0$ denotes the expectation operator assuming that the agent does not exert effort to distort the outcome (see Appendix D.1 for details). As (1) does not depend on e^*, the unique effort level satisfying the first-order condition is then $e^* = \frac{\beta}{1-\beta} \cdot MV$. Throughout this paper, we will assume that the first-order approach is valid, so that e^* constitutes the unique equilibrium effort choice.

We measure welfare using a standard criterion that treats both the outcome variable Y and the agent's effort cost $C(e)$ as welfare-relevant.[11] An agent whose type is θ and effort choice is e thus generates welfare

$$w(\theta, e) \equiv \mathbb{E}^e(Y \mid \theta) - C(e) = \theta + e - \frac{1}{2}e^2. \tag{2}$$

This function is strictly concave in effort and maximized at the "first-best" effort level $e^{FB} = 1$ no matter the agent's type. Appendix A extends our analysis to alternative welfare specifications incorporating learning-by-doing and unproductive gaming.

---

[11]The assumption that Y contributes to social welfare is appropriate for settings such as the labor market, in which the outcome captures productive output or some other socially valuable activity. We do not include the agent's equilibrium reputational payoff in the welfare calculation because on average that payoff is fixed at $\mathbb{E}^{e^*}(\mathbb{E}^{e^*}(\theta \mid Y)) = \mu$, independent of the equilibrium effort level. (This property continues to hold when the model is augmented with data.).

2.2 *Data and beliefs*

We now augment the basic model by supposing that θ and ε are predictable from underlying (and potentially measurable) *covariates*, with data revealing a subset of these covariates. We refer to those covariates which predict θ as *attributes*, denoted by the random variables $(a_1, \ldots, a_J)$, and those covariates which predict ε as *circumstances*, denoted by the random variables $(c_1, \ldots, c_K)$. Specifically, the type θ and shock ε satisfy

$$\theta = f^1(a_1) + \cdots + f^J(a_J) + u_\theta$$

$$\varepsilon = g^1(c_1) + \cdots + g^K(c_K) + u_\varepsilon$$

where each f^j, $j \in \{1, \ldots, J\}$, and g^k, $k \in \{1, \ldots, K\}$, is a deterministic and one-to-one effect size function. (This specification nests the standard linear regression model as a special case when all effect size functions are affine.) For convenience, we will define the *type components* $\theta_j \equiv f^j(a_j)$ for each $j = 1, \ldots, J$ and *shock components* $\varepsilon_k \equiv g^k(c_k)$ for each $k = 1, \ldots, K$.[12]

The idiosyncratic noise terms u_θ and u_ε are independent of one another and of all covariates, and have full support on the reals.[13] We allow for correlation between attributes and between circumstances, but assume that the vector of attributes is independent of the vector of circumstances, i.e., $(a_1, \ldots, a_J) \perp\!\!\!\perp (c_1, \ldots, c_K)$, implying in particular that $\theta \perp\!\!\!\perp \varepsilon$. (We consider covariates which are correlated with both the type and shock in Section 4.1.)

Some covariates are measured, making them observable to the agent and the market. We use $\mathcal{J} \subseteq \{1, \ldots, J\}$ to denote the set of measured attributes and $\mathcal{K} \subseteq \{1, \ldots, K\}$ to denote the set of measured circumstances. All measured covariates are observed at the outset of the interaction, leading the agent and market to share a common belief that the agent's type and shock follow their distributions conditional on the agent's measured covariate values. We view symmetric uncertainty as a natural conceptual benchmark that allows us to cleanly disentangle moral hazard from issues of selection. In Section 2.3, we discuss how our results would change if either side had additional private information.

The interaction then proceeds as described in Section 2.1, with appropriate adjustments to the calculation of equilibrium effort. Conditioning on the measured covariates, the agent's marginal value of effort changes from (1) to the quantity

$$MV_{\mathcal{J},\mathcal{K}} \equiv \mathbb{E}^0\left(\frac{\partial}{\partial Y} \mathbb{E}^0(\theta \mid Y, a_{\mathcal{J}}, c_{\mathcal{K}}) \mid a_{\mathcal{J}}, c_{\mathcal{K}} \right) \tag{3}$$

and equilibrium effort becomes

$$e^*_{\mathcal{J},\mathcal{K}} = \frac{\beta}{1 - \beta} \cdot MV_{\mathcal{J},\mathcal{K}}. \tag{4}$$

---

[12]Invertibility of the effect size functions implies that observation of a new covariate a_j or c_k is equivalent to observation of the corresponding type or shock component θ_j or ε_k. Some of our results, particularly those involving strong homoskedasticity, do not depend on invertibility.

[13]The full support assumption ensures that the distributions of θ and ε conditional on any family of measured covariates have full support, simplifying our proofs. All of our results continue to hold in the absence of full support, and so we will make free use of examples which do not feature full-support idiosyncratic noise terms.

Note that both the marginal value of effort and equilibrium effort may vary with the values of the agent's measured covariates, and they are therefore both random quantities. We interpret this randomness from a population perspective, by supposing that the market interacts with a continuum of agents possessing varying attributes and circumstances. From this perspective, random variation in $e^*_{\mathcal{J},\mathcal{K}}$ corresponds to a distribution of effort across the population of agents.

Aggregate welfare given measured covariates $(\mathcal{J},\mathcal{K})$ is the expectation of realized welfare $w(\theta, e^*_{\mathcal{J},\mathcal{K}})$ as defined in (2), averaging over variation in the type θ and effort $e^*_{\mathcal{J},\mathcal{K}}$ across the population:

$$W(\mathcal{J},\mathcal{K}) \equiv \mathbb{E}\big(w\big(\theta, e^*_{\mathcal{J},\mathcal{K}}\big)\big) = \mu + \mathbb{E}\left(e^*_{\mathcal{J},\mathcal{K}} - \frac{1}{2}\big(e^*_{\mathcal{J},\mathcal{K}}\big)^2\right).$$

(Recall that $\mu \equiv \mathbb{E}(\theta)$ is the unconditional average quality in the population.) Aggregate welfare is maximized when all agents exert the first-best level $e^{FB} = 1$.

Our main results compare effort and welfare when the set of measured covariates changes from some *baseline* family $(\mathcal{J},\mathcal{K})$ to an *expanded* family $(\mathcal{J} \cup \{j'\}, \mathcal{K})$ or $(\mathcal{J}, \mathcal{K} \cup \{k'\})$ containing one additional covariate. To simplify exposition, throughout the main text we develop our results assuming that the baseline family is $(\mathcal{J},\mathcal{K}) = (\emptyset, \emptyset)$, while the expanded family is either $(\{1\}, \emptyset)$ or $(\emptyset, \{1\})$, corresponding to measurement of attribute 1 or circumstance 1. (Our results extend straightforwardly to general baselines—see Appendix B for details.) In this context, we define $\eta \equiv \sum_{j>1} f^j(a_j) + u_\theta$ and $\delta \equiv \sum_{k>1} g^k(c_k) + u_\varepsilon$ and decompose the type and shock as

$$\theta = \theta_1 + \eta, \qquad \varepsilon = \varepsilon_1 + \delta,$$

so that the type component θ_1 and shock component ε_1 summarize the information revealed by a new measurement, while η and δ are the residual unknowns.

We impose a set of standard regularity conditions on the distributions of these variables.

ASSUMPTION 1 (Admissibility). *The random variables θ and ε have log-concave density functions, and the conditional random variables $\eta \mid a_1$ and $\delta \mid c_1$ have log-concave density functions for every realization of a_1 and c_1.*

ASSUMPTION 2 (Differentiability). *For every effort level e, the derivative $\frac{\partial}{\partial Y}\mathbb{E}^e(\theta \mid Y)$ exists and is uniformly bounded across all realizations of Y. Additionally, for every effort level e and realization of (a_1, c_1), the derivatives $\frac{\partial}{\partial Y}\mathbb{E}^e(\theta \mid Y, a_1)$ and $\frac{\partial}{\partial Y}\mathbb{E}^e(\theta \mid Y, c_1)$ exist and are uniformly bounded across all realizations of Y.*

In additive statistical inference models, log-concavity is a canonical assumption ensuring that better outcomes correspond to improved inferences about latent variables.[14]

---

[14]Specifically, if an analyst observes an outcome Z which is decomposable as $Z = X + Y$, where both X and Y are unobserved, and if X and Y are statistically independent and have log-concave density functions, then upon observing Z his posterior beliefs about X and Y are higher (in the first-order stochastic dominance order) for larger realizations of Z (Milgrom (1981))..

Assumption 1 ensures that in both the baseline and the expanded environments, better (worse) realizations of Y lead to higher (lower) posterior beliefs about both the type and shock. Assumption 2 ensures that conditional expectations are sufficiently smooth that we can take derivatives and exchange derivatives and expectations where required.

2.3 *Discussion of modeling choices*

Private information on the side of the market   Our model assumes that the market does not know more about an agent's type or shock than the agent does himself. In some applications, private information on the side of the market is possible if the market has access to data on past outcomes for other agents with similar covariates. (This view of big data is embedded in, for instance, the "inverse selection" model of Brunnermeier, Lamba, and Segura-Rodriguez (2021).) We show in Section 4.2 that our results continue to hold under this sort of informational asymmetry, so long as measurement of a new covariate leads the agent to believe that the market has gained new information about his type or shock.

Private information on the side of the agent   Another possibility is that the agent knows more about his type and shock than the market. This asymmetry implies not only that the agent knows more about his own covariates, but also that he can discern how these covariates impact his type and shock distributions (via the effect size functions f^j and g^k), a demanding assumption in many applications. Nevertheless, our results continue to hold with private information on the agent's side whenever the market's posterior type expectation is linear in the outcome signal (for instance, whenever $(\theta_1, \ldots, \theta_J)$ and $(\varepsilon_1, \ldots, \varepsilon_K)$ follow elliptical distributions). Beyond these settings, the agent's perceived marginal value of effort may vary with private information about his type or shock, complicating the analysis, but we conjecture that our main results extend more broadly.

Exogeneity of covariates   Our model contrasts covariates, which are fixed characteristics of the agent (at least in the short run); and the outcome Y, which is susceptible to manipulation by effort. We view this dichotomy as a useful one for several reasons. First, θ and ε may be determined by the aggregation of many covariates, each of which individually plays only a small role. In such contexts, our exercise can be viewed as focusing on the agent's incentives to influence the relatively informative outcome signal Y, while abstracting from any costly distortion of less-informative individual covariates. Second, to compute the value of manipulating a covariate, the agent must know the precise shape of the effect size function describing how that covariate impacts the outcome, which is more demanding than the knowledge requirements that we impose.

3. Main results

Our main results characterize the impact of measuring a new covariate on population effort and welfare. We first show that for a broad class of covariates, measuring a new attribute decreases population effort on average, while measuring a new circumstance increases it (Section 3.1). However, outside a narrower class of covariates, new measurements may yield effort responses of heterogeneous magnitude and even direction

across agents (Section 3.2). In Section 3.3, we combine these insights to study the effect of measuring a new covariate on aggregate welfare, taking into account both the resulting average effect on effort as well as the induced heterogeneity.

3.1 *The average impact of new measurements*

Our first main result demonstrates that in a wide class of models, measuring a new covariate leads to a systematic shift in average effort across the population. Further, the direction of this shift depends solely on whether the covariate is an attribute or a circumstance: Attributes reduce average effort, while circumstances increase it.

Our result applies to covariates that satisfy the following condition.

DEFINITION 1 (Affiliation). Attribute 1 is *affiliated* if (θ_1, η) are statistically affiliated.[15] Circumstance 1 is *affiliated* if (ε_1, δ) are statistically affiliated.

Affiliation describes settings in which a good realization of a measured covariate (e.g., θ_1) also implies good realizations (on average) of unmeasured covariates (e.g., η). Examples of settings in which attribute 1 is affiliated include:

EXAMPLE 1. $(\theta_1, \ldots, \theta_J)$ follow a multivariate normal distribution, and all correlation coefficients are nonnegative.   ◊

EXAMPLE 2. $(\theta_1, \ldots, \theta_J)$ are i.i.d. draws from an exponential distribution with rate parameter λ, where $\lambda \sim \text{Gamma}(\alpha_0, \beta_0)$ with $\alpha_0 \geq 1$.   ◊

The following result establishes that a newly measured affiliated attribute lowers average effort, while a newly measured affiliated circumstance raises it.[16]

THEOREM 1. *Suppose Assumptions 1–2 hold. If attribute* 1 *is affiliated, then measuring it weakly reduces average effort. If circumstance* 1 *is affiliated, then measuring it weakly increases average effort.*

To understand the result, consider the simplest setting in which all covariates are independent of one another. (Independence is a sufficient condition for attribute 1 and circumstance 1 to be affiliated.) In this case, measuring a new attribute reduces the market's ex post uncertainty about the agent's type. As a result, the market infers less about the agent's type from the outcome, decreasing the marginal value of improving the outcome through costly effort and therefore also the equilibrium level of effort. Measuring a new circumstance instead reduces the market's ex post uncertainty about the agent's

---

[15]A pair of random variables (Z_1, Z_2) is *statistically affiliated* if $Z_2 \mid Z_1$ satisfies the monotone likelihood ratio property with respect to Z_1. When (Z_1, Z_2) possess a strictly positive, twice-differentiable joint density function $\rho(z_1, z_2)$, this condition is equivalent to $\partial^2 \log \rho / \partial z_1 \partial z_2 \geq 0$ everywhere.

[16] We formally establish weak monotonicity. At the cost of a more involved proof, it can further be shown that when a covariate is nondegenerate, affiliation implies strict monotonicity.

shock, leading it to infer more about the agent's type from the outcome and boosting the agent's equilibrium effort.

When covariates are correlated, measuring a new covariate need not reduce the market's ex post uncertainty about all agents in the population. (See Section 3.3 for an example.) To sign the average change in effort, we must show that even if uncertainty about some agents increases, their effort does not change by so much as to outweigh the opposing effort changes by all other agents. Affiliation ensures enough comovement of covariates to guarantee this result.

Theorem 1 is closely related to a result of Dewatripont, Jewitt, and Tirole (1999) (hereafter DJT) in a related setting where a single agent exerts effort prior to the realization of an auxiliary signal. (Recall that in contrast, covariates in our model are observed before each agent chooses their effort, so the market holds heterogeneous beliefs about agents at the time of effort choice.) DJT show that, under certain statistical conditions, the agent's effort drops in anticipation of a signal about his type and rises in anticipation of a signal about the shock.[17]

Both results employ the same basic logic, but they rely on different statistical conditions. Our condition is formulated for our additive setting and highlights the role played by correlation between covariates, while DJT derive a more abstract condition in a setting with less structure on signals. Additionally, DJT assume technical regularity conditions that we are able to relax through an alternative proof technique.[18] Finally, in our setting the change in average effort induced by a new measurement is only one aspect of its impact on the population distribution of effort and on aggregate welfare. Our next results explore these implications in more depth.

3.2 *Redistribution of uncertainty*

We now show that *redistribution of uncertainty* across agents is necessary for measurement of a new covariate to impact higher moments of the population effort distribution. Formally, whenever a covariate reduces uncertainty about the corresponding outcome component in the same way for all agents, the average effect identified in Theorem 1 is the sole effect of a new measurement.

The condition we identify requires that the variance as well as all higher moments of the unmeasured component are independent of the covariate realization. It strengthens the notion of homoskedasticity commonly imposed in linear regression models, which requires independence only of the residual variance. By analogy, we refer to it as *strong homoskedasticity*.

DEFINITION 2 (Strong homoskedasticity). Attribute 1 satisfies *strong homoskedasticity* if the random variable $\eta - \mathbb{E}(\eta \mid a_1)$ is independent of a_1. Circumstance 1 satisfies *strong homoskedasticity* if the random variable $\delta - \mathbb{E}(\delta \mid c_1)$ is independent of c_1.

---

[17]See, for example, Example 5.3 in DJT for an instance of this result in an additive model.

[18]Specifically, DJT implicitly assume sufficient regularity of the distributions of all random variables to permit two successive exchanges of an effort derivative and an expectation (see their Proposition 2.2). Our proof avoids one of these exchanges, and as a result we require only the weaker regularity condition imposed in Assumption 2.

Strong homoskedasticity does not eliminate correlation between covariates, since the expectations of η and δ are allowed to depend (respectively) on a_1 and c_1. It does, however, rule out the possibility that the spread of an unmeasured covariate may depend on the measured covariate. To give a concrete example, suppose the measured covariate "zip code" is correlated with the unmeasured covariate "income." Strong homoskedasticity allows the average income level to differ across zip codes, but is violated if income levels are more heterogeneous in some zip codes than in others.

Additional examples of strongly homoskedastic models include:

EXAMPLE 3. $(\theta_1, \ldots, \theta_J)$ and $(\varepsilon_1, \ldots, \varepsilon_K)$ each follow multivariate normal distributions.
$$\Diamond$$

EXAMPLE 4. The type and shock components satisfy $\theta_j = F_j(\theta_1) + X_j$ for $j \geq 2$ and $\varepsilon_k = G_k(\varepsilon_1) + Z_k$ for $k \geq 2$, for functions F_j and G_k and random variables $X_j \perp\!\!\!\perp \theta_1$ and $Z_k \perp\!\!\!\perp \varepsilon_1$.
$$\Diamond$$

Under strong homoskedasticity, a strong version of Theorem 1 ensures a uniform change of effort across agents.[19] Since strong homoskedasticity encompasses correlation structures outside the class of affiliated covariates (for instance, multivariate normal models with negative correlation coefficients), this result further broadens the finding of Theorem 1.

THEOREM 2. *Suppose Assumptions 1–2 hold. If attribute 1 satisfies strong homoskedasticity, then measuring it weakly reduces every agent's effort. If circumstance 1 satisfies strong homoskedasticity, then measuring it weakly increases every agent's effort.*

In either case, the magnitude of the effort change is the same for every agent, irrespective of their covariate realization.

When strong homoskedasticity fails, measurement of a new covariate induces heterogeneity in the spread of the market's beliefs about different agents, an effect we refer to as redistribution of uncertainty. This redistribution can lead to heterogeneous effort changes across the population, and may even exert a directional effect on effort which for some agents overturns the average effect identified in Theorem 1. We provide an illustration of this possibility through an example in Section 3.3.

Our results therefore highlight an important connection between redistribution of uncertainty and inequality generated by data usage. When a new measurement redistributes uncertainty across agents, its payoff impact may differ across agents with different covariate realizations, leading some agents to become better off at the expense of others.[20] A social planner may care intrinsically about this disparate impact, especially

---

[19] As in Theorem 1, we formally establish only weak monotonicity. When a covariate is additionally nondegenerate, strong homoskedasticity further implies strict monotonicity. (See footnote 16.).

[20]If agents do not receive any up-front payments in period 1, they become better or worse off to the extent their effort falls or rises. In some contexts, agents might be compensated for the value generated by their equilibrium effort. (See, for instance, the competitive labor market of Holmström (1999).) In that case, agents' payoffs increase in effort over some range. Nonetheless, redistribution of uncertainty still benefits some agents at the expense of others.

since these payoff changes may be correlated with sensitive social or demographic char-
acteristics via the measured covariates.[21] In the next section, we show that even when
the social planner does not have any intrinsic distributional concerns, this dispersion
harms aggregate welfare.

3.3 *Social welfare and data regulation*

An important question faced by regulators is whether to allow firms in particular mar-
kets to use new covariates for forecasting. We now apply our positive results to answer
this normative question. We characterize the welfare impact of measuring *regular co-
variates*, a class encompassing all covariates exhibiting the systematic impact on average
effort identified in Theorems 1 and 2.[22]

DEFINITION 3. Attribute 1 is *regular* if measuring it weakly reduces the marginal value
of effort on average. Circumstance 1 is *regular* if measuring it weakly increases the
marginal value of effort on average. If the change is strict, we call the covariate *strictly*
regular.

It follows from Theorems 1 and 2 that all affiliated and strongly homoskedastic co-
variates are regular.[23]

To develop some intuition for the impact of a new measurement on welfare, sup-
pose first that effort in both the baseline and expanded environments is deterministic,
e.g., because the newly measured covariate is strongly homoskedastic. Let $MV \equiv MV_{\emptyset,\emptyset}$
and $MV_{+A} \equiv MV_{\{1\},\emptyset}$ denote the marginal value of effort before and after attribute 1 is
measured. We will similarly let e and e_{+A} denote effort in these two environments. The
agent's effort is related to his marginal value of effort via

$$e = \frac{\beta}{1-\beta} \cdot MV, \qquad e_{+A} = \frac{\beta}{1-\beta} \cdot MV_{+A}.$$

If attribute 1 is regular, then $MV_{+A} \leq MV$ and so correspondingly $e_{+A} \leq e$. Whether
this change is welfare-improving depends on the size of β. If β is sufficiently large, then
measuring the attribute moves the agent's effort closer to the first-best level $e^{FB} = 1$,
increasing welfare; in contrast, if β is sufficiently small, then effort moves away from
first-best and welfare is reduced. This logic is reversed for circumstances. See Figure 1
for an illustration.

When agents react heterogeneously to a newly measured covariate, aggregate wel-
fare depends on details of the distribution of effort beyond the mean. Signing the effect

---

[21]Automated prediction algorithms have recently come under scrutiny for unintentionally discriminat-
ing against protected social groups such as racial minorities on the basis of such correlations (Kearns and
Roth (2019)). Our analysis highlights a new channel through which such correlations might harm disad-
vantaged groups.

[22]To facilitate nonquadratic effort costs (see Section 4.3), we define regularity with respect to the
marginal value of effort rather than effort itself. Since equilibrium effort is proportional to the marginal
value of effort when costs are quadratic, the two notions are equivalent in our baseline setting.

[23]As noted in footnotes 16 and 19, when a covariate is additionally nondegenerate, it can be shown to be
strictly regular.

baseline effort

$$W(e)$$

0

Effort e

e^{FB}

new attribute   new circumstance

(a) Weak reputational incentives (low β)

baseline effort

$$W(e)$$

0

Effort e

e^{FB}

new attribute   new circumstance

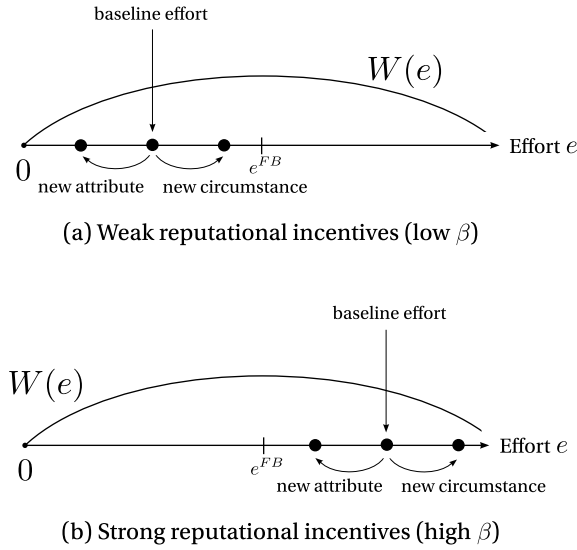(b) Strong reputational incentives (high β)

FIGURE 1. The welfare impact of measuring a new covariate when effort is deterministic.

of a covariate on welfare therefore requires aggregating the systematic effort effect identified above and the additional impact of redistributing uncertainty.

The following theorem shows that even in the presence of disparate impact, the effect of a newly measured covariate on aggregate welfare depends in a simple way on the size of the reputation weight. Specifically, for every regular attribute, there is a threshold reputation weight β^* such that measuring the covariate is welfare-improving only for reputation weights above β^*.[24] Analogously, for every regular circumstance, there is a threshold reputation weight β_* such that measuring the covariate is welfare-improving only for all reputation weights below β_*.

THEOREM 3. *Suppose attribute 1 is regular. Then there exists a threshold reputation weight $\beta^* \in (0, 1]$ such that measuring the attribute is welfare-improving if and only if $\beta > \beta^*$. Moreover, $\beta^* < 1$ if and only if*

$$\mathbb{E}\big(MV_{+A}^2\big) < \mathbb{E}\big(MV^2\big). \tag{5}$$

Suppose circumstance 1 is regular. Then there exists a threshold reputation weight $\beta_* \in [0, 1)$ such that measuring the circumstance is welfare-improving if and only if $\beta < \beta_*$. Moreover, $\beta_* > 0$ if and only if the circumstance is strictly regular.

Despite the apparent symmetry between the two parts of this result, they are not perfectly mirrored. Every (strictly) regular circumstance is welfare-improving for sufficiently small β, but some attributes fail to improve welfare even for large β. This asymmetry stems from the fact that a larger reputation weight leads to increased effort dispersion in the presence of uncertainty redistribution, and hence increased aggregate effort

---

[24]We say that a covariate is welfare-improving if aggregate welfare *strictly* increases when the covariate is measured.

costs. For attributes, this effect is largest precisely when the average effort decrease from measuring the covariate is most beneficial to welfare, leading to an ambiguous relationship between reputational concerns and total welfare. For circumstances, by contrast, the dispersion effect is most pronounced when the effort increase from measurement is most harmful to welfare, and so the two forces reinforce one another.

Condition (5) characterizes when the dispersion effect overwhelms the average effect of measuring a new attribute. If the attribute decreases the marginal value of effort for all agents, then a fall in the expected marginal value of effort implies that the expectation of its square falls as well, satisfying (5). The theorem then guarantees that the attribute improves welfare for all sufficiently large reputation weights. By contrast, when the newly measured attribute increases effort for some agents and decreases it for others, the dispersion of the marginal value of effort MV_{+A} generated by redistribution of uncertainty may be sufficiently large that (5) fails. In this case, measuring the attribute diminishes welfare regardless of the weight agents place on their future reputation.

This result highlights the role that redistribution of uncertainty plays in determining welfare. Even if measuring a new covariate leads to a welfare-improving shift in aggregate effort, it may be optimal to prohibit the use of this covariate for forecasting if the aggregate shift is achieved at the cost of large welfare losses associated with particular groups of agents. We conclude our analysis with a simple example demonstrating this possibility.

Example Worker productivity θ is a function of two attributes: residential stability $a_1 \in [0, 1]$ and worker reliability $a_2 \in [0, 1]$. These attributes determine the worker's productivity via $\theta = f^1(a_1) + f^2(a_2)$, where

$$f^1(a_1) = a_1/10, \qquad f^2(a_2) = 10 \cdot a_2$$

There are no measurable circumstances, and the total shock to output ε follows the standard normal distribution $\mathcal{N}(0, 1)$.

Residential stability and worker reliability are positively correlated, with residential stability a_1 distributed uniformly on $[0, 1]$, while the conditional distribution of worker reliability a_2 given a_1 is

$$a_2 \mid a_1 \sim \begin{cases} U([0.9, 1]), & \forall a_1 \geq 0.05 \\ U([0, 1]), & \forall a_1 < 0.05 \end{cases}$$

That is, workers with very low residential stability (bottom 5%) are less reliable on average, but are also substantially more heterogeneous.[25]

In a labor market in which neither attribute is measured, the market learns about worker productivity solely through output. All workers then exert a common level of effort, which can be numerically computed to be $e^* \approx 0.16 \cdot \frac{\beta}{1-\beta}$. If the market begins measuring residential stability, equilibrium effort for workers with high residential stability

---

[25]On the one hand, some workers may move frequently due to evictions, reflecting low reliability. On the other hand, the exceptionally prolific mathematician Paul Erdős famously possessed no permanent residence (Cofield, 2013).

$(a_1 > 0.05)$ falls by roughly 50%, to $e_H^{**} \approx 0.077 \cdot \frac{\beta}{1-\beta}$, while effort for all remaining workers rises fivefold to $e_L^{**} \approx 0.82 \cdot \frac{\beta}{1-\beta}$. Intuitively, although residential stability has only a small direct impact on job performance, it has a large impact on the market's uncertainty about worker reliability. The market's uncertainty about workers who move very frequently (residential stability falls below the bottom 5% percentile) *increases*, while the market's uncertainty about the type of all remaining workers decreases. Since, moreover,

$$\mathbb{E}\big(MV_{+A}^2\big) \approx 0.039 > MV^2 \approx 0.026,$$

Theorem 3 implies that aggregate welfare decreases upon measuring attribute 1 for any reputation weight β.

4. Extensions

We now analyze several extensions of our framework. Section 4.1 studies the effect of covariates, which are correlated with both the type and shock. Section 4.2 establishes that our results are robust to agent uncertainty about the market's beliefs. Section 4.3 relaxes the assumption that the agent's effort costs are quadratic.

4.1 *General covariates*

Our main results have assumed that individual covariates are informative about the agent's type θ or shock ε, but not both. In some applications, covariates may plausibly lie somewhere between these two extremes. We now show how our results can be adapted to accommodate such covariates.

  As in our main results, we focus on a baseline in which no covariates are observed and an expanded data set consisting of a single covariate, whose value we will denote by the random variable X. We allow this covariate to be correlated with both θ and ε in a very general way, which we summarize by its effect on the conditional mean of the outcome. Let $Y^0 \equiv \theta + \varepsilon$ be the baseline outcome ignoring the agent's effort, and define the random variable $\bar{Y}^0 \equiv \mathbb{E}(\theta + \varepsilon \mid X)$ to be the conditional mean of the baseline outcome given the covariate. We maintain an invertibility assumption ensuring that measuring X is equivalent to observing the conditional mean outcome $\bar{Y}^0$.

ASSUMPTION 3 (General invertibility). $\mathbb{E}(\theta + \varepsilon \mid X = x)$ *is a one-to-one function of* x.

  This condition is analogous to the invertibility assumptions imposed on the effect size functions f^j and g^k in the baseline model, and it serves the same purpose.

  We additionally impose admissibility and differentiability assumptions analogous to Assumptions 1 and 2 in the baseline analysis.

ASSUMPTION 4 (General admissibility). $(\bar{Y}^0, Y^0)$ *are statistically affiliated.*

ASSUMPTION 5 (General differentiability). *For every effort level e, the derivative $\frac{\partial}{\partial Y}\mathbb{E}^e(\theta \mid Y)$ exists and is uniformly bounded across all realizations of Y. For every effort level e and realization of X, the derivative $\frac{\partial}{\partial Y}\mathbb{E}^e(\theta \mid Y, X)$ exists and is uniformly bounded across all realizations of Y.*

Assumption 4 serves the same role as Assumption 1 in ensuring that better outcomes correspond to improved inferences about latent variables. In particular, if the "net residual" $\Delta Y^0 \equiv Y^0 - \bar{Y}^0$ is independent of the realization of X, affiliation of $(\bar{Y}^0, Y^0)$ reduces to log-concavity of the density function of ΔY^0. Meanwhile, Assumption 5 is a straightforward adaptation of Assumption 2.

Finally, we impose an assumption ensuring that θ and ε are correlated only through the covariate X. We maintain it to focus on the simplest context in which correlation between the type and shock might arise.

ASSUMPTION 6 (General independence). (θ, ε) *are independent conditional on* X.

We now derive conditions under which measuring X increases or decreases effort, extending the results of Theorem 1 to this setting.

PROPOSITION 1. *Suppose that Assumptions 3–6 hold.*

1. *If* $(\bar{Y}^0, \theta)$ *and* $(\bar{Y}^0, -\varepsilon)$ *are each statistically affiliated, then measuring X reduces average effort.*

2. *If* $(\bar{Y}^0, -\theta)$ *and* $(\bar{Y}^0, \varepsilon)$ *are each statistically affiliated, then measuring X increases average effort.*

This result establishes that a covariate which is positively associated with one component of the outcome, and is simultaneously negatively associated with the remaining component, has an unambiguous impact on the expected marginal value of effort. The positive association condition here is a direct analog of the affiliation condition in Theorem 1, and is needed for the same reason. Meanwhile, the negative association condition rules out scenarios in which a good covariate realization implies both a high type *and* a high shock. Since these inferences have conflicting effects on the marginal value of effort, the net effect of measuring such a covariate is inherently ambiguous. By contrast, if a good covariate realization suggests a high type and a low shock, or vice versa, the two effects reinforce and the measurement has an unambiguous impact on average effort.

This result can be strengthened to obtain a uniform effect on effort under homoskedasticity conditions similar to those imposed in Theorem 2. In particular, let $\Delta\theta \equiv \theta - \mathbb{E}(\theta \mid X)$ be the residual unobserved type component after measuring X. Define $\Delta\varepsilon$ similarly with respect to the shock. Then if the joint distribution of $(\Delta\theta, \Delta\varepsilon)$ is independent of the realization of X, measuring X affects effort uniformly across all agents.

To illustrate these forces concretely, we analyze the effect of measuring a new covariate in a multivariate Gaussian setting. Suppose that θ and ε are decomposable as

$$\theta = \mu + b \cdot X + Z, \qquad \varepsilon = d \cdot X + W$$

where $X \sim N(0, \sigma_x^2)$, $Z \sim N(0, \sigma_z^2)$, and $W \sim N(0, \sigma_w^2)$ are mutually independent and μ, b, and d are known constants. The following lemma ensures that the regularity assumptions imposed in Proposition 1 are satisfied in this setting whenever $b + d \neq 0$, a condition we will maintain going forward.[26]

LEMMA 1. *Assumptions 3–6 are satisfied in a multivariate Gaussian setting with a general covariate whenever $b + d \neq 0$.*

We now check when the conditions identified in Proposition 1 under which measuring X increases or decreases effort are satisfied. $(\bar{Y}^0, \theta)$ and $(\bar{Y}^0, \varepsilon)$ are each jointly Gaussian, and $(\bar{Y}^0, \pm\theta)$ are positively correlated if and only if $\text{sign}(b + d) = \pm\text{sign}(b)$. Similarly, $(\bar{Y}^0, \pm\varepsilon)$ are positively correlated if and only if $\text{sign}(b + d) = \pm\text{sign}(d)$. Then whenever $b > 0$, Proposition 1 implies that measuring X reduces effort if $b + d > 0 \geq d$, i.e., $d \in (-b, 0]$. Similarly, whenever $d > 0$, measuring X increases effort if $b + d > 0 \geq b$, i.e., $b \in (-d, 0]$.

The bounds $d \leq 0$ and $b \leq 0$ illustrate the general point made earlier: Measuring X has an unambiguous effect on effort only if its informativeness about one component of the outcome is reinforced rather than opposed by its informativeness about the remaining component. The remaining condition $b + d > 0$ ensures that better outcomes correspond to improved inferences about the type or shock in the baseline, without which the expected directional effect of a measurement can reverse.

We can verify these results by explicitly calculating MV and MV_+, the marginal value of effort before and after measuring X. The following result summarizes the calculation.

PROPOSITION 2. $\text{sign}(MV - MV_+) = \text{sign}((b + d)(\frac{b}{\sigma_z^2} - \frac{d}{\sigma_w^2}))$.

If $b > 0$ and $d \in (-b, 0]$, this result implies that $MV_+ < MV$, in line with the prediction of Proposition 1. Similarly, if $d > 0$ and $b \in (-d, 0]$, then $MV_+ > MV$. Conversely, if both b and d are positive, the sign of $MV - MV_+$ is ambiguous. Depending on the sizes of these coefficients relative to the residual uncertainty about θ and ε, measuring X could move the marginal value of effort in either direction.

4.2 *Model uncertainty and misspecification*

Suppose that, contrary to our assumptions in the baseline model, the agent is subjectively uncertain about the market's perceived distribution of (θ, ε) given his measured covariates. Such a situation may arise if he does not know which set of covariates the market observes, or if he does not know how the market maps his covariate values into perceived type and shock distributions.

This subjective uncertainty can be modeled by supposing the agent possesses beliefs over possible joint distributions of (θ, ε) that the market might hold when forecasting the agent's type. (It is not important that the market's true model be contained in the

---

[26]If $b + d = 0$, then X does not impact Y and cannot be estimated by observing the outcome. As a result, measuring it has no impact on the marginal value of effort.

support of the agent's beliefs, so the agent may be misspecified.) We will continue to maintain the assumption that the agent is not asymmetrically informed about his type, and so his own subjective belief about the distribution of his type and outcome is the expectation of his belief about the market's distribution.

In this setting, all of our results extend in the following sense: If the agent becomes convinced that the market's statistical model has become "better-informed" about the agent's type or shock, his effort will move in the direction predicted by our results, so long as the corresponding statistical assumptions hold for each model in the support of the agent's beliefs. More precisely, an agent believes the market has become "better-informed" if he thinks that, regardless of what statistical model it is in fact using, the market has gained access to an additional attribute or additional covariate. In that case, the marginal value of effort moves in the same direction for every model in the support of the agent's beliefs, and the expected marginal value of effort therefore moves in this direction as well. Our main results therefore continue to hold in this environment.

4.3 *General convex cost functions*

We have established our main results under the assumption that effort costs take the form $C(e) = \frac{1}{2}e^2$. Under this cost function, equilibrium effort is identical to the marginal value of effort, allowing us to characterize the former by analyzing the latter. More generally, when C is a strictly convex cost function, equilibrium effort is a uniquely determined, strictly increasing function of the marginal value of effort:

$$e^*_{\mathcal{J},\mathcal{K}} = (C')^{-1}\left(\frac{\beta}{1-\beta} \cdot MV_{\mathcal{J},\mathcal{K}}\right),$$

where $MV_{\mathcal{J},\mathcal{K}}$ is as defined in (3). As a result, under such a cost function, a deterministic shift in the marginal value of effort implies a change in effort in the same direction. This implies in particular that the results of Theorem 2 under strong homoskedasticity extend immediately.

Theorem 1 for affiliated covariates extends so long as all agents change their effort in the same direction, and more generally under a condition on the third derivative of the effort cost function.[27] (In Appendix C, we present similar, but more restrictive, generalizations of the welfare results from Section 3.3.)

PROPOSITION 3. *Suppose Assumptions 1–2 hold.*

(a) *If attribute* 1 *is affiliated, then measuring it reduces average effort if $C''' \geq 0$ or all agents change their effort in the same direction.*

(b) *If circumstance* 1 *is affiliated, then measuring it increases average effort if $C''' \leq 0$ or all agents change their effort in the same direction.*

---

[27]The proof of this result is a straightforward application of the proof of Theorem 1, combined with the logic of the discussion following the proposition statement.

The new force which arises under general cost functions is that average effort may respond to mean-preserving spreads of the marginal value of effort. To illustrate this possibility, consider any cost function $C(e) \propto e^k$, where $k > 1$. If $k > 2$, then under such a cost function the marginal cost of effort is convex, so equilibrium effort is a concave function of the marginal value of effort. Hence, any mean-preserving spread of the marginal value of effort reduces average effort. Conversely, if $2 > k > 1$, effort is a convex function of the marginal value of effort, and a mean-preserving spread of the marginal value of effort increases average effort.

Measuring a new affiliated covariate has two effects: It shifts the *average* marginal value of effort, and (whenever strong homoskedasticity fails) it may additionally introduce a *spread* in the distribution of marginal values. If the marginal cost of effort is convex, this second effect tends to reduce equilibrium effort. Thus, when a new attribute is measured, these two forces work together to lower average effort, and the results of Theorem 1 continue to hold. A similar outcome holds when the marginal cost of effort is concave and a new circumstance is measured. When the two forces conflict, the net effect on effort is ambiguous. In particular, if agents change their effort in different directions, average effort could move in the opposite direction from the average marginal value of effort.

5. Conclusion

As firms and governments move toward collecting large consumer data sets as inputs to decision-making, the question of whether and how to regulate the usage of personal data has emerged as an important policy question. Recent regulations, such as the European Union's General Data Protection Regulation, have focused on protecting consumer privacy and improving transparency regarding what kind of data is being collected. An important complementary consideration is how data impacts economic outcomes. In this paper, we have focused on one such factor—the effect that market access to novel covariates has on incentives for hidden effort.

Our results indicate that forecasting from data on enduring personal attributes decreases average effort across the population, while conversely data reflecting short-lived circumstances boosts effort. It is therefore important to distinguish between these two classes of data when regulating data usage. Further, new data may lead to increased variation in effort across workers, an outcome which has a first-order impact on welfare. This finding suggests that regulators should also take into account the distributional effects of new data when deciding whether to permit its use in particular markets.

One way to interpret the attributes and circumstances in our model is as standins for covariates with different levels of persistence in a dynamic model, where the agent exerts effort over multiple periods and his type evolves over time. Generalizing our results to a many-period setting is technically challenging under non-Gaussian information structures, since effort deviations today may distort future returns to effort. Nonetheless, doing so would permit a richer study of the welfare implications of forecasting from data with varying persistence, making it an important avenue for future research.

Appendix A: Alternative welfare specifications

In this appendix, we extend our welfare analysis to consider alternative environments in which effort improves future as well as current outcomes ("learning-by-doing") or is partially dissipative ("gaming" effort).

A.1 *Learning-by-doing*

In some applications, effort may improve future as well as current outcomes, for instance in labor market settings featuring learning-by-doing. In that case, the agent's type is not constant over time but instead improves with past effort, and effort has socially beneficial effects over multiple periods.

Our model can be modified to accommodate this feature by allowing the agent's type $\theta^{(t)}$, which determines the average outcome in period t, to be time-dependent. Concretely, we will suppose that $\theta^{(2)} = \theta^{(1)} + \gamma \cdot e$, where $\gamma > 0$ is a learning-by-doing parameter. Period-1 output is

$$Y = e + \theta^{(1)} + \varepsilon,$$

while the agent's period-2 reputational reward is $\mathbb{E}(\theta^{(2)} \mid Y)$.

The presence of learning by doing does not affect equilibrium effort, because the agent's reputational reward is based on the market's forecast of his effort (which is fixed) rather than his true effort. This expectation is

$$\mathbb{E}^{e^*}\big(\theta^{(2)} \mid Y\big) = (1 + \gamma) \cdot e^* + \mathbb{E}^{e^*}\big(\theta^{(1)} \mid Y\big).$$

Exerting additional effort is therefore valuable to the agent only insofar as it improves the market's forecast of $\theta^{(1)}$, exactly as in our main model. Thus, equation (4) continues to characterize equilibrium effort.

The socially optimal effort level, however, becomes $e^{FB} = 1 + \gamma$ in this model. Equilibrium effort therefore falls below the first-best level for a broader range of reputation weights β as the learning-by-doing parameter γ increases. An analogue of Theorem 3 continues to hold, where the threshold reputation weights β^* and β_* are increasing in γ. In other words, increased learning-by-doing makes circumstances (which boost effort) more attractive and attributes (which reduce it) less so at any given reputational weight.

A.2 *"Gaming" effort*

In other applications, effort may be dissipative and serve to distort a signal of quality without producing social value. This possibility may arise, for instance, in labor market settings in which a worker can spend time performing "influence activities" to increase the visibility of his accomplishments (as in Milgrom and Roberts (1988)). It may also arise in educational settings where the outcome variable is a test score that can be improved by test prep with no further educational value (as in Frankel and Kartik (2022)).

To accommodate this possibility, our welfare criterion can be modified to discount the welfare benefits of effort:

$$w(\theta, e) = \theta + \delta \cdot e - \frac{1}{2}e^2,$$

where $\delta \in [0, 1)$ measures the proportion of effort which is socially beneficial. When $\delta = 0$, effort is totally unproductive, while $\delta \in (0, 1)$ captures situations in which some fraction of effort contributes social value.

The dissipative nature of effort has no impact on equilibrium effort, but reduces the first-best effort level to $e^{FB} = \delta$. Equilibrium effort will therefore exceed the first-best level for a broader range of reputation weights β as effort becomes increasingly dissipative. A result analogous to Theorem 3 can be established in this setting, with the threshold reputation weights β^* and β_* increasing in δ. One interesting case is $\delta = 0$, in which effort is fully dissipative effort. In that case, measuring new attributes improves welfare while measuring new circumstances diminishes it, regardless of the reputation weight β. (The one exception is for an attribute with significant disparate impact, which may still be welfare-reducing for all β.)

Appendix B: Results for a general baseline

The results of Section 3 can be straightforwardly generalized to accommodate settings in which some covariates are initially measured by the market. Given any sets $\mathcal{J} \subseteq \{1, \ldots, J\}$ of measured attributes and $\mathcal{K} \subseteq \{1, \ldots, K\}$ of measured circumstances, define

$$\eta_{\mathcal{J}} \equiv \sum_{j \notin \mathcal{J}} \theta_j + u_\theta, \qquad \delta_{\mathcal{K}} \equiv \sum_{k \notin \mathcal{K}} \varepsilon_k + u_\varepsilon$$

to be the sums of all *unmeasured* components of the agent's type and shock.

Fix a baseline family $(\mathcal{J}, \mathcal{K})$ of measured covariates. Affiliation and strong homoskedasticity may be generalized to this environment as follows.

DEFINITION B.1 (Affiliation). The attribute $j' \notin \mathcal{J}$ is $\mathcal{J}$-*affiliated* if $(\theta_{j'}, \eta_{\mathcal{J} \cup \{j'\}})$ is affiliated conditional on $a_{\mathcal{J}}$. The circumstance $k' \notin \mathcal{K}$ is $\mathcal{K}$-*affiliated* if $(\varepsilon_{k'}, \delta_{\mathcal{K} \cup \{k'\}})$ is affiliated conditional on $c_{\mathcal{K}}$.

DEFINITION B.2 (Strong homoskedasticity). The attribute $j' \notin \mathcal{J}$ satisfies $\mathcal{J}$-*strong homoskedasticity* if $\eta_{\mathcal{J} \cup \{j'\}} - \mathbb{E}(\eta_{\mathcal{J} \cup \{j'\}} \mid a_{\mathcal{J} \cup \{j'\}})$ is independent of $a_{j'}$ conditional on $a_{\mathcal{J}}$. The circumstance $k' \notin \mathcal{K}$ satisfies $\mathcal{K}$-*strong homoskedasticity* if $\delta_{\mathcal{K} \cup \{k'\}} - \mathbb{E}(\delta_{\mathcal{K} \cup \{k'\}} \mid c_{\mathcal{K} \cup \{k'\}})$ is independent of $c_{k'}$ conditional on $c_{\mathcal{K}}$.

As formulated, these definitions apply across all $(\mathcal{J}, \mathcal{K})$-*subpopulations* of agents, where each subpopulation consists of all agents sharing a particular realization of $(a_{\mathcal{J}}, c_{\mathcal{K}})$. They could alternatively be formulated more narrowly to apply only for a particular set of realized covariates, if the analyst is primarily interested in the impact of a new covariate on a particular subpopulation of agents.

Assumptions 1 and 2, which imposed log-concavity on latent variables and bound-edness of derivatives of conditional expectations, must also be extended for a general set of baseline measured covariates. We split these conditions into an assumption we maintain in the baseline environment, and a condition imposed on newly measured covariates.

ASSUMPTION B.1 (Baseline admissibility). *The conditional distributions $\theta \mid a_{\mathcal{J}}$ and $\varepsilon \mid c_{\mathcal{K}}$ have log-concave density functions, and for every effort level e and realization of $(a_{\mathcal{J}}, c_{\mathcal{K}})$, the derivative $\frac{\partial}{\partial Y}\mathbb{E}^e(\theta \mid Y, a_{\mathcal{J}}, c_{\mathcal{K}})$ exists and is uniformly bounded across all realizations of Y.*

DEFINITION B.3 (Admissible covariates). An attribute $j' \notin \mathcal{J}$ is $\mathcal{J}$-*admissible* if: (1) $\eta_{\mathcal{J}\cup\{j'\}} \mid a_{\mathcal{J}\cup\{j'\}}$ has a log-concave density function for every realization of $a_{\mathcal{J}\cup\{j'\}}$; and (2) for every effort level e and realization of covariates $(a_{\mathcal{J}\cup\{j'\}}, c_{\mathcal{K}})$, the derivative $\frac{\partial}{\partial Y}\mathbb{E}^e(\theta \mid Y, a_{\mathcal{J}\cup\{j'\}}, c_{\mathcal{K}})$ exists and is uniformly bounded across all realizations of Y.

A circumstance $k' \notin \mathcal{K}$ is $\mathcal{K}$-*admissible* if: (1) $\delta_{\mathcal{K}\cup\{k'\}} \mid c_{\mathcal{K}\cup\{k'\}}$ has a log-concave density function for every realization of $c_{\mathcal{K}\cup\{k'\}}$; and (2) for every every effort level e and realization of covariates $(a_{\mathcal{J}}, c_{\mathcal{K}\cup\{k'\}})$, the derivative $\frac{\partial}{\partial Y}\mathbb{E}^e(\theta \mid Y, a_{\mathcal{J}}, c_{\mathcal{K}\cup\{k'\}})$ exists and is uniformly bounded across all realizations of Y.

With these concepts, we can generalize Theorems 1 and 2 as follows.

THEOREM B.1. *Suppose Assumption B.1 holds.*

(a) *If attribute j' is $\mathcal{J}$-admissible and satisfies $\mathcal{J}$-affiliation, then measuring it weakly reduces average effort within each $(\mathcal{J}, \mathcal{K})$-subpopulation.*

(b) *If circumstance k' is $\mathcal{K}$-admissible and satisfies $\mathcal{K}$-affiliation, then measuring it weakly increases average effort within each $(\mathcal{J}, \mathcal{K})$-subpopulation.*

THEOREM B.2. *Suppose Assumption B.1 holds.*

(a) *If attribute j' is $\mathcal{J}$-admissible and satisfies $\mathcal{J}$-strong homoskedasticity, then measuring it weakly reduces every agent's effort. Further, the magnitude of the effort change is the same for every agent in each $(\mathcal{J}, \mathcal{K})$-subpopulation.*

(b) *If circumstance k' is $\mathcal{K}$-admissible and satisfies $\mathcal{K}$-strong homoskedasticity, then measuring it weakly increases every agent's effort. Further, the magnitude of the effort change is the same for every agent in each $(\mathcal{J}, \mathcal{K})$-subpopulation.*

These results can be applied repeatedly to assess the impact of measuring multiple covariates, so long as admissibility and the corresponding statistical condition (affiliation or strong homoskedasticity) holds for each of the measured covariates relative to its respective baseline. Note in particular that the exponential and multivariate normal settings of Examples 2 and 3 satisfy affiliation and strong homoskedasticity, respectively, for any baseline and newly measured covariate.

Our welfare results also hold under general baselines using appropriate notions of regularity and strict regularity.

Definition B.4. Fix a baseline family of measured covariates $(\mathcal{J}, \mathcal{K})$. Then:

- An attribute $j' \notin \mathcal{J}$ is $(\mathcal{J}, \mathcal{K})$-*regular* if, conditional on any realization of $(a_{\mathcal{J}}, c_{\mathcal{K}})$, measuring j' weakly reduces the marginal value of effort on average. It is *strictly $(\mathcal{J}, \mathcal{K})$-regular* if the reduction is strict for a positive fraction of realizations of $(a_{\mathcal{J}}, c_{\mathcal{K}})$.

- A circumstance $k' \notin \mathcal{K}$ is $(\mathcal{J}, \mathcal{K})$-*regular* if, conditional on any realization of $(a_{\mathcal{J}}, c_{\mathcal{K}})$, measuring k' weakly increases the marginal value of effort on average. It is *strictly $(\mathcal{J}, \mathcal{K})$-regular* if the increase is strict for a positive fraction of realizations of $(a_{\mathcal{J}}, c_{\mathcal{K}})$.

Weak regularity imposes monotonicity separately on each subpopulation of agents. Strict regularity imposes the stronger requirement of strict monotonicity for a positive fraction of agents. (In the special case of a baseline with no observed covariates, strict regularity trivially implies strict monotonicity for all agents, corresponding to Definition 3.)

The following result generalizes Theorem 3 to general baselines.

Theorem B.3. *Fix a baseline family of measured covariates $(\mathcal{J}, \mathcal{K})$.*

(a) *For every $(\mathcal{J}, \mathcal{K})$-regular attribute $j' \notin \mathcal{J}$, there exists a threshold reputation weight $\beta^* \in (0, 1]$ such that measuring j' is welfare-improving if and only if $\beta > \beta^*$. Moreover, $\beta^* < 1$ if and only if*

$$\mathbb{E}\big(MV^2_{\mathcal{J} \cup \{j'\}, \mathcal{K}}\big) < \mathbb{E}\big(MV^2_{\mathcal{J}, \mathcal{K}}\big) \tag{B.1}$$

where $MV_{\mathcal{J}, \mathcal{K}}$ is as defined in (3).

(b) *For every $(\mathcal{J}, \mathcal{K})$-regular circumstance $k' \notin \mathcal{K}$, there exists a threshold reputation weight $\beta_* \in [0, 1)$ such that measuring k' is welfare-improving if and only if $\beta < \beta_*$. Moreover, $\beta_* > 0$ if and only if k' is strictly $(\mathcal{J}, \mathcal{K})$-regular.*

Appendix C: Welfare under general convex costs

Theorem 3, our main welfare result, can be extended to nonquadratic effort cost functions under the same conditions as Proposition 3, assuming that effort costs follow a power law. We state and prove this result for general baselines, as in the analysis of Appendix B.

Proposition C.1. *Suppose that $C(e) = Ae^k$ for some $A > 0$ and $k > 1$. Fix a baseline family of measured covariates $(\mathcal{J}, \mathcal{K})$.*

(a) *Suppose that the market measures the additional regular attribute $j' \notin \mathcal{J}$. If either $k \geq 2$ or else all agents change their effort in the same direction, then there exists a threshold reputation weight $\beta^* \in (0, 1]$ such that the measurement is welfare-improving if and only if $\beta > \beta^*$. If j' is strictly regular and all agents change their effort in the same direction, then $\beta^* < 1$.*

(b) *Suppose that the market measures the additional regular circumstance $k' \notin \mathcal{K}$. If either $k \leq 2$ or else all agents change their effort in the same direction, then there exists a threshold reputation weight $\beta_* \in [0, 1)$ such that the measurement is welfare-improving if and only if $\beta < \beta_*$. If k' is strictly regular, then $\beta_* > 0$.*

For general cost functions, fully characterizing how aggregate welfare changes with β becomes intractable. However, it can be shown that if the effort cost function is approximately quadratic near zero, then when β is small, newly measured regular attributes reduce aggregate welfare and newly measured regular circumstances increase them; while for large β, these effects reverse.

PROOF OF PROPOSITION C.1. If a newly measured covariate is regular but not strictly regular, then expected effort is unchanged while the distribution of effort in each subpopulation undergoes a mean-preserving spread under the measurement. Then since effort costs are strictly convex, aggregate welfare must at least weakly decrease no matter the value of β, corresponding to $\beta^* = 1$ for an attribute and $\beta_* = 0$ for a circumstance. For the remainder of the proof, we assume that the newly measured attribute is strictly regular.

Let $\delta \equiv \beta/(1 - \beta)$. Note that $e^*_{\mathcal{J},\mathcal{K}}$ depends on β only through δ, and we will write $e^*_{\mathcal{J},\mathcal{K}}(\delta)$ to make this dependence explicit.

We first consider the case in which the market measures a new attribute j'. All notation is as in the proof of Theorem B.3. Define

$$\Delta \overline{W}(\delta) \equiv \mathbb{E}\big[w_0\big(e^*_{\mathcal{J}',\mathcal{K}}(\delta)\big) - w_0\big(e^*_{\mathcal{J},\mathcal{K}}(\delta)\big)\big],$$

where

$$w_0(e) \equiv e - C(e).$$

Recall that given any family of measured covariates, equilibrium effort in a given subpopulation satisfies $e^* = (C')^{-1}(\delta \cdot MV)$, where MV is the corresponding subpopulation marginal value of effort. Thus,

$$\frac{\partial e^*}{\partial \delta} = \frac{MV}{C''\big((C')^{-1}(\delta \cdot MV)\big)}$$

and

$$\frac{\partial}{\partial \delta} w(e^*) = \big(1 - C'(e^*)\big)\frac{\partial e^*}{\partial \delta} = (1 - \delta \cdot MV) \cdot MV \cdot \frac{1}{C''\big((C')^{-1}(\delta \cdot MV)\big)}.$$

When $C(e) = Ae^k$, we have

$$\frac{1}{C''\big((C')^{-1}(\delta \cdot MV)\big)} = A' \cdot (\delta \cdot MV)^{\frac{1}{k-1}-1},$$

where $A' \equiv 1/\big((k-1)(Ak)^{1/(k-1)}\big) > 0$. Hence,

$$\frac{\partial}{\partial \delta} w(e^*) = A' \cdot \delta^{\frac{1}{k-1}-1} \cdot (1 - \delta \cdot MV) \cdot MV^{\frac{1}{k-1}}.$$

Using this identity, the derivative of $\Delta \overline{W}$ may be written

$$\frac{\partial}{\partial \delta}\Delta \overline{W}(\delta) = A' \cdot \delta^{\frac{1}{k-1}-1} \cdot \mathbb{E}\big[(1 - \delta \cdot MV_{\mathcal{J}',\mathcal{K}}) \cdot MV_{\mathcal{J}',\mathcal{K}}^{\frac{1}{k-1}}$$

$$- (1 - \delta \cdot MV_{\mathcal{J},\mathcal{K}}) \cdot MV_{\mathcal{J},\mathcal{K}}^{\frac{1}{k-1}}\big].$$

This expression crosses zero at most once for $\delta > 0$, and

$$\lim_{\delta \to 0} \delta^{1-\frac{1}{k-1}} \frac{\partial}{\partial \delta}\Delta \overline{W}(\delta) = \mathbb{E}\big[MV_{\mathcal{J}',\mathcal{K}}^{\frac{1}{k-1}} - MV_{\mathcal{J},\mathcal{K}}^{\frac{1}{k-1}}\big] \tag{C.1}$$

while

$$\lim_{\delta \to \infty} \delta^{-\frac{1}{k-1}} \frac{\partial}{\partial \delta}\Delta \overline{W}(\delta) = \mathbb{E}\big[MV_{\mathcal{J},\mathcal{K}}^{\frac{k}{k-1}} - MV_{\mathcal{J}',\mathcal{K}}^{\frac{k}{k-1}}\big]. \tag{C.2}$$

Strict regularity of j' requires that $\mathbb{E}[MV_{\mathcal{J}',\mathcal{K}} \mid a_{\mathcal{J}}, c_{\mathcal{K}}] \leq MV_{\mathcal{J},\mathcal{K}}$ for all realizations of $(a_{\mathcal{J}}, c_{\mathcal{K}})$, with the inequality strict with positive probability. If all agents change their effort in the same direction, $MV_{\mathcal{J}',\mathcal{K}} \leq MV_{\mathcal{J},\mathcal{K}}$ everywhere. This inequality combined with the previous implication of strict regularity implies that $MV_{\mathcal{J}',\mathcal{K}} < MV_{\mathcal{J},\mathcal{K}}$ with positive probability, in which case (C.1) is strictly negative while (C.2) is strictly positive. And in general, if $k \geq 2$ then Jensen's inequality and strict regularity imply that

$$\mathbb{E}\big[MV_{\mathcal{J}',\mathcal{K}}^{\frac{1}{k-1}} \mid a_{\mathcal{J}}, c_{\mathcal{K}}\big] \leq \mathbb{E}[MV_{\mathcal{J}',\mathcal{K}} \mid a_{\mathcal{J}}, c_{\mathcal{K}}]^{\frac{1}{k-1}} \leq MV_{\mathcal{J},\mathcal{K}}^{\frac{1}{k-1}},$$

with the final inequality strict with positive probability. Hence,

$$\mathbb{E}\big[MV_{\mathcal{J}',\mathcal{K}}^{\frac{1}{k-1}}\big] < \mathbb{E}\big[MV_{\mathcal{J},\mathcal{K}}^{\frac{1}{k-1}}\big].$$

Thus, the derivative of $\Delta \overline{W}$ is strictly negative near zero if either $k \geq 2$ or all agents change their effort in the same direction, and in the latter case it eventually becomes positive for large δ.

Next, observe that when $\delta = 0$, $e^* = 0$ no matter the value of MV, and so $\Delta \overline{W}(0) = 0$. Then given single-crossing of the derivative of $\Delta \overline{W}(\delta)$, it must be that $\Delta \overline{W}(\delta)$ crosses zero at most once for $\delta > 0$, and any crossing point is from below. Let $\delta^* > 0$ denote this crossing point, with $\delta^* = \infty$ in the case that no crossing occurs. Then $\Delta \overline{W}(\delta) < 0$ for $\delta < \delta^*$, while $\overline{W}(\delta) > 0$ for $\delta > \delta^*$. Additionally, if all agents change their effort in the same direction, for large δ the derivative of $\Delta \overline{W}(\delta)$ is positive and approximately proportional to $\delta^{\frac{1}{k-1}}$, and thus becomes unboundedly large. It follows that eventually $\Delta \overline{W}(\delta) > 0$ for δ sufficiently large, so $\delta^* < \infty$ in this case. Letting $\beta^* \equiv \delta^*/(1 + \delta^*)$ yields the desired reputation weight threshold.

The case of a newly measured circumstance k' follows along similar lines, with two main differences. First, strict regularity and Jensen's inequality imply that

$$\mathbb{E}\big[MV_{\mathcal{J},\mathcal{K}'}^{\frac{k}{k-1}} \mid a_{\mathcal{J}}, c_{\mathcal{K}}\big] \geq \mathbb{E}[MV_{\mathcal{J},\mathcal{K}'} \mid a_{\mathcal{J}}, c_{\mathcal{K}}]^{\frac{k}{k-1}} \geq MV_{\mathcal{J},\mathcal{K}}^{\frac{k}{k-1}},$$

with the final inequality strict with positive probability. Hence,

$$\mathbb{E}\big[MV_{\mathcal{J},\mathcal{K}}^{\frac{k}{k-1}} - MV_{\mathcal{J},\mathcal{K}'}^{\frac{k}{k-1}}\big] < 0$$

and the derivative of $\Delta\overline{W}$ is negative for large δ, whether or not all agents change their effort in the same direction. Thus, $\delta_* < \infty$ always. Second, the parameter restriction under which Jensen's inequality implies that the derivative of $\Delta\overline{W}$ is positive near zero is $k \leq 2$.   □

APPENDIX D: PROOFS OF RESULTS FROM THE MAIN TEXT

D.1 *Preliminary work: Characterization of MV*

Fix a family of measured covariates $(\mathcal{J}, \mathcal{K})$.

LEMMA D.1. *Suppose that Assumption B.1 holds. Then the equilibrium marginal value of effort is*

$$MV_{\mathcal{J},\mathcal{K}} = \mathbb{E}\left[\frac{\partial}{\partial Y^0}\mathbb{E}\big[\eta_{\mathcal{J}} \mid Y^0, a_{\mathcal{J}}, c_{\mathcal{K}}\big]\,\bigg|\, a_{\mathcal{J}}, c_{\mathcal{K}}\right],$$

where $Y^0 \equiv \theta + \varepsilon$.

PROOF. Throughout this proof, fix a set of realizations of $(a_{\mathcal{J}}, c_{\mathcal{K}})$, and condition all distributions on these realizations. To economize on notation, we will suppress explicit conditioning on these covariates.

Let $\rho_{y|\eta,e}(y \mid t, e)$ be the conditional density of $Y \mid \eta_{\mathcal{J}}, e$ and $\rho_{y|e}(y \mid e)$ be the conditional density of $Y \mid e$. Because effort affects the outcome as an additive shift, $\rho_{y|\eta,e}(y \mid t, e) = \rho_{y|\eta,e}(y - e \mid t, 0)$ and $\rho_{y|e}(y \mid e) = \rho_{y|e}(y - e \mid 0)$ for every (y, t, e). So, let $\rho_{\eta|y,e}(t \mid y, e)$ be the conditional density of $\eta_{\mathcal{J}} \mid Y, e$, and let $\rho_\eta(t)$ be the conditional density of $\eta_{\mathcal{J}}$. Then by Bayes' rule,

$$\rho_{\eta|y,e}(t \mid y, e) = \frac{\rho_{y|\eta,e}(y \mid t, e)\rho_\eta(t)}{\rho_{y|e}(y \mid e)} = \frac{\rho_{y|\eta,e}(y - e \mid t, 0)\rho_\eta(t)}{\rho_{y|e}(y - e \mid 0)} = \rho_{\eta|y,e}(t \mid y - e, 0).$$

Hence,

$$\begin{aligned}
\mathbb{E}^{e^*}[\theta \mid Y = y] &= \sum_{j \in \mathcal{J}} \theta_j + \mathbb{E}^{e^*}[\eta_{\mathcal{J}} \mid Y = y] \\
&= \sum_{j \in \mathcal{J}} \theta_j + \int t\rho_{\eta|y,e}(t \mid y, e^*)\,dt \\
&= \sum_{j \in \mathcal{J}} \theta_j + \int t\rho_{\eta|y,e}(t \mid y - e^*, 0)\,dt \\
&= \sum_{j \in \mathcal{J}} \theta_j + \mathbb{E}^0[\eta_{\mathcal{J}} \mid Y = y - e^*].
\end{aligned}$$

Under the measure corresponding to $e = 0$, the variable Y is equal to Y^0 almost surely. So,

$$\mathbb{E}^0[\eta_{\mathcal{J}} \mid Y = y - e^*] = \mathbb{E}[\eta_{\mathcal{J}} \mid Y^0 = y - e^*].$$

Now,

$$\begin{aligned}
\mathbb{E}^e[\mathbb{E}^{e^*}[\theta \mid Y]] &= \int dy\, \rho_{y|e}(y \mid e)\mathbb{E}^{e^*}[\theta \mid Y = y] \\
&= \int dy\, \rho_{y|e}(y \mid e)\left(\sum_{j \in \mathcal{J}} \theta_j + \mathbb{E}[\eta_{\mathcal{J}} \mid Y^0 = y - e^*]\right) \\
&= \int dy\, \rho_{y|e}(y - e \mid 0)\left(\sum_{j \in \mathcal{J}} \theta_j + \mathbb{E}[\eta_{\mathcal{J}} \mid Y^0 = y - e^*]\right),
\end{aligned}$$

and so by making the variable substitution $y' = y - e$ we may write

$$\mathbb{E}^e[\mathbb{E}^{e^*}[\theta \mid Y]] = \int dy'\, \rho_{y|e}(y' \mid 0)\left(\sum_{j \in \mathcal{J}} \theta_j + \mathbb{E}[\eta_{\mathcal{J}} \mid Y^0 = y' - e^* + e]\right).$$

Differentiating with respect to e and invoking Assumption B.1 to justify applying the dominated convergence theorem yields

$$\frac{\partial}{\partial e}\mathbb{E}^e[\mathbb{E}^{e^*}[\theta \mid Y]]\bigg|_{e=e^*} = \int dy'\, \rho_{y|e}(y' \mid 0)\frac{\partial}{\partial Y^0}\mathbb{E}[\eta_{\mathcal{J}} \mid Y^0]\bigg|_{Y^0 = y'}.$$

Recall that $Y = Y^0$ conditional on $e = 0$, so $\rho_{y|e}(y' \mid 0)$ is the density of Y^0. The rhs of the previous expression may therefore be written

$$\frac{\partial}{\partial e}\mathbb{E}^e[\mathbb{E}^{e^*}[\theta \mid Y]]\bigg|_{e=e^*} = \mathbb{E}\left[\frac{\partial}{\partial Y^0}\mathbb{E}[\eta_{\mathcal{J}} \mid Y^0]\right],$$

as desired. $\square$

D.2 *Proofs of Theorems 1 and B.1*

We prove Theorem B.1, from which Theorem 1 follows immediately as a corollary.

D.2.1 *Part (a)*   Fix a baseline family of measured covariates $(\mathcal{J}, \mathcal{K})$. As established in Lemma D.1, the marginal value of effort in the baseline is

$$MV_{\mathcal{J},\mathcal{K}} = \mathbb{E}\left[\frac{\partial}{\partial Y^0}\mathbb{E}[\eta_{\mathcal{J}} \mid Y^0, a_{\mathcal{J}}, c_{\mathcal{K}}]\,\middle|\, a_{\mathcal{J}}, c_{\mathcal{K}}\right],$$

where $Y^0 \equiv \theta + \varepsilon$ is the baseline value of the outcome after subtracting out the agent's effort.

Now suppose the market additionally observes the additional attribute $j' \notin \mathcal{J}$, and let $\mathcal{J}' \equiv \mathcal{J} \cup \{j'\}$. Under the expanded family of measured covariates, the marginal value

of effort becomes

$$MV_{\mathcal{J}',\mathcal{K}} = \mathbb{E}\left[\frac{\partial}{\partial Y^0}\mathbb{E}\left[\eta_{\mathcal{J}'} \mid Y^0, a_{\mathcal{J}'}, c_{\mathcal{K}}\right] \,\middle|\, a_{\mathcal{J}'}, c_{\mathcal{K}}\right].$$

Note that conditional on $(a_{\mathcal{J}}, c_{\mathcal{K}})$, $MV_{\mathcal{J}',\mathcal{K}}$ is a random variable whose value is a function of the realization of $a_{j'}$.

Because $f^{j'}$ is a one-to-one mapping, conditioning on the value of $a_{j'}$ is equivalent to conditioning on the value of $\theta_{j'} = f^{j'}(a_{j'})$. So, we may equivalently write the agent's marginal value of effort under the expanded set of covariates as

$$MV_{\mathcal{J}',\mathcal{K}} = \mathbb{E}\left[\frac{\partial}{\partial Y^0}\mathbb{E}\left[\eta_{\mathcal{J}'} \mid Y^0, \theta_{j'}, a_{\mathcal{J}}, c_{\mathcal{K}}\right] \,\middle|\, \theta_{j'}, a_{\mathcal{J}}, c_{\mathcal{K}}\right].$$

LEMMA D.2. $(\eta_{\mathcal{J}}, \theta_{j'}, Y^0)$ *are affiliated conditional on* $(a_{\mathcal{J}}, c_{\mathcal{K}})$.

PROOF. Fix a set of realizations of $(a_{\mathcal{J}}, c_{\mathcal{K}})$, and condition all distributions on their values. To economize on notation, explicit conditioning on these covariates will be suppressed. Let $\rho_{\eta,\theta,Y}(u, t, y)$ be the joint density of $(\eta_{\mathcal{J}}, \theta_{j'}, Y^0)$. We will show that $\rho_{\eta,\theta,Y}$ is log-supermodular.[28]

Let $\rho_\theta(t)$ be the density of $\theta_{j'}$, $\rho_{\eta|\theta}(u \mid t)$ be the conditional density of $\eta_{\mathcal{J}} \mid \theta_{j'}$, and $\rho_{Y|\eta}(y \mid u)$ be the conditional density of $Y^0 \mid \eta_{\mathcal{J}}$. Note that conditional on $\eta_{\mathcal{J}}$, Y^0 is independent of $\theta_{j'}$, and so

$$\rho_{\eta,\theta,Y}(u, t, y) = \rho_\theta(t)\rho_{\eta|\theta}(u \mid t)\rho_{Y|\eta}(y \mid u).$$

It is therefore sufficient to show that $\rho_{Y|\eta}$ and $\rho_{\eta|\theta}$ are log-supermodular.

First, consider $\rho_{Y|\eta}$. Define

$$\mu_{(\mathcal{J},\mathcal{K})} \equiv \sum_{j \in \mathcal{J}} \theta_j + \sum_{k \in \mathcal{K}} \varepsilon_k + \eta_{\mathcal{J}}.$$

Then Y^0 may be decomposed as

$$Y^0 = \mu_{(\mathcal{J},\mathcal{K})} + \eta_{\mathcal{J}} + \delta_{\mathcal{K}},$$

where $\mu_{(\mathcal{J},\mathcal{K})}$ is a constant. Let $\rho_\delta(z)$ be the density of $\delta_{\mathcal{K}}$. Then

$$\rho_{Y|\eta}(y \mid u) = \rho_\delta(y - \mu_{(\mathcal{J},\mathcal{K})} - u).$$

Under Assumption 1, ρ_δ is log-concave, meaning $\rho_{Y|\eta}$ is log-supermodular.

As for $\rho_{\eta|\theta}$, let $\rho_{\eta'|\theta}(w \mid t)$ be the conditional density of $\eta_{\mathcal{J}'} \mid \theta_{j'}$. As $\eta_{\mathcal{J}} = \theta_{j'} + \eta_{\mathcal{J}'}$, it follows that

$$\rho_{\eta|\theta}(u \mid t) = \rho_{\eta'|\theta}(u - t \mid t).$$

---

[28]A vector of random variables possessing a joint density function is affiliated if and only if its density function is log-supermodular, that is, the logarithm of its density function is supermodular. We make use of the following well-known facts: any product of log-supermodular functions is log-supermodular, and a twice continuously differentiable, strictly positive function $f(\mathbf{x})$ is log-supermodular if and only if $\partial^2 \log f/\partial x_i x_j \geq 0$ for every pair of components $i \neq j$ and all $\mathbf{x}$.

Hence, by the chain rule,

$$\frac{\partial^2}{\partial u \partial t} \log \rho_{\eta|\theta}(u \mid t) = \left[\frac{\partial^2}{\partial w \partial t} \log \rho_{\eta'|\theta}(w \mid t) - \frac{\partial^2}{\partial w^2} \log \rho_{\eta'|\theta}(w \mid t) \right]_{w=u-t}.$$

Under Assumption 1, $\rho_{\eta'|\theta}$ is log-concave and so the second term is nonnegative. Meanwhile, $\mathcal{J}$-affiliation of j' implies that the first term is also nonnegative. Hence,

$$\frac{\partial^2}{\partial u \partial t} \log \rho_{\eta|\theta}(u \mid t) \geq 0,$$

establishing the desired log-supermodularity.   □

The marginal value of effort given observation of attribute j', $MV_{\mathcal{J}',\mathcal{K}}$, can be compared to the marginal value of effort in the baseline, $MV_{\mathcal{J},\mathcal{K}}$, as follows. Fix a set of realizations of $(a_{\mathcal{J}}, c_{\mathcal{K}})$, and define

$$F_\theta(t \mid y) \equiv \Pr\left(\theta_{j'} \leq t \mid Y^0 = y, a_{\mathcal{J}}, c_{\mathcal{K}}\right)$$

to be the conditional distribution function of $\theta_{j'}$ given the outcome Y^0, and

$$\phi(y, t) \equiv \mathbb{E}\left[\eta_{\mathcal{J}'} \mid Y^0 = y, \theta_{j'} = t, a_{\mathcal{J}}, c_{\mathcal{K}}\right]$$

to be the conditional expectation of $\eta_{\mathcal{J}'}$ given Y^0 and $\theta_{j'}$.

By the law of total probability,

$$\mathbb{E}\left[\eta_{\mathcal{J}} \mid Y^0 = y, a_{\mathcal{J}}, c_{\mathcal{K}}\right] = \int dF_\theta(t \mid y)(t + \phi(y, t)),$$

and so the change in the conditional expectation of the unobserved $\eta_{\mathcal{J}}$ as Y^0 moves from y to $y' > y$ is

$$\mathbb{E}\left[\eta_{\mathcal{J}} \mid Y^0 = y', a_{\mathcal{J}}, c_{\mathcal{K}}\right] - \mathbb{E}\left[\eta_{\mathcal{J}} \mid Y^0 = y, a_{\mathcal{J}}, c_{\mathcal{K}}\right]$$
$$= \int dF_\theta(t \mid y')(t + \phi(y', t)) - \int dF_\theta(t \mid y)(t + \phi(y, t)). \tag{D.1}$$

This difference can be signed using Lemma D.2: Since $(\eta_{\mathcal{J}}, \theta_{j'})$ are affiliated conditional on Y^0, the expression $t + \phi(y, t)$ is nondecreasing in t. And since $(\theta_{j'}, Y^0)$ are affiliated, the expectation $\mathbb{E}[\pi(\theta_{j'}) \mid Y^0, a_{\mathcal{J}}, c_{\mathcal{K}}]$ is nondecreasing in Y^0 for any increasing function π. Thus,

$$\int dF_\theta(t \mid y')(t + \phi(y, t))$$

is nondecreasing in y', and so the expression in (D.1) can be bounded below by

$$\int dF_\theta(t \mid y)(\phi(y', t) - \phi(y, t)).$$

It follows that

$$\frac{\partial}{\partial y}\mathbb{E}\big[\eta_{\mathcal{J}} \mid Y^0 = y, a_{\mathcal{J}}, c_{\mathcal{K}}\big] \geq \int dF_{\theta}(t \mid y)\frac{\partial \phi}{\partial y}(y, t).$$

The lhs is the marginal improvement in the posterior expectation of $\eta_{\mathcal{J}}$ when the realization of Y^0 is increased. The rhs is the expected marginal improvement in the posterior expectation of $\eta_{\mathcal{J}'}$ when it is conditioned on the manipulated realization of Y^0 as well as the *unmanipulated* realization of $\theta_{j'}$.

To complete the proof, rewrite this inequality as

$$\frac{\partial}{\partial Y^0}\mathbb{E}\big[\eta_{\mathcal{J}} \mid Y^0, a_{\mathcal{J}}, c_{\mathcal{K}}\big] \geq \mathbb{E}\left[\frac{\partial}{\partial Y^0}\mathbb{E}\big[\eta_{\mathcal{J}'} \mid Y^0, \theta_{j'}, a_{\mathcal{J}}, c_{\mathcal{K}}\big] \,\bigg|\, Y^0, a_{\mathcal{J}}, c_{\mathcal{K}}\right].$$

Taking the expectation of each side conditional on $(a_{\mathcal{J}}, c_{\mathcal{K}})$ yields

$$MV_{\mathcal{J},\mathcal{K}} \geq \mathbb{E}\left[\frac{\partial}{\partial Y^0}\mathbb{E}\big[\eta_{\mathcal{J}'} \mid Y^0, \theta_{j'}, a_{\mathcal{J}}, c_{\mathcal{K}}\big] \,\bigg|\, a_{\mathcal{J}}, c_{\mathcal{K}}\right].$$

By the law of iterated expectations, the rhs may be expanded as

$$\mathbb{E}\left[\frac{\partial}{\partial Y^0}\mathbb{E}\big[\eta_{\mathcal{J}'} \mid Y^0, \theta_{j'}, a_{\mathcal{J}}, c_{\mathcal{K}}\big] \,\bigg|\, a_{\mathcal{J}}, c_{\mathcal{K}}\right]$$
$$= \mathbb{E}\left[\mathbb{E}\left[\frac{\partial}{\partial Y^0}\mathbb{E}\big[\eta_{\mathcal{J}'} \mid Y^0, \theta_{j'}, a_{\mathcal{J}}, c_{\mathcal{K}}\big] \,\bigg|\, \theta_{j'}, a_{\mathcal{J}}, c_{\mathcal{K}}\right] \,\bigg|\, a_{\mathcal{J}}, c_{\mathcal{K}}\right]$$
$$= \mathbb{E}[MV_{\mathcal{J}',\mathcal{K}} \mid a_{\mathcal{J}}, c_{\mathcal{K}}].$$

Therefore,

$$MV_{\mathcal{J},\mathcal{K}} \geq \mathbb{E}[MV_{\mathcal{J}',\mathcal{K}} \mid a_{\mathcal{J}}, c_{\mathcal{K}}].$$

This inequality holds for every realization of $(a_{\mathcal{J}}, c_{\mathcal{K}})$. Thus, the marginal value of effort in every $(\mathcal{J}, \mathcal{K})$-subpopulation is weakly larger than the expected marginal value once attribute j' is additionally measured, as desired.

D.2.2 *Part (b)*   Suppose that the market observes the additional circumstance $k' \notin \mathcal{K}$. Let $\mathcal{K}' \equiv \mathcal{K} \cup \{k'\}$. Under the expanded family of measured covariates, the marginal value of effort becomes

$$MV_{\mathcal{J},\mathcal{K}'} = \mathbb{E}\left[\frac{\partial}{\partial Y^0}\mathbb{E}\big[\eta_{\mathcal{J}'} \mid Y^0, a_{\mathcal{J}}, c_{\mathcal{K}'}\big] \,\bigg|\, a_{\mathcal{J}}, c_{\mathcal{K}'}\right].$$

Because $g^{k'}$ is a one-to-one-mapping, conditioning on $c_{k'}$ is equivalent to conditioning on $\varepsilon_{k'} = g^{k'}(c_{k'})$. We may therefore equivalently write the agent's marginal value of effort under the expanded set of covariates as

$$MV_{\mathcal{J},\mathcal{K}'} = \mathbb{E}\left[\frac{\partial}{\partial Y^0}\mathbb{E}\big[\eta_{\mathcal{J}'} \mid Y^0, \varepsilon_{k'}, a_{\mathcal{J}}, c_{\mathcal{K}}\big] \,\bigg|\, \varepsilon_{k'}, a_{\mathcal{J}}, c_{\mathcal{K}}\right].$$

LEMMA D.3. $(\varepsilon_{k'}, \delta_{\mathcal{K}}, Y^0)$ *are affiliated conditional on* $(a_{\mathcal{J}}, c_{\mathcal{K}})$.

Proof. This is established along very similar lines to the proof of Lemma D.2. Fix realizations of $(a_{\mathcal{J}}, c_{\mathcal{K}})$, condition all distributions on their values, and suppress explicit conditioning. Let $\rho_\eta(u)$ be the density of $\eta_{\mathcal{J}}$ and $\rho_{\delta'|\varepsilon}(x \mid z)$ be the conditional density of $\delta_{-\mathcal{K}'} \mid \varepsilon_{k'}$. The conditions required for the steps of the proof of Lemma D.2 to go through are that $\rho_\eta(u)$ is log-concave, $\rho_{\delta'|\varepsilon}(x \mid z)$ is log-concave in x for all z, and $(\delta_{\mathcal{K}'}, \varepsilon_{k'})$ are affiliated. The first two properties follow from Assumption B.1, while the final property holds by $\mathcal{K}$-affiliation of circumstance k'. $\square$

We compare $MV_{\mathcal{J},\mathcal{K}'}$ with $MV_{\mathcal{J},\mathcal{K}}$ in a manner very similar to the case of an additional attribute. Fix realizations of $(a_{\mathcal{J}}, c_{\mathcal{K}})$, and define

$$F_\varepsilon(z \mid y) \equiv \Pr\big(\varepsilon_{k'} \leq z \mid Y^0 = y, a_{\mathcal{J}}, c_{\mathcal{K}}\big)$$

to be the conditional CDF of $\varepsilon_{k'}$ given the outcome Y^0. Decompose Y^0 as

$$Y^0 = \mu_{(\mathcal{J},\mathcal{K})} + \eta_{\mathcal{J}} + \delta_{\mathcal{K}}.$$

Taking expectations of each side conditional on $(Y^0, \varepsilon_{k'}, a_{\mathcal{J}}, c_{\mathcal{K}})$ yields

$$Y^0 = \mu_{(\mathcal{J},\mathcal{K})} + \mathbb{E}\big[\eta_{\mathcal{J}} \mid Y^0, \varepsilon_{k'}, a_{\mathcal{J}}, c_{\mathcal{K}}\big] + \mathbb{E}\big[\delta_{\mathcal{K}} \mid Y^0, \varepsilon_{k'}, a_{\mathcal{J}}, c_{\mathcal{K}}\big].$$

Hence,

$$\int dF_\varepsilon(z \mid y)\mathbb{E}\big[\eta_{\mathcal{J}} \mid Y^0 = y, \varepsilon_{k'} = z, a_{\mathcal{J}}, c_{\mathcal{K}}\big]$$

$$= y - \mu_{(\mathcal{J},\mathcal{K})} - \int dF_\varepsilon(z \mid y)\mathbb{E}\big[\delta_{\mathcal{K}} \mid Y^0 = y, \varepsilon_{k'} = z, a_{\mathcal{J}}, c_{\mathcal{K}}\big]. \qquad (\text{D.2})$$

Lemma D.3 directly implies that

$$\int dF_\varepsilon(z \mid y)\mathbb{E}\big[\delta_{\mathcal{K}} \mid Y^0 = y, \varepsilon_{k'} = z, a_{\mathcal{J}}, c_{\mathcal{K}}\big]$$

is weakly increasing in y, so (D.2) is weakly decreasing in y.

Following the same logic as in the attributes case, monotonicity of (D.2) implies that

$$\frac{\partial}{\partial Y^0}\mathbb{E}\big[\eta_{\mathcal{J}} \mid Y^0, a_{\mathcal{J}}, c_{\mathcal{K}}\big] \leq \mathbb{E}\left[\frac{\partial}{\partial Y^0}\mathbb{E}\big[\eta_{\mathcal{J}} \mid Y^0, \varepsilon_{k'}, a_{\mathcal{J}}, c_{\mathcal{K}}\big] \,\Big|\, Y^0, a_{\mathcal{J}}, c_{\mathcal{K}}\right],$$

and it follows that

$$MV_{\mathcal{J},\mathcal{K}} \leq \mathbb{E}[MV_{\mathcal{J},\mathcal{K}'} \mid a_{\mathcal{J}}, c_{\mathcal{K}}].$$

Thus, the marginal value of effort in each subpopulation in the baseline is weakly lower than the expected marginal value of effort when the circumstance k' is additionally measured.

D.3 *Proofs of Theorems 2 and B.2*

We prove Theorem B.2, from which Theorem 2 follows immediately as a corollary.

D.3.1 *Part (a)*   Fix a baseline family of measured covariates $(\mathcal{J}, \mathcal{K})$. As established in Lemma D.1, the marginal value of effort is

$$MV(\mathcal{J}, \mathcal{K}) = \mathbb{E}\left[\frac{\partial}{\partial Y^0} \mathbb{E}\big[\eta_{\mathcal{J}} \mid Y^0, a_{\mathcal{J}}, c_{\mathcal{K}}\big] \,\Big|\, a_{\mathcal{J}}, c_{\mathcal{K}} \right],$$

where $Y^0 \equiv \theta + \varepsilon$ is the baseline value of the outcome after subtracting out the agent's effort.

Now suppose the market additionally observes the attribute $j' \notin \mathcal{J}$, and let $\mathcal{J}' \equiv \mathcal{J} \cup \{j'\}$. Under the expanded family of measured covariates, the marginal value of effort becomes

$$MV_{\mathcal{J}', \mathcal{K}} = \mathbb{E}\left[\frac{\partial}{\partial Y^0} \mathbb{E}\big[\eta_{\mathcal{J}'} \mid Y^0, a_{\mathcal{J}'}, c_{\mathcal{K}}\big] \,\Big|\, a_{\mathcal{J}'}, c_{\mathcal{K}} \right],$$

where, conditional on $(a_{\mathcal{J}}, c_{\mathcal{K}})$, $MV_{\mathcal{J}', \mathcal{K}}$ is a random variable whose value is a function of the realization of $a_{j'}$.

The outcome Y^0 may be decomposed as

$$Y^0 = \mu_{(\mathcal{J}, \mathcal{K})} + \eta_{\mathcal{J}} + \delta_{\mathcal{K}}, \tag{D.3}$$

where

$$\mu_{(\mathcal{J}, \mathcal{K})} \equiv \sum_{j \in \mathcal{J}} \theta_j + \sum_{k \in \mathcal{K}} \varepsilon_k$$

is constant conditional on $(a_{\mathcal{J}'}, c_{\mathcal{K}})$. The residual type component $\eta_{\mathcal{J}}$ may be further decomposed as

$$\eta_{\mathcal{J}} = \bar{\eta}_{\mathcal{J}|j'} + \Delta\eta_{\mathcal{J}'},$$

where

$$\bar{\eta}_{\mathcal{J}|j'} \equiv \mathbb{E}[\eta_{\mathcal{J}} \mid a_{\mathcal{J}'}], \quad \Delta\eta_{\mathcal{J}'} \equiv \theta - \mathbb{E}[\theta \mid a_{\mathcal{J}'}].$$

We may therefore rewrite (D.3) as

$$Y^0 = \mu_{(\mathcal{J}, \mathcal{K})} + \bar{\eta}_{\mathcal{J}|j'} + \Delta\eta_{\mathcal{J}'} + \delta_{\mathcal{K}}.$$

Now, note that

$$\eta_{\mathcal{J}'} - \mathbb{E}[\eta_{\mathcal{J}'} \mid a_{\mathcal{J}'}] = \sum_{j \in \mathcal{J}'} \theta_{j'} + \eta_{\mathcal{J}'} - \mathbb{E}\left[\sum_{j \in \mathcal{J}'} \theta_{j'} + \eta_{\mathcal{J}'} \mid a_{\mathcal{J}'} \right] = \Delta\eta_{\mathcal{J}'}.$$

Hence, $\mathcal{J}$-strong homoskedasticity of attribute j' is equivalent to the assumption that $\Delta\eta_{\mathcal{J}'}$ is independent of $a_{j'}$ conditional on $(a_{\mathcal{J}}, c_{\mathcal{K}})$. Therefore, under $\mathcal{J}$-strong homoskedasticity, Y^0 depends on $a_{j'}$ only through $\bar{\eta}_{\mathcal{J}|j'}$. It follows that under $\mathcal{J}$-strong homoskedasticity, $\mathbb{E}[\Delta\eta_{\mathcal{J}'} \mid Y^0, a_{\mathcal{J}'}, c_{\mathcal{K}}] = \mathbb{E}[\Delta\eta_{\mathcal{J}'} \mid Y^0, \bar{\eta}_{\mathcal{J}|j'}, a_{\mathcal{J}}, c_{\mathcal{K}}]$, and the latter expectation depends on $a_{j'}$ only through $\bar{\eta}_{\mathcal{J}|j'}$.

Using this fact, we may write

$$\mathbb{E}\big[\eta_{\mathcal{J}'} \mid Y^0, a_{\mathcal{J}'}, c_{\mathcal{K}}\big] = \bar{\eta}_{\mathcal{J}|j'} + \mathbb{E}\big[\Delta\eta_{\mathcal{J}'} \mid Y^0, \bar{\eta}_{\mathcal{J}|j'}, a_{\mathcal{J}}, c_{\mathcal{K}}\big]$$

and

$$MV_{\mathcal{J}',\mathcal{K}} = \mathbb{E}\left[\frac{\partial}{\partial Y^0} \mathbb{E}\left[\Delta\eta_{\mathcal{J}'} \mid Y^0, \bar{\eta}_{\mathcal{J}|j'}, a_{\mathcal{J}}, c_{\mathcal{K}} \right] \,\middle|\, \bar{\eta}_{\mathcal{J}|j'}, a_{\mathcal{J}}, c_{\mathcal{K}} \right].$$

The theorem holds if can we show that the conditional expectation of $\Delta\eta_{\mathcal{J}'}$ is less responsive to the realization of the outcome Y than the conditional expectation of the original residual $\eta_{\mathcal{J}}$. Note that $\eta_{\mathcal{J}}$ is the sum of the (conditionally) independent variables $\bar{\eta}_{\mathcal{J}|j'}$ and $\Delta\eta_{\mathcal{J}'}$, so uncertainty about $\Delta\eta_{\mathcal{J}'}$ is mechanically lower than uncertainty about $\eta_{\mathcal{J}}$. But this does not directly translate into a statement that the posterior expectation of $\Delta\eta_{\mathcal{J}'}$ is less sensitive to the realization of Y. In general, we are not even guaranteed that higher realizations of Y lead to higher inferences about $\Delta\theta_{\mathcal{J}}$ once we have conditioned on the realization of $\bar{\eta}_{\mathcal{J}|j'}$.[29] We next prove a key technical lemma, which will imply an analogue of admissibility for our transformed environment.

LEMMA D.4. $(\eta_{\mathcal{J}}, \bar{\eta}_{\mathcal{J}|j'}, Y^0)$ *are affiliated conditional on* $(a_{\mathcal{J}}, c_{\mathcal{K}})$.

PROOF. Fix a set of realizations of $(a_{\mathcal{J}}, c_{\mathcal{K}})$, and condition all distributions on these values. To economize on notation, we suppress explicit conditioning on these co-variates throughout this proof. Let $\tilde{\rho}_{\eta,\theta,Y}(u, t, y)$ be the conditional joint density of $(\eta_{\mathcal{J}}, \bar{\eta}_{\mathcal{J}|j'}, Y^0)$. We will show that $\tilde{\rho}_{\eta,\theta,Y}$ is log-supermodular.

Use $\tilde{\rho}_{\theta}(t)$ to denote the density of $\bar{\eta}_{\mathcal{J}|j'}$, $\tilde{\rho}_{\eta|\theta}(u \mid t)$ to denote the conditional density of $\eta_{\mathcal{J}} \mid \bar{\eta}_{\mathcal{J}|j'}$, and $\tilde{\rho}_{Y|\eta}(y \mid u)$ to denote the conditional density of $Y^0 \mid \eta_{\mathcal{J}}$. Note that Y^0 is independent of $\bar{\eta}_{\mathcal{J}|j'}$ conditional on $\eta_{\mathcal{J}}$. So, $\tilde{\rho}_{\eta,\theta,Y}$ may be decomposed as

$$\tilde{\rho}_{\eta,\theta,Y}(u, t, y) = \tilde{\rho}_{\theta}(t)\tilde{\rho}_{\eta|\theta}(u \mid t)\tilde{\rho}_{Y|\eta}(y \mid u).$$

It is therefore sufficient to show that $\tilde{\rho}_{Y|\eta}$ and $\tilde{\rho}_{\eta|\theta}$ are log-supermodular.

First, consider $\tilde{\rho}_{Y|\eta}$. Decompose Y^0 as

$$Y^0 = \mu_{(\mathcal{J},\mathcal{K})} + \eta_{\mathcal{J}} + \delta_{\mathcal{K}}.$$

Let $\rho_{\delta}(z)$ be the density of $\delta_{\mathcal{K}}$. Then

$$\tilde{\rho}_{Y|\eta}(y \mid u) = \rho_{\delta}(y - \mu_{(\mathcal{J},\mathcal{K})} - u).$$

Under Assumption B.1, ρ_{δ} is log-concave, meaning $\tilde{\rho}_{Y|\eta}$ is log-supermodular.

As for $\tilde{\rho}_{\eta|\theta}$, let $\tilde{\rho}_{\eta'}(w)$ be the density of $\Delta\eta_{\mathcal{J}'}$. Decompose $\eta_{\mathcal{J}}$ as

$$\eta_{\mathcal{J}} = \bar{\eta}_{\mathcal{J}|j'} + \Delta\eta_{\mathcal{J}'},$$

and recall that if j' is $\mathcal{J}$-strongly homoskedastic, then $\Delta\eta_{\mathcal{J}'}$ is independent of $a_{j'}$ and hence $\bar{\eta}_{\mathcal{J}|j'}$. It follows that

$$\tilde{\rho}_{\eta|\theta}(u \mid t) = \tilde{\rho}_{\eta'}(u - t),$$

---

[29]Recall that our admissibility assumptions are imposed on the original type component $\theta_{j'}$, and not on the constructed $\bar{\eta}_{\mathcal{J}|j'}$.

and hence,

$$\frac{\partial^2}{\partial u \partial t} \log \tilde{\rho}_{\eta|\theta}(u \mid t) = -\frac{\partial^2}{\partial w^2} \log \tilde{\rho}_{\eta'}(w)\bigg|_{w=u-t} = -\frac{\partial^2}{\partial u^2} \log \tilde{\rho}_{\eta|\theta}(u \mid t).$$

Now, let $\rho_{\eta|a}(u \mid \alpha)$ denote the conditional density of $\eta_{\mathcal{J}} \mid a_{j'}$. Define

$$\zeta(\alpha) \equiv f^{j'}(\alpha) + \mathbb{E}[\eta_{\mathcal{J}'} \mid a_{j'} = \alpha],$$

so that

$$\eta_{\mathcal{J}} = \zeta(a_{j'}) + \Delta \eta_{\mathcal{J}'}.$$

Strong homoskedasticity implies that

$$\rho_{\eta|a}(u \mid \alpha) = \tilde{\rho}_{\eta'}\big(u - \zeta(\alpha)\big) = \tilde{\rho}_{\eta|\theta}\big(u \mid \zeta(\alpha)\big).$$

Let $\tilde{\Theta} \equiv \{t : \zeta(\alpha) = t \text{ for some } \alpha \in A_{j'}\}$ denote the support of $\bar{\eta}_{\mathcal{J}|j'}$. Fix any $t \in \tilde{\Theta}$. Then for all u and every $\alpha \in A_{j'}$ such that $\zeta(\alpha) = t$,

$$\frac{\partial^2}{\partial u^2} \log \tilde{\rho}_{\eta|\theta}(u \mid t) = \frac{\partial^2}{\partial u^2} \log \rho_{\eta|a}(u \mid \alpha).$$

Let $\rho_{\eta'|a}$ denote the conditional density of $\eta_{\mathcal{J}'} \mid a_{j'}$. Then $\rho_{\eta|a}(u \mid \alpha) = \rho_{\eta'|a}(u - f^{j'}(\alpha) \mid \alpha)$, so that

$$\frac{\partial^2}{\partial u^2} \log \tilde{\rho}_{\eta|\theta}(u \mid t) = \frac{\partial^2}{\partial w^2} \log \rho_{\eta'|a}(w \mid \alpha)\bigg|_{w=u-f^{j'}(\alpha)}.$$

Under Assumption B.1, $\rho_{\eta'|a}$ is log-concave and this final derivative is nonpositive, meaning

$$\frac{\partial^2}{\partial u \partial t} \log \tilde{\rho}_{\eta|\theta}(u \mid t) = -\frac{\partial^2}{\partial u^2} \log \tilde{\rho}_{\eta|\theta}(u \mid t) \geq 0$$

for every u and $t \in \tilde{\Theta}$. Hence, $\tilde{\rho}_{\eta|\theta}$ is log-supermodular, as desired.   □

Following arguments identical to those used for the proof of Theorem B.1 (with $\Delta \eta_{\mathcal{J}'}$ and $\bar{\eta}_{\mathcal{J}|j'}$ playing the roles of $\eta_{\mathcal{J}'}$ and $\theta_{j'}$, respectively), Lemma D.4 implies

$$MV_{\mathcal{J},\mathcal{K}} \geq \mathbb{E}[MV_{\mathcal{J}',\mathcal{K}} \mid a_{\mathcal{J}}, c_{\mathcal{K}}].$$

Thus, the marginal value of effort in each subpopulation in the baseline setting is weakly higher than the expected marginal value once attribute j' is additionally measured, for any realizations of $(a_{\mathcal{J}}, c_{\mathcal{K}})$.

To complete the proof, we must establish that monotonicity holds uniformly across realizations of $a_{j'}$, and not just on average. This follows immediately from the fact that $MV_{\mathcal{J}',\mathcal{K}}$ is independent of $a_{j'}$ conditional on $(a_{\mathcal{J}}, c_{\mathcal{K}})$. To see this, decompose Y as

$$Y = e + \mu_{(\mathcal{J},\mathcal{K})} + \bar{\eta}_{\mathcal{J}|j'} + \Delta \eta_{\mathcal{J}'} + \delta_{\mathcal{K}}.$$

Strong homoskedasticity implies that $\Delta\eta_{\mathcal{J}'}$ is independent of $a_{j'}$ conditional on $(a_{\mathcal{J}}, c_{\mathcal{K}})$. Hence, $a_{j'}$ enters the market's inference problem only via a known additive shift $\bar{\eta}_{\mathcal{J}|j'}$ to the agent's type distribution and, its value therefore does not impact incentives for effort. So, these incentives must be independent of $a_{j'}$, as claimed.

D.3.2 *Part (b)*   Suppose the market observes the additional circumstance $k' \notin \mathcal{K}$. Let $\mathcal{K}' \equiv \mathcal{K} \cup \{k'\}$. Under the expanded family of measured covariates, the marginal value of effort is

$$MV_{\mathcal{J},\mathcal{K}'} = \mathbb{E}\left[\frac{\partial}{\partial Y^0}\mathbb{E}\left[\eta_{\mathcal{J}} \mid Y^0, a_{\mathcal{J}}, c_{\mathcal{K}'}\right] \,\middle|\, a_{\mathcal{J}}, c_{\mathcal{K}'}\right].$$

Define

$$\bar{\delta}_{\mathcal{K},k'} \equiv \mathbb{E}[\delta_{\mathcal{K}} \mid c_{\mathcal{K}'}], \quad \Delta\delta_{\mathcal{K}'} \equiv \varepsilon - \mathbb{E}[\varepsilon \mid c_{\mathcal{K}'}].$$

Then Y^0 may be decomposed as

$$Y^0 = \mu_{(\mathcal{J},\mathcal{K})} + \eta_{\mathcal{J}} + \bar{\delta}_{\mathcal{K},k'} + \Delta\delta_{\mathcal{K}'}.$$

Under strong homoskedasticity, $\Delta\delta_{\mathcal{K}'}$ is independent of $c_{k'}$ conditional on $(a_{\mathcal{J}}, c_{\mathcal{K}})$, and so the distribution of Y^0 depends on $c_{k'}$ only through $\bar{\delta}_{\mathcal{K},k'}$. Thus,

$$\mathbb{E}\left[\eta_{\mathcal{J}} \mid Y^0, a_{\mathcal{J}}, c_{\mathcal{K}'}\right] = \mathbb{E}\left[\eta_{\mathcal{J}} \mid Y^0, \bar{\delta}_{\mathcal{K},k'}, a_{\mathcal{J}}, c_{\mathcal{K}}\right],$$

and the latter random variable depends on $c_{k'}$ only through $\bar{\delta}_{\mathcal{K},k'}$. Therefore, in a manner analogous to the attribute case, the marginal value of effort after measuring j' depends on $c_{k'}$ only through $\bar{\delta}_{\mathcal{K},k'}$ and may be written

$$MV_{\mathcal{J},\mathcal{K}'} = \mathbb{E}\left[\frac{\partial}{\partial Y^0}\mathbb{E}\left[\eta_{\mathcal{J}} \mid Y^0, \bar{\delta}_{\mathcal{K},k'}, a_{\mathcal{J}}, c_{\mathcal{K}}\right] \,\middle|\, \bar{\delta}_{\mathcal{K},k'}, a_{\mathcal{J}}, c_{\mathcal{K}}\right].$$

LEMMA D.5. $(\bar{\delta}_{\mathcal{K},k'}, \delta_{\mathcal{K}}, Y^0)$ *are affiliated conditional on* $(a_{\mathcal{J}}, c_{\mathcal{K}})$.

PROOF. This proof follows along very similar lines to the proof of Lemma D.4. Fix realizations of $(a_{\mathcal{J}}, c_{\mathcal{K}})$, condition all distributions on their values, and suppress explicit conditioning on these covariates. Let $\rho_\eta(u)$ be the conditional density of $\eta_{\mathcal{J}}$ and $\rho_{\delta'|c}(x \mid z)$ be the conditional density of $\delta_{\mathcal{K}'} \mid c_{k'}$. The conditions required for the steps of the proof of Lemma D.4 to go through are that $\rho_\eta(u)$ is log-concave, $\rho_{\delta'|c}(x \mid z)$ is log-concave in x for all z, and $\Delta\delta_{\mathcal{K}'}$ is independent of $\bar{\delta}_{\mathcal{K}|k'}$. The first two properties are ensured by Assumption B.1, while the final property holds under strong homoskedasticity. □

We compare $MV_{\mathcal{J},\mathcal{K}'}$ and $MV_{\mathcal{J},\mathcal{K}}$ in a manner very similar to the attribute case. Fix realizations of $(a_{\mathcal{J}}, c_{\mathcal{K}})$. Define

$$\widetilde{F}_\varepsilon(z \mid y) \equiv \Pr\left(\bar{\delta}_{\mathcal{K},k'} \leq z \mid Y^0 = y, a_{\mathcal{J}}, c_{\mathcal{K}}\right)$$

to be the distribution function of $\bar{\delta}_{\mathcal{K},k'}$ conditional on the outcome Y^0. Decompose Y^0 as

$$Y^0 = \mu_{(\mathcal{J},\mathcal{K})} + \eta_{\mathcal{J}} + \delta_{\mathcal{K}}.$$

Taking expectations of each side conditional on $(Y^0, \bar{\delta}_{\mathcal{K},k'}, a_{\mathcal{J}}, c_{\mathcal{K}})$ yield

$$Y^0 = \mu_{(\mathcal{J},\mathcal{K})} + \mathbb{E}[\eta_{\mathcal{J}} \mid Y^0, \bar{\delta}_{\mathcal{K},k'}, a_{\mathcal{J}}, c_{\mathcal{K}}] + \mathbb{E}[\delta_{\mathcal{K}} \mid Y^0, \bar{\delta}_{\mathcal{K},k'}, a_{\mathcal{J}}, c_{\mathcal{K}}].$$

Hence,

$$\int d\widetilde{F}_{\varepsilon}(z \mid y') \mathbb{E}[\eta_{\mathcal{J}} \mid Y^0 = y, \bar{\delta}_{\mathcal{K},k'} = z, a_{\mathcal{J}}, c_{\mathcal{K}}]$$

$$= y - \mu_{(\mathcal{J},\mathcal{K})} - \int d\widetilde{F}_{\varepsilon}(z \mid y') \mathbb{E}[\delta_{\mathcal{K}} \mid Y^0 = y, \bar{\delta}_{\mathcal{K},k'} = z, a_{\mathcal{J}}, c_{\mathcal{K}}]. \qquad (\text{D.4})$$

Lemma D.5 directly implies that

$$\int d\widetilde{F}_{\varepsilon}(z \mid y) \mathbb{E}[\delta_{\mathcal{K}} \mid Y^0 = y, \bar{\delta}_{\mathcal{K},k'} = z, a_{\mathcal{J}}, c_{\mathcal{K}}]$$

is nondecreasing in y, so (D.4) is nonincreasing in y.

Following the same logic as in the attributes case, monotonicity of (D.4) implies that

$$\frac{\partial}{\partial Y^0} \mathbb{E}[\eta_{\mathcal{J}} \mid Y^0, a_{\mathcal{J}}, c_{\mathcal{K}}] \leq \mathbb{E}\left[\frac{\partial}{\partial Y^0} \mathbb{E}[\eta_{\mathcal{J}} \mid Y^0, \bar{\delta}_{\mathcal{K},k'}, a_{\mathcal{J}}, c_{\mathcal{K}}] \,\middle|\, Y^0, a_{\mathcal{J}}, c_{\mathcal{K}}\right],$$

and it follows that

$$MV_{\mathcal{J},\mathcal{K}} \leq \mathbb{E}[MV_{\mathcal{J},\mathcal{K}'} \mid a_{\mathcal{J}}, c_{\mathcal{K}}].$$

Thus, the marginal value of effort in each subpopulation in the baseline is weakly lower than the expected marginal value of effort when the circumstance k' is additionally measured.

The final step in the proof is to establish that monotonicity holds uniformly across realizations of the additional circumstance, and not just on average. This follows from arguments nearly identical to those used for the attributes case.

D.4 *Proof of Theorem 3*

Define $\delta \equiv \beta/(1 - \beta)$. Then given any set of measured covariates, equilibrium effort can be written as $e^*_{\mathcal{J},\mathcal{K}} = \delta \cdot MV_{\mathcal{J},\mathcal{K}}$. We will characterize the aggregate welfare change from measuring a new covariate as a function of δ, which can be mapped onto an equivalent characterization in terms of β.

Consider first the case in which the market observes the new attribute j'. Let $\mathcal{J}' \equiv \mathcal{J} \cup \{j'\}$. The aggregate change in welfare from measuring j', as a function of δ, is

$$\Delta\overline{W}(\delta) = \mathbb{E}[w_0(\delta \cdot MV_{\mathcal{J}',\mathcal{K}}) - w_0(\delta \cdot MV_{\mathcal{J},\mathcal{K}})],$$

where

$$w_0(e) \equiv e - \frac{1}{2}e^2.$$

The aggregate change in welfare is a quadratic function of δ, whose first derivative is

$$\frac{d}{d\delta}\Delta\overline{W}(\delta) = \mathbb{E}[MV_{\mathcal{J}',\mathcal{K}} - MV_{\mathcal{J},\mathcal{K}}] - \delta \cdot \mathbb{E}[MV^2_{\mathcal{J}',\mathcal{K}} - MV^2_{\mathcal{J},\mathcal{K}}].$$

Observe that $\overline{W}(0) = 0$ and

$$\frac{d}{d\delta}\Delta\overline{W}(\delta)|_{\delta=0} = \mathbb{E}[MV_{\mathcal{J}',\mathcal{K}} - MV_{\mathcal{J},\mathcal{K}}\Big] \leq 0$$

with the inequality implied by regularity of j'. So, $\Delta\overline{W}(\delta)$ is a quadratic function that vanishes at $\delta = 0$ and is nonincreasing there. Suppose first that j' is strictly regular, so that $\Delta\overline{W}$ is strictly decreasing at 0. Then its shape must then satisfy either of:

- $\Delta\overline{W}(\delta)$ is strictly convex and intersects zero exactly once for $\delta > 0$,

- $\Delta\overline{W}(\delta)$ is weakly concave and does not intersect zero for any $\delta > 0$.

Define $\delta^* \equiv \inf\{\delta > 0 : \Delta\overline{W}(\delta) \geq 0\}$. This threshold lies in $(0, \infty)$, and has the property that $\Delta\overline{W}(\delta) > 0$ for $\delta > \delta^*$, and $\Delta\overline{W}(\delta) \leq 0$ for $\delta \in (0, \delta^*]$. Further, $\delta^* < \infty$ if and only if $\Delta\overline{W}$ is strictly convex. Convexity is determined by the sign of

$$\frac{d^2}{d\delta^2}\Delta\overline{W}(\delta) = \mathbb{E}[MV^2_{\mathcal{J},\mathcal{K}} - MV^2_{\mathcal{J}',\mathcal{K}}].$$

Hence, $\delta^* < \infty$ if and only if $\mathbb{E}[MV^2_{\mathcal{J},\mathcal{K}}] > \mathbb{E}[MV^2_{\mathcal{J}',\mathcal{K}}]$. Letting $\beta^* \equiv \delta^*/(1 + \delta^*)$ therefore yields a reputation weight threshold with the desired properties.

Suppose instead that j' is weakly but not strictly regular. Write

$$\mathbb{E}[MV^2_{\mathcal{J}',\mathcal{K}} \mid a_{\mathcal{J}}, c_{\mathcal{K}}]$$

$$= \mathbb{E}[(MV_{\mathcal{J}',\mathcal{K}} - \mathbb{E}[MV_{\mathcal{J}',\mathcal{K}} \mid a_{\mathcal{J}}, c_{\mathcal{K}}])^2 \mid a_{\mathcal{J}}, c_{\mathcal{K}}] + \mathbb{E}[MV_{\mathcal{J}',\mathcal{K}} \mid a_{\mathcal{J}}, c_{\mathcal{K}}]^2$$

$$= \mathrm{Var}(MV_{\mathcal{J}',\mathcal{K}} \mid a_{\mathcal{J}}, c_{\mathcal{K}}) + \mathbb{E}[MV_{\mathcal{J}',\mathcal{K}} \mid a_{\mathcal{J}}, c_{\mathcal{K}}]^2.$$

Hence,

$$\mathbb{E}[MV^2_{\mathcal{J}',\mathcal{K}} \mid a_{\mathcal{J}}, c_{\mathcal{K}}] - MV^2_{\mathcal{J},\mathcal{K}}$$

$$= \mathbb{E}[MV_{\mathcal{J}',\mathcal{K}} \mid a_{\mathcal{J}}, c_{\mathcal{K}}]^2 - MV^2_{\mathcal{J},\mathcal{K}} + \mathrm{Var}(MV_{\mathcal{J}',\mathcal{K}} \mid a_{\mathcal{J}}, c_{\mathcal{K}}).$$

Weak but not strict regularity implies that $MV_{\mathcal{J},\mathcal{K}} = \mathbb{E}[MV_{\mathcal{J}',\mathcal{K}} \mid a_{\mathcal{J}}, c_{\mathcal{K}}]$ with probability 1, so that

$$\mathbb{E}[MV^2_{\mathcal{J}',\mathcal{K}} \mid a_{\mathcal{J}}, c_{\mathcal{K}}] - MV^2_{\mathcal{J},\mathcal{K}} = \mathrm{Var}(MV_{\mathcal{J}',\mathcal{K}} \mid a_{\mathcal{J}}, c_{\mathcal{K}}) \geq 0$$

with probability 1. Therefore, $\Delta\overline{W}$ is at least weakly concave. Weak but not strict regularity additionally implies that $\Delta\overline{W}$ has zero slope at $\delta = 0$. Hence, in this case

$\Delta \overline{W}(\delta) \leq 0$ for all $\delta > 0$, and we set $\delta^* = \infty$ and $\beta^* = \infty$. We may summarize the work for the strictly and weakly regular cases with the conclusion that $\beta^* < \infty$ if and only if $\mathbb{E}[MV_{\mathcal{J},\mathcal{K}}^2] > \mathbb{E}[MV_{\mathcal{J}',\mathcal{K}}^2]$. Indeed, in the strictly regular case, this equivalence was directly established, while in the weakly regular case, it is true both that $\beta^* = \infty$ and that the latter inequality cannot hold.

Now suppose the market observes the new circumstance k'. Calculations very similar to those for the attribute case show that $\Delta \overline{W}$ is a quadratic function of δ which vanishes at $\delta = 0$ and is weakly increasing there, and strictly increasing if k' is strictly regular. It is concave if and only if $\mathbb{E}[MV_{\mathcal{J},\mathcal{K}'}^2] \geq [MV_{\mathcal{J},\mathcal{K}}^2]$, with the concavity strict if and only if the inequality is.

If k' is only weakly regular, then the calculations for the attribute case show that $\mathbb{E}[MV_{\mathcal{J},\mathcal{K}'}^2] \geq \mathbb{E}[MV_{\mathcal{J},\mathcal{K}}^2]$, meaning that $\Delta \overline{W}$ must be at least weakly concave, in which case $\Delta \overline{W}(\delta) \leq 0$ for all $\delta > 0$. In this case, we set $\delta_* = 0$ and $\beta_* = \delta_*/(1+\delta_*) = 0$.

Going forward, suppose that k' is strictly regular. Defining $\delta_* \equiv \inf\{\delta > 0 : \Delta \overline{W}(\delta) \leq 0\}$ yields a threshold in $(0, \infty]$ with the property that $\Delta \overline{W}(\delta) < 0$ for all $\delta > \delta_*$ and $\Delta \overline{W}(\delta) > 0$ for all $\delta \in (0, \delta_*)$. This threshold is finite if and only if $\Delta \overline{W}$ is a strictly concave function. Calculations very similar to the attribute case imply that

$$\mathbb{E}\big[MV_{\mathcal{J},\mathcal{K}'}^2 \mid a_{\mathcal{J}}, c_{\mathcal{K}}\big] - MV_{\mathcal{J},\mathcal{K}}^2$$
$$= \mathbb{E}[MV_{\mathcal{J},\mathcal{K}'} \mid a_{\mathcal{J}}, c_{\mathcal{K}}]^2 - MV_{\mathcal{J},\mathcal{K}}^2 + \mathrm{Var}(MV_{\mathcal{J},\mathcal{K}'} \mid a_{\mathcal{J}}, c_{\mathcal{K}}).$$

Strict regularity implies that the difference between the first two terms on the rhs is nonnegative for every realization of $(a_{\mathcal{J}}, c_{\mathcal{K}})$, and positive with positive probability. Further, the conditional variance of $MV_{\mathcal{J},\mathcal{K}'}$ must be nonnegative. Therefore,

$$\mathbb{E}\big[MV_{\mathcal{J},\mathcal{K}'}^2 \mid a_{\mathcal{J}}, c_{\mathcal{K}}\big] \geq MV_{\mathcal{J},\mathcal{K}}^2,$$

with the inequality strict with positive probability. Taking the unconditional expectation of both sides yields the desired strict concavity of $\Delta \overline{W}$, implying $\delta_* < \infty$. Letting $\beta_* \equiv \delta_*/(1+\delta_*)$ yields a reputation weight with the desired properties.

D.5 *Proof of Proposition 1*

We prove the first part of the result, with the second following from nearly identical arguments. To streamline notation, throughout this proof we will drop superscripts on Y^0 and $\bar{Y}^0$.

Under Assumption 3, conditioning on the value of X is equivalent to conditioning on the value of $\bar{Y}$. The desired result can therefore be established using a technique very similar to the one employed in the proof of Theorem B.1(a) subsequent to Lemma D.2, with $\bar{Y}$ playing the role of $\theta_{j'}$. That argument requires two conditions: (A) $(\bar{Y}, Y)$ are statistically affiliated, and (B) $(\bar{Y}, \theta)$ are statistically affiliated conditional on Y. We maintain condition A by hypothesis, so it remains only to establish condition B.

Let $\rho_{Y|\bar{Y},\theta}(y \mid y', t)$ be the conditional density of Y given $(\bar{Y}, \theta)$, $\rho_{\bar{Y},\theta}$ be the joint density of $(\bar{Y}, \theta)$, and ρ_Y be the marginal density of Y. Then by Bayes' rule,

$$\rho_{\bar{Y},\theta|Y}(y', t \mid y) = \frac{\rho_{Y|\bar{Y},\theta}(y \mid y', t)\rho_{\bar{Y},\theta}(y', t)}{\rho_Y(y)}.$$

Because $(\bar{Y}, \theta)$ are affiliated, the density $\rho_{\bar{Y},\theta}$ is log-supermodular. To establish condition B, it is therefore sufficient to show that $\rho_{Y|\bar{Y},\theta}(y \mid y', t)$ is log-supermodular in (y', t), holding y fixed.

Let $\rho_{\varepsilon|\bar{Y},\theta}(e \mid y', t)$ and $\rho_{\varepsilon|\bar{Y}}(e \mid y')$ be the conditional densities of ε given $(\bar{Y}, \theta)$ and $\bar{Y}$, respectively. Given the conditional independence of θ and ε given X and the invertibility of $\mathbb{E}(\bar{Y} \mid X)$, it follows that θ and ε are independent conditional on $\bar{Y}$. Therefore,

$$\rho_{\varepsilon|\bar{Y},\theta}(e \mid y', t) = \rho_{\varepsilon|\bar{Y}}(e \mid y')$$

for all t, allowing us to write

$$\rho_{Y|\bar{Y},\theta}(y \mid y', t) = \rho_{\varepsilon|\bar{Y},\theta}(y - t \mid y', t) = \rho_{\varepsilon|\bar{Y}}(y - t \mid y').$$

Let $\rho_{\varepsilon,\bar{Y}}(e, y')$ be the joint density of $(\varepsilon, \bar{Y})$ and $\rho_{\bar{Y}}(y')$ be the marginal density of $\bar{Y}$. Then using Bayes' rule, we have

$$\rho_{Y|\bar{Y},\theta}(y \mid y', t) = \frac{\rho_{\varepsilon,\bar{Y}}(y - t, y')}{\rho_{\bar{Y}}(y')}.$$

Condition B therefore follows if $\rho_{\varepsilon,\bar{Y}}$ is log-submodular. Equivalently, $\rho_{-\varepsilon,\bar{Y}}$ must be log-supermodular, where $\rho_{-\varepsilon,\bar{Y}}$ is the joint density of $(-\varepsilon, \bar{Y})$. This condition is equivalent to affiliation of $(-\varepsilon, \bar{Y})$, as hypothesized.

D.6 *Proof of Lemma 1*

In a multivariate Gaussian environment, the conditional mean of output is

$$\bar{Y}^0 = \mu + (b + d)X,$$

satisfying Assumption 3 whenever $b + d \neq 0$. Since (X, Z, W) are jointly Gaussian and $\bar{Y}^0$ and Y^0 are each linear combinations of (X, Z, W), the pair $(\bar{Y}^0, Y^0)$ are also jointly Gaussian. Additionally,

$$Y^0 = \bar{Y}^0 + Z + W,$$

and since $\bar{Y}^0$ is independent of $Z + W$ it follows that Y^0 and $\bar{Y}^0$ are positively correlated. Therefore, $(\bar{Y}^0, Y^0)$ are affiliated, satisfying Assumption 4. Moreover, conditional expectations in multivariate Gaussian environments are linear in the conditioning variable, ensuring that Assumption 5 is satisfied. Finally, mutual independence of X, Z, and W implies that (θ, ε) are independent conditional on X, satisfying Assumption 6.

D.7 *Proof of Proposition 2*

Prior to the measurement, (θ, Y^0) have joint distribution

$$\begin{pmatrix} \theta \\ Y^0 \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \mu \\ \mu \end{pmatrix}, \begin{pmatrix} \sigma_\theta^2 & \sigma_{\theta,Y} \\ \sigma_{\theta,Y} & \sigma_Y^2 \end{pmatrix} \right)$$

where

$$\sigma_\theta^2 = b^2\sigma_x^2 + \sigma_z^2, \quad \sigma_{\theta,Y} = b(b+d)\sigma_x^2 + \sigma_z^2, \quad \sigma_Y^2 = (b+d)^2\sigma_x^2 + \sigma_z^2 + \sigma_w^2.$$

The market's type forecast given equilibrium effort e^* is therefore

$$\mathbb{E}^{e^*}(\theta \mid Y) = \mathbb{E}(\theta \mid Y^0 = Y - e^*) = \mu + \frac{\sigma_{\theta,Y}}{\sigma_Y^2} \cdot (Y - \mu - e^*),$$

implying that the baseline marginal value of effort is

$$MV = \frac{\sigma_{\theta,Y}}{\sigma_Y^2} = \frac{b(b+d)\sigma_x^2 + \sigma_z^2}{(b+d)^2\sigma_x^2 + \sigma_z^2 + \sigma_w^2}.$$

Meanwhile, after measuring X, the conditional joint distribution of (θ, Y^0) is

$$\begin{pmatrix} \theta \\ Y^0 \end{pmatrix} \bigg| X \sim \mathcal{N}\left(\begin{pmatrix} \mu + bX \\ \mu + (b+d)X \end{pmatrix}, \begin{pmatrix} \sigma_z^2 & \sigma_z^2 \\ \sigma_z^2 & \sigma_z^2 + \sigma_w^2 \end{pmatrix} \right).$$

Given equilibrium effort e^{**}, the market's posterior type forecast after measuring X is

$$\mathbb{E}^{e^{**}}(\theta \mid X, Y) = \mathbb{E}(\theta \mid X, Y^0 = Y - e^{**}) = \mu + bX + \frac{\sigma_z^2}{\sigma_z^2 + \sigma_w^2}(Y - e^{**} - \mu - (b+d)X).$$

The agent's marginal value of effort under the expanded data set is therefore

$$MV_+ = \frac{\sigma_z^2}{\sigma_z^2 + \sigma_w^2}.$$

Comparing the expressions for MV and MV_+ just derived yields the identity in the proposition statement.

References

Ball, Ian (2022), "Scoring strategic agents." Working Paper. [409]

Bergemann, Dirk, Alessandro Bonatti, and Tan Gan (2022), "The economics of social data." *RAND Journal of Economics*, 53, 263–296. [408]

Bonatti, Alessandro and Gonzalo Cisternas (2020), "Consumer scores and price discrimination." *The Review of Economic Studies*, 87, 750–791. [409]

Braverman, Mark and Sylvain Chassang (2022), "Data-driven incentive alignment in capitation schemes." *Journal of Public Economics*, 207, 104584. [408]

Brunnermeier, Markus, Rohit Lamba, and Carlos Segura-Rodriguez (2021), "Inverse selection." Working Paper. https://www.businessinsider.com/china-social-credit-system -punishments-and-rewards-explained-2018-4. [408, 414]

Canales, Katie and Aaron Mok (November 28, 2022), "China's 'social credit' system ranks citizens and punishes them with throttled internet speeds and flight bans if the Communist Party deems them untrustworthy." *Insider*. [407]

CareerBuilder (August 9, 2018), "More Than Half of Employers Have Found Content on Social Media That Caused Them NOT to Hire a Candidate, According to Recent Career-Builder Survey." *PR Newswire*. https://www.prnewswire.com/news-releases/number-of-employers-using-social-media-to-screen-candidates-at-all-time-high-finds-latest-careerbuilder-study-300474228.html. [408]

Cofield, Calla (March 26, 2013), "An Arbitrary Number of Years Since Mathematician Paul Erdős's Birth." *Scientific American*. https://www.scientificamerican.com/article/an-arbitrary-number-of-years-since-mathematicians-birth/. [420]

Dewatripont, Mathias, Ian Jewitt, and Jean Tirole (1999), "The economics of career concerns, Part I: Comparing information structures." *The Review of Economic Studies*, 66, 183–198. [409, 416]

Eliaz, Kfir and Ran Spiegler (2019), "The model selection curse." *American Economic Review: Insights*, 1, 127–140. [409]

Eliaz, Kfir and Ran Spiegler (2022), "On incentive-compatible estimators." *Games and Economic Behavior*, 122, 204–220. [409]

Elliott, Matthew, Andrea Galeotti, and Andrew Koh (2022), "Market segmentation through information." Working Paper. [408]

European Commission (April 21, 2021), "Europe fit for the Digital Age: Commission proposes new rules and actions for excellence and trust in Artificial Intelligence." https://ec.europa.eu/commission/presscorner/detail/en/ip_21_1682 [408]

Frankel, Alex and Navin Kartik (2022), "Improving information from manipulable data." *Journal of the European Economic Association*, 20, 79–115. [409, 426]

Gage, Deborah (May 22, 2012), "Can Social Networks Help You Get A Loan? Lenddo Thinks So." *Wall Street Journal*. https://www.wsj.com/articles/BL-DGB-24461. [407]

Gomes, Renato and Alessandro Pavan (2022), "Price customization and targeting in matching markets." RAND Journal of Economics. (Forthcoming). [408]

Haghtalab, Nika, Nicole Immorlica, Brendan Lucier, and Jack Wang (2020), "Maximizing welfare with incentive-aware evaluation mechanisms." In *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence*, 160–166. [409]

Hidir, Sinem and Nikhil Vellodi (2021), "Privacy, personalization and price discrimination." *Journal of the European Economic Association*, 19, 1342–1363. [408]

Holmström, Bengt (1999), "Managerial incentive problems: A dynamic perspective." *The Review of Economic Studies*, 66, 169–182. [408, 409, 410, 417]

Hörner, Johannes and Nicolas S. Lambert (2021), "Motivational ratings." *The Review of Economic Studies*, 88, 1892–1935. [409]

Hu, Lily, Nicole Immorlica, and Jennifer Wortman Vaughan (2019), "The disparate effects of strategic manipulation." In *Proceedings of the Conference on Fairness, Accountability, and Transparency*, 259–268. [409]

Ichihashi, Shota (2019), "Online privacy and information disclosure by consumers." *American Economic Review*, 110, 569–595. [408]

Kearns, Michael and Aaron Roth (2019), *The Ethical Algorithm: The Science of Socially Aware Algorithm Design*. Oxford University Press. [418]

Lippert, John (October 11, 2024), "ZestFinance issues small, high-rate loans, uses big data to weed out deadbeats." *Washington Post*. https://www.washingtonpost.com/business/zestfinance-issues-small-high-rate-loans-uses-big-data-to-weed-out-deadbeats/2014/10/10/e34986b6-4d71-11e4-aa5e-7153e466a02d_story.html. [407]

Meyer, Margaret A. and John Vickers (1997), "Performance comparisons and dynamic incentives." *Journal of Political Economy*, 105, 547–581. [409]

Milgrom, Paul (1981), "Good news and bad news: Representation theorems and applications." *The Bell Journal of Economics*, 12, 380–391. [413]

Milgrom, Paul and John Roberts (1988), "An economic approach to influence activities in organizations." *American Journal of Sociology*, 94, S154–S179. [426]

Rodina, David (2018), "Information design and career concerns." Working Paper. [409]

Smolin, Alex (2021), "Dynamic evaluation design." *American Economic Journal: Microeconomics*, 13, 300–331. [409]

Tirole, Jean (2021), "Digital dystopia." *American Economic Review*, 111, 2007–2048. [409]

Yang, Kai Hao (2022), "Selling consumer data for profit: Optimal market-segmentation design and its consequences." *American Economic Review*, 112, 1364–1393. [408]