Allocating Scarce Information*

Richard Holden and Anup Malani†

September 11, 2013

Abstract

This paper takes up the question of how a social planner should allocate scarce information. We consider a model with one sender (the planner) and two receivers (agents). The receivers are in different states of the world but the sender may send only one signal about states. We compare the optimal information allocation to two benchmarks: reporting simple statistics (e.g., reporting the average or median state across agents) and the optimal allocation of ordinary goods. We show: (1) reporting a simple statistic is rarely optimal; (2) unlike the allocation of goods, the optimal allocation of information is neither continuous nor monotonic in agents' endowments; (3) information allocations have distributive consequences; and (4) more accurate signals may reduce welfare if there are negative spillovers from agents' actions. Finally, we relate the problem of allocating scarce information to the fundamental welfare theorems and the Coase Theorem.

*Preliminary and Incomplete. Please do not circulate.

†Holden: University of New South Wales and NBER. email: richard.holden@unsw.edu.au. Malani: Univeristy of Chicago and NBER. email: amalani@uchicago.edu. We thank Emir Kamenica for helpful discussions.

1 Introduction

Information is inherently scarce. Because it is a public good, the market may produce a suboptimal level of information and the social planner may want to subsidize the production of information. However, the production of information is costly and thus remains scarce even with optimal government subsidies. This raises the question we take up in this paper: how should the planner allocate scarce information? We offer a theoretical framework in which to analyze this problem and offer some initial guidance for optimal policy. We compare the optimal allocation of information to two benchmarks: the optimal allocation of ordinary goods and an equal sharing rule. The latter captures the common sense solution of providing average or median statistics for a population.

To be more concrete, consider the following hypothetical. Different neighborhoods in a city have different crime rates. Each neighborhood wants to know its local crime rate so they it residents can determine whether it is safe to go out at night and whether they need burglar alarms. The government cannot, however, publish a separate crime rate for each neighborhood. Either it is difficult to precisely measure crime rates for each neighborhood (costly acquisition), it is difficult to obtain TV time or newspaper space to publish neighborhood-level crime rates (costly transmission), or it is too costly for neighborhoods to identify from a neighborhood-level database exactly what their particular crime rate is (costly absorption). Suppose the government can only publish one crime rate for the entire city. Should it publish the rate for one particular neighborhood or should it publish a crime rate that is, say, the average across neighborhoods?[1]

[1] A predicate question – though not the focus of this paper – is why would people believe the information that a government provides? One answer is that they naively believe the government. We find this approach unappealing given the high level of skepticism of government in popular discourse. Therefore, we take individuals as highly rational and focus on a mechanism that still allows the government to persuade people. Specifically, our model assumes that the government can undertake factual inquires that favor some areas over others, but can commit to revealing the full results of these inquiries. For example, it can tell the police to report crime only for the low crime neighborhood, but cannot stop the police form reporting the crime level in that neighborhood. Kamenica & Gentzkow (2011) show that the commitment to revelation of results is sufficient to move the posteriors of purely Bayesian receivers of information. Even with *known*

One intuition that many have is that the government should publish mean or median crime rates. One of our central results is that this intuition is false: the optimal strategy is to allocate information that equates the marginal utility from information. Only under very strict conditions that are highly unlikely to hold will it be the case that the allocation that happens to equate marginal utilities happens to align with a particluar simple statistic like a notable moment or percentile. Indeed, it may even be optimal for the government to target information at one particular area. The reason is that announcing, say, the mean crime rate will cause both the high crime area to employ insufficient precautions and the low crime area to employ excessive precautions. Announcing the crime rate in just the high crime area, however, would cause the high crime area to take appropriate precautions. The low crime area, being familiar with the government's preferences, will presume the government has announced the rate for the high crime area and ignore the announcement. This could minimize overall distortions.

Another intuition that many may have is that, given well-behaved utility functions, the government's allocation of information to a neighborhood may be continuous in that neighborhood's initial endowment. In the social planner's problem with divisible goods and concave utility, an incremental increase in initial endowment lowers the marginal utility of an agent, which in turn incrementally lowers the planner's optimal allocation of goods to that individual. When allocating information, the planner continues to care about marginal utility, but utility may not be continuous in information even if preferences are well-behaved. The reason is that government information affects the receiver's beliefs, which in turn affect her actions, and that in turn affects her utility. Even if the mapping from information to belief and from action to utility is continuous, the mapping from belief to actions may not be. Nor do we think it reasonable to assume it is.

This somewhat technical point highlights a more fundamental way in which allocating information and allocating other goods are different problems. A planner

distortions in the structure of inquiries, receivers can use results from inquiries to learn something about the underlying probability over different states. Simply put, even if the receiver would have prefered a different inquiry, the information she receives is still valuable. Going forward, we shall call government inquires "signals" and their results "realizations."

seeking to maximize aggregate welfare by allocating consumption across agents need only know their endowments and preferences, so as to equate marginal utilities. The same is true in the Mirlees (1971) optimal income taxation setting, but subject to incentive compatibility constraints. Allocating scarce information involves two additional steps. When a planner allocates information she affects the *beliefs* of the receiver(s). These beliefs depend on both receiver's prior and the nature of the information. This information, in turn, affects the optimal action of receiver(s). It is these actions that map into utilities.[2] Thus, to equate the marginal utilities of the receivers is a significantly more complex problem for a planner than allocating scarce resources.

Given that government-provided information affects behavior, it is apparent that the information policy has important distributional effects. One might suppose that the optimal strategy to minimize allocative distortions might result in uneven incremental distribution of wealth. For example, if the optimal strategy is to announce the crime rates for the high crime area, then only the high crime area obtains value from information. Thus, rationing of information may involve a familiar trade-off between allocative and distributive benefits. We show this is true except when one area's state or actions have a positive spillover on another area. For instance, suppose criminals operate throughout a city, there are increasing returns to crime and criminals have nearly the same wage in alternative, non-criminal occupations as crime. When the government targets the high crime area, that area may begin taking more precautions. This may lead criminals to switch to alternate occupations. This could benefit the low crime area as well. In short, while there are distributive effects of information allocation, the distributive consequences may not track the targeting if there are positive spillovers across receivers.

If targeting information at one receiver is optimal, is it also optimal to maximize the accuracy of information to that receiver? In most cases, the answer is yes. However, if there are negative spillovers, then it may be optimal to target but with garbled information. For example, if accurately informing just the high-crime area

[2] That said, there do exist models in which beliefs directly affect utility holding actions constant. [CITE anticipation, placebo effect papers.]

substantially increases their level of precaution and that precaution encourages burglars to move to the previously low-crime neighborhood rather than outside the city,[3] then such information may not maximize welfare. However, providing less accurate information that induces only slightly higher precaution in the high crime area may not cause a shift among burglars and thus no spillover on low crime areas. Indeed, in the extreme, it may be optimal for the government not to provide any information even though costless information is a public good.

An obvious caveat to our findings concerning distribution and accuracy is that they depend on the nature of spillovers. As such they will depend on marginal direct value of information to targeted recipients versus the marginal indirect costs of information due to spillovers. In contrast, our finding concerning targeting is quite robust to spillovers. Regardless of the nature of spillovers, it is unlikely that the optimality requirement to equate marginal utility from information yields reporting of a simple statistic, even after accounting for spillovers.

While our model and examples focus on the government's role in allocating information, our analysis applies broadly to any entity that produces information but must ration it amongst receivers. For example, because the local news cannot report temperature readings for each location in a city it must decide which locations to target in broadcasts. Another example is Consumer Reports, which has information on multiple dimensions of quality for a given class of consumer products such as TVs. If different types of consumers value different dimensions of quality, how should it weight those dimensions and thus allocate its information when giving an up or down rating for products? Our findings apply to these problems, even though the private

[3] In our illustration, spillovers operate as negative externalities (additional crime in the low crime area) from actions in the high crime area. This is not the only mechanism by which spillovers may operate. Actions in the high crime area may induce responsive action in the low crime area (additional precautions in the low crime area to offset the flow of criminals). Moreover, spillovers may also operate through changes in the beliefs of the high crime area. For example, if the low crime area believes that the correlation between crime in their neighborhood and the other neighborhood is more negative than it really is, then a signal to the high crime area may cause the low crime area underestimate the crime in their area. As a result they will take inadequate precautions. Importantly, our claims about the effect of spillovers on distribution and accuracy do not depend on the mechanism of the spillover, whether due to negative externalities or strategic response to actions or whether due to actions or beliefs.

entities that ration information may have different preferences than the benevolent planner we consider in this paper.

We see our paper as a cautionary tale. A large and growing literature suggests that planners can make people better off by manipulating their information set (see, for instance, Sunstein & Thaler (2008)). Moreover, this literature celebrates the simplicity of the interventions that are allegedly required. We sharply question this by emphasizing the significant informational burdens on a planner allocating scarce information in a second-best world.

We proceed as follows. Section 2 outlines our model and provides a brief discussion of our modeling assumptions. In section 3 we offer a very simple example which illustrates how the government optimally decides what signal to send. Section 4 is the heart of the paper, providing results on targeting, distribution and accuracy. We connect our framework to welfare theorems and discuss the robustness of our model in section 5. Section 6 concludes.

2 The Model

2.1 Overview

There are three parties: a government and two receivers. For each receiver there is a true "state of the world", but they do not know what it is. The government sends the receivers some information and the receivers decide whether to update their beliefs about states. The receivers take an action, and their payoffs depend on the actions both take, as well as the true state of the world for each. Our question is what sort of information the government should send.

2.2 Statement of the problem

2.2.1 Signals

There are two receivers (R1 and R2) indexed by $i = 1, 2$. For each receiver there is a state space $\Omega_i = \{X, Y\}$. Define $\Omega = \Omega_1 \times \Omega_2$. No party directly observes the true state. Both receivers have a common prior μ_0 over Ω.

The government (G) shares this prior, and has access to two signal structures (π_1, π_2) which provide information to it about the true state. A signal structure is comprised of a realization space $S = \{x, y\}$ and a family of conditional distributions $\{\pi_i(\cdot|\omega)\}_{\omega \in \Omega_i}$ that map true states into realizations observed by the government.[4] Signal structure 1 provides the best information for R1, and signal structure 2 provides the best information to R2. Further, define the *accuracy* of a signal structure π_i as $\zeta_i = \frac{1}{2}(\Pr(x|X) + \Pr(y|Y))$. Note that a minimally accurate signal has accuracy 1/2 and a maximally accurate signal has accuracy 1.

2.2.2 Blended signals

The government can create a third signal structure that is a mixture of distributions of signals from (π_1, π_2) with weight $\gamma \in [0, 1]$ on π_1. The government chooses γ to create this mixture. For the bulk of the paper we will stick to this one-parameter (i.e. γ) means of obtaining signal structure 3, but we note at this point that the set of possible signal structures is far richer than this. While the government has access to signal structures 1 and 2, it is only signal structure 3 – a mixture of the first two – that the government can convey to receivers. This limitation is how we capture the idea that information is scarce.

Choosing a γ weight equal to 0 or 1 is defined as targeting a signal to R1 or R2, respectively. Choosing any other weight in $(0, 1)$ is defined as blending signals. Because π_1 and π_2 provide more accurate information to R1 and R2, respectively, one can speak of the government's choice of γ as a choice of the amount of accurate information to send to each receiver, with higher γ implying the allocation of more accuracy to R1. Determining this the optimal weight choice is how we operationalize allocating scarce information as between R1 and R2.

The weight γ also corresponds to a class of simple statistics, specifically weighted averages. At $\gamma = 1/2$, the government is simply reporting a mean. We can capture other simple statistics with a function $\theta(\pi_1, \pi_2)$. To keep the intuition behind our

[4] In an abuse of notation, we will sometimes refer to π_i as a realization from signal structure i where it will not cause confusion.

results simple, we focus on simple statistics captured by a choice of γ. Our results can be generalized to the more general class of simple statistics. We will distinguish between the governments optimal choice of a weight and a simple statistic by calling the former γ^* and the latter $\hat{\gamma}$. When $\gamma^* = \hat{\gamma}$, we say the government's optimal choice is to send the simple statistic $\hat{\gamma}$.

2.2.3 Updating, actions and welfare

Each receiver has preferences which can represented by a von Neumann-Morgenstern utility function $u_i(a_i, a_{-i}, \omega_i, \omega_{-i})$. After observing the realization of the signal, R1 and R2 update their beliefs according to Bayes rule, leading to posterior beliefs μ_s. Each receiver then, simultaneously and non-cooperatively, takes an action $a_i^*(\mu_s) \in \arg\max_{a \in A} E_{\mu_s^i}\left[u_i\left(a_i, a_{-i}^*(\mu), \omega_i, \omega_{-i}\right)\right]$, where $A \in \{\alpha, \beta\}$. We assume that if a receiver is indifferent between the two actions, given her posterior beliefs, she randomizes 50:50 between the two actions. The government chooses a weight γ to maximize welfare $V(u_1, u_2)$ subject to the constraints of how receivers update after observing a realization and how posterior beliefs map onto actions taken by the receivers. All primitives of the model are common knowledge among the three players.

Given a signal structure, each realization induces a posterior belief $\mu_s \in \Delta(\Omega)$ and, importantly, a distribution of posteriors $\tau \in \Delta(\Delta(\Omega))$, given by

$$\mu_s(\omega) = \frac{\pi(s|\omega)\mu_0(\omega)}{\sum_{\omega' \in \Omega} \pi(s|\omega')\mu_0(\omega')}, \forall s, \omega,$$

$$\tau(\mu) = \sum_{s:\mu_\mu} \sum_{\omega' \in \Omega} \pi(s|\omega')\mu_0(\omega'), \forall \mu.$$

We shall say that a distribution of posteriors satisfies the *Bayesian Budget Constraint* if the expected posterior probability is equal to the prior probability.

Kamenica & Gentzkow (2011) demonstrate that a commitment to report the realization from a signal structure that satisfies the Bayesian Budget Constraint ensures that receivers of the report will update their beliefs in response to the realization.

We will use this result to ensure the government's signal is credible. Specifically, we assume the government can credibly commit to reveal the realization from a mixture of these two signal structures to the receivers. Moreover, we assume that the signal structures (π_1, π_2) each satisfy the Bayesian Budget Contraint. Because the mean of random variable distributed with a mixture distribution $\gamma\pi_1 + (1-\gamma)\pi_2$ is $\gamma\mu_1 + (1-\gamma)\mu_2$, where μ_i is the mean of a random variable from with distribution π_i, the mixture also satisfies the Bayesian Budget Constraint. Thus the government's credible commitment to reveal the realization from such a mixture is sufficient to ensure the receivers update their beliefs in response to the government report.

2.3 Discussion of the Model

We pause briefly to discuss some of our modeling assumptions. We have intentionally kept a number of primitives of the model simple: the state space and the action space is binary for each receiver. These assumptions are for parsimony and ease of exposition. We do not believe that any substantive result in the paper would be materially altered if we let the state spaces Ω_i be arbitrary compact metric spaces (even non-finite), and the set of feasible actions be some compact set. We define the product set $\Omega = \Omega_1 \times \Omega_2$ and endow the players with prior beliefs over that space in order to allow for arbitrary correlation between the beliefs of the receivers about their own state space. For instance, knowing that my neighborhood has high crime rates may be informative about crime rates in your neighborhood. The assumption that receivers share a common prior over Ω is restrictive in the sense that it does not allow a receiver to know more about her state space than the other receiver. It is natural to think, for example, that I know more about the crime rate in my neighborhood than yours. Later in the paper we discuss allowing each receiver to obtain a private signal, which can differ in its informativeness, about her own state space, in addition to the government signal. As this requires some more complicated formalism, and it is not one of our central results, we do not introduce the additional formalism.[5]

Note that the government does not decide which signal to send after observing the realizations of both signals. Rather, G selects a signal structure. This implies that at the time G takes her action, she has no information which R1 and R2 do not have themselves. Thus there is no inference for the receivers to draw from G's action. Thus we have a dynamic game of complete information and can use subgame perfect equilibrium as the solution concept. If G had some private, payoff relevant information when she took her action then we would have a dynamic game of incomplete information, and would need to use sequential or perfect Bayesian equilibrium as a solution concept. This typically leads to multiple equilibria which are very sensitive to choices of out of equilibrium beliefs. The latter, we feel, is not the appropriate model for the questions at hand.

One key assumption we make is that the government is constrained in her communication abilities. She can send either a single signal to both receivers, rather than two signals: one tailored to each receiver. Alternatively, she can access a *blended* signal structure which represents a way of combining signals (as akin to avergaing them, as a special case).[6] As discussed in the introduction, this is a way of modeling the idea that there are either significant research, communication or information processing costs.

Our second key assumption is that the government can commit to reporting the realization of a mixture of (π_1, π_2), each of which satisifies the Bayesian Budget Constraint. A necessary primitive for a model of how to allocate scarce information is that there is some information that the government can send that receivers believe.

[5]We also assume receivers update according to Bayes Rule. This is also more strict than required. Kamenica & Gentzkow (2011) show that an assumption that the government can commit to conducting and reporting the results of a particular inquiry (signal structure) implies that Receivers will update in response to the Sender's reports so long the Receivers follow Bayes Rule and the inquiry conforms to Bayesian Budget Contraint. They do not show Bayes Rule is necessary. Nor is precommitment to a signal structure necessary for Receivers to believe signals from a Sender. We do not explore alternatives as the purpose of this paper is not to ask how the government can send credible signals but how the government should allocate information when it can send credible signals.

[6]Tehcnically, we allow Sender to form a third signal structure which is a convolution of the family of conditional distributions from signal structures one and two.

In the absence of such a primitive, the question of how to allocate scarce information is trivial because information does not impact receivers beliefs and thereby utility. In a framework with Bayes rational receivers, the fact that the government has information that receivers believe – in the absence of futher assumptions about, e.g., costly signalling – means that the information must satisfy the Bayesian Budget Constraint. In short, if government information matters to receivers and receivers are Bayesian, our second key assumption seems the most parsimonious primitive for our model.

3 An Example: Crime Rates

We now provide a very simple example which highlights how the model works. R1 (who we will call Rachel) and R2 (Rosalind) are residents of different neighborhoods within a city (which consists for these purposes of two neighborhoods–say, the north side and south side of Chicago respectively). The state space Ω_i for each is that crime can be H(igh) or L(ow), so that $\Omega = \{H, L\} \times \{H, L\}$. There are thus four states of the world: (H, H)–crime rates are high in both neighborhoods, (L, H)–the crime rate is low on the north side and high on the south side, (H, L)–high on the north side and low on the south side, or (L, L)–low in both neighborhoods. Let the common prior μ_0 be 0.3 that the true state ω is (H, H), 0.4 that $\omega = (L, H)$, 0.1 that $\omega = (H, L)$, and 0.2 that $\omega = (L, L)$.

G has access to two signal structures. Signal structure π_1 provides information about crime rates on the north side, and signal structure π_2 provides information about crime rates on the south side. Suppose that $\theta_1 = \theta_2 = 0$, so that information about the North side is irrelevant for the South side and *vice versa.* For both π_1 and π_2 the realization space $S = \{h, l\}$, with the interpretation that h is a signal that crime is high and l is a signal that crime is low. The signals, however, are not perfect. The family of conditional distributions for π_1 is as follows

$$\pi_1\left(h|H \text{ on north side}\right) = 3/4 = \pi_1\left(l|L\right),$$
$$\pi_1\left(h|L\right) = 1/4 = \pi_1\left(l|H\right).$$

For π_2 the conditional distributions are

$$\pi_2\left(h|H \text{ on south side}\right) = 3/5 = \pi_2\left(l|L\right),$$
$$\pi_2\left(h|L\right) = 2/5 = \pi_2\left(l|H\right).$$

We will interpret action α as being C(autious), and action β as being N(ormal). Suppose that Rachel and Rosalind's utility functions are the same and depend only on their own action and their state of the world. Further, suppose that taking action C entails a cost of 1, whereas N is costless. Also suppose that the benefit of being cautious when the neighborhood is high crime is 3 and of behaving normal when when the neighborhood is high crime is -2. Otherwise benefits are 0. We thus have

$$u\left(C,H\right) = 3-1 = 2$$
$$u\left(C,L\right) = 0-1 = -1$$
$$u\left(N,H\right) = -2-0 = -2$$
$$u\left(N,L\right) = 0-0 = 0$$

for both receivers. We assume $v = u_1 + u_2$.

If either Rache or Rosalind got no signal, they would each take precaution. Given Rachel's prior about crime on the north side of Chicago, her payoff from taking precaution is

$$\begin{aligned} u_1\left(C\right) &= u\left(C,H\right)\mu_0\left(H,\cdot\right) + u\left(C,L\right)\mu_0\left(L,\cdot\right) \\ &= \left(2\times 0.4\right) + \left(-1\times 0.6\right) = 0.2 \end{aligned}$$

Whereas if she took no precaution, her payoff would be

$$u_1(N) = u(N, H)\mu_0(H, \cdot) + u(N, L)\mu_0(L, \cdot) = -0.8$$

Taking precaution offers a higher utility payoff. The same is true for Rosalind:

$$u_2(C) = u(C, H)\mu_0(\cdot, H) + u(C, L)\mu_0(\cdot, L) = 1.1 > -1.4 = u_2(N)$$

If Rachel gets the signal, her Bayesian posterior beliefs about crime rates in her neighborhood are

$$\begin{aligned}
\mu_1(H|h) &= \frac{\pi_1(h|H)\mu_0(H, *)}{\pi_1(h|H)\mu_0(H, *) + \pi_1(h|L)\mu_0(L, *)} \\
&= \frac{3/4 \times 0.4}{3/4 \times 0.4 + 1/4 \times 0.6} = 2/3 \\
\mu_1(L|h) &= 1/3 \\
\mu_1(H|l) &= 0.18 \\
\mu_1(L|l) &= 0.82.
\end{aligned}$$

We can now determine what action Rachel will take if she gets the signal h or l:

$$\begin{aligned}
u_1(C, h) &= u(C, H)\mu_1(H|h) + u(C, L)\mu_1(L|h) \\
&= 1 > -0.45 = u_2(N, h) \\
u_1(C, l) &= -2.1 < -0.36 = u_2(N, l)
\end{aligned}$$

Thus Rachel takes precaution if she observes h and behaves normal if she observe l. Sending π_1 does not change Rachel's behavior if she observes h, but does prevent her from wastefully taking precaution if she observes l.

By contrast, if Rosalind gets the signal, she does not change her behavior. If Rosalind gets the signal, her Bayesian posterior beliefs are crime in her neighborhood are

$$\mu_1(H|h) = 0.78 \quad \mu_1(L|h) = .22$$
$$\mu_1(H|l) = .61 \quad \mu_1(L|l) = 0.39.$$

Whether she observes h or l, she'll take precautions because

$$u_2(C,h) = 1.34 > -1.56 = u_2(N,h)$$
$$u_2(C,l) = .17 > -0.78 = u_2(N,l)$$

Given that the Rosalind takes precaution whether or not she gets a signal, the government can only improve its payoff by sending a signal to Rachel. Rachel and Rosalind can figure that out, and so only Rachel will update her beliefs and change her actions following the signal.

4 General Results

In this section we offer a number of general results, centered around four themes: targeting v. blending, the value of simple statistics, distributive impacts and the value of accuracy. Before doing so, however, we establish a benchmark result which provide conditions under which the first-best is attainable. First, let us define the first-best payoff to the government as V^{FB}. This is the payoff in the absence of any communication constraints. To that end, define $\hat{u}_i(\pi_j) = u_i\left(a_i^*(\mu_s), a_{-i}^*(\mu_s), \omega_i, \omega_{-i}\right)$ when signal π_j is sent. We can then state

Proposition 1 *Suppose* $dV(u_1,u_2)/du_i = 0$ *or* $\hat{u}_i(\pi_i) = \hat{u}_i(\pi_{-i})$, *then* $lim_{\zeta_i \to 1}(V(u_1,u_2)) \to V^{FB}$.

Proof of Proposition 1. By the supposition of the proposition we can, w.l.o.g. write $V(u_1,u_2)$ as $V(u_1)$. As $\zeta_1 \to 1$ receiver 1's action approaches a_1^*. Routine continuity arguments imply that $u_1 \to u_1^{FB}$ and hence $V(u_1,u_2) \to V^{FB}$ as required. ■

In words, if the government only cares about the preferences of one receiver, then it can achieve the first best by sending a perfectly informative signal to that party.

Alternatively, if targeting a signal at a receiver does not change that receiver's utility, then the government can achieve the first best by targeting the other receiver. Either way, the communication constraints in our model are irrelevant. This result is far from surprising, but provides a useful benchmark for understanding deviations from it, as we will see below.

Our second proposition explains when the government may want to target information at one particular receiver.

Proposition 2 *If V is linear in u_i and u_i is linear in γ, then the optimal action for the government is to send a signal to one receiver or to none, but not to blend signals. If V is not linear then it may be optimal for the government to blend signals.*

This is, again, a rather elementary result, but helps to ground ideas. When the government's payoff is linear in the accuracy of the signal it can send to either receiver then it wishes to target the information to one or other receiver. The more striking conclusion is just how demanding this linearity condition is and that, when it does not hold, the government wishes to blend signals.

The next proposition highlights how challenging it is to perform optimal blending using simple statistics such as a mean.

Proposition 3 *Suppose there exists an optimal weighting γ^* and simple statistic $\hat{\gamma}$ such that, for a given set of continuous functions $\{V(\cdot), u_1(\cdot), u_2(\cdot), a_1(\cdot), a_2(\cdot)\}$, $\gamma^* = \hat{\gamma}$. There does not exist for generic (π_1, π_2) an ε perturbation of any of the functions $V(\cdot)$, $u_1(\cdot)$, $u_2(\cdot)$, $a_1(\cdot)$, and $a_2(\cdot)$ such that $\gamma^* \neq \hat{\gamma}$.*

The government's optimal weight equates the welfare-weighted marginal utility of accuracy across receivers: abusing notation, $(\partial V/\partial u_1)(\partial u_1/\partial \gamma) = (\partial V/\partial u_2)(\partial u_2/\partial \gamma)$. Even when this aligns with a specific simple statistic $\hat{\gamma}$, a small perturbation of the functions (V, u, a) that dictate the mapping from R1 weighting γ to welfare V can change marginal utilities at $\hat{\gamma}$ such that the simple statistic no longer equates weighted marginal utilities. Another way to put this is that simple statistics are not defined by their ability to equate weighted marginal utilities of more information and thus should not be expected to provide an optimal allocation of information.

A feature of the planner's problem that can help illustrate why the allocation of information is different than the allocation of divisible goods is that, when welfare functions and utility functions are continuous, the optimal allocation of divisible goods to an agent is continuous in that agents' initial endowment of that good. Yet similarly strong assumptions on welfare and utility do not imply that allocations of information are continuous in agents' initial endowments.

Proposition 4 *Suppose* $(V(u_i, u_{-i}), u_i(a_i, a_{-i}, \omega_i, \omega_{-i}), \mu_s(\mu_0, \gamma))$ *are well-behaved, including continuous in their arguments. It is not the case that* γ^* *is continuous in* ζ_i *and* ζ_{-i}*.*

Proof. This follows directly from the fact that a_i need not be continuous in the realization of $\gamma\pi_1 + (1-\gamma)\pi_2$. ■

The optimal allocation of information, i.e., the government's choice of weight depends on the mapping from information to posterior to action to utility to welfare. Even if posteriors, utility and welfare is continuous, the mapping from beliefs to actions may not be continuous. This result is not sensitive to our assumption of a binary action space. Even if actions were continuous, a necessary condition for a continuous mapping from accuracy endowments to weights is a continuous mapping from beliefs to actions. In environments with non-continuous payoffs to action, environments which are not hard to imagine when there are strategic interactions, it is unreasonable to assume a continuous mapping from beliefs to actions.

In the setting we are considering, with communication constraints, provision of information typically has distributive consequences. A simple intuition, which is misleading, is that the party that receives the signal is necessarily better off. This is false, as we formalize in the next proposition. First, define three mutually exclusive cases: (i) if the government chooses to report π_i, then u_j is unaffected relative to when the government reports no information; (ii) if the government chooses signal π_i, then u_j is lower relative to when the government reports no information (negative spillovers); (iii) if the government chooses signal π_i, then u_j is higher relative to when the government reports no information (positive spillovers).

Proposition 5 *In cases (i) and (ii), if the government chooses a higher γ, then the utility of R1 (weakly) rises more than the utiltiy of R2.*

This may seem counter-intuitive at first. After all, if the signal is making someone worse off shouldn't she just ignore it. The answer is that it is not optimal to do so. The information is valuable, even though different information might have been *more* valuable. This is similar to the results on Bayesian persuasion, where a receiver can have her action "manipulated" by the sender by a signal altering the distribution of posterior beliefs, holding the mean fixed (Kamenica & Gentzkow 2011).

It is not hard to provide examples to this effect, as one has the freedom to choose any utility functions, signal structures, and mapping from signals to optimal actions. A useful avenue for further work might be to find meaningful restrictions on the environment which rule out the above result.

We now turn to the effects of the accuracy of signal structures. Given the three cases just defined, we can now state

Proposition 6 *In cases (ii), if the government chooses signal π_i, then it could be optimal for the government to also Blackwell garble the signal π_i.*

It would seem that better information is beneficial. After all, better information might cause a receiver to take a better action from the government's perspective. That is, abusing notation, $\partial V(u_1, u_2)/\partial \zeta_i \geq 0, \forall i$. This reasoning ignores the possibility that the signal to i might lower u_j. For example, the signal to i might cause j to update her beliefs about her environment in a manner that is erroneous (though consistent with Bayes theorem). The resulting action could lower j's utility more than the signal increased i's utility. It is also possible that i's action has negative spillovers. In our crime example, perhaps when i adds an alarm to her house, criminals move over and attack's j's house. Finally, it is possible there are strategic aspects to the problem. When one receiver changes her action it may well change the optimal action of the other receiver–they are playing a game.

So the real question is what is the sign of the following object: $dV(u_1, u_2)/d\zeta_i$? It turns out that it can be positive or negative. The reason is that there is both a direct and indirect effect of providing better information to a receiver[7] The direct effect is certainly positive, but the indirect effect can be negative and, moreover, can be larger that the direct effect. That is, from the government's perspective, the negative consequences of the change in the action of receiver j from providing better information to receiver i can offset and even outweigh the benefits from an improved action from receiver i.[8]

Perhaps it is not surprising that negative spillovers might make it better to send less accurate information. A similar thing can be said about ordinary goods: the government should not send i more of a good if doing so has a larger negative effect on j. However, with information we can show something similar even in the absence of spillovers. For example, we can state the following.

Proposition 7 *Define $\sigma(\mu) = \int [\omega - E_\mu(\omega)]^2 \mu(\omega)\, d\omega$ as the variance of (a random variable with) distribution μ. Suppose π_i increases (decreases) the variance of the posterior, i.e., $\sigma(\mu_s) > \sigma(\mu_0)$. If u_i is declining (increasing) in the variance of the posterior, then more it may be better to send no signal to i.*

Proof. See Kamenica & Gentzkow (2011), Remark 1 and Proposition 2. ■

A government signal meets the Bayesian Budget Constraint – and is thus credible – so long as it does not shift the expected the mean of the posterior. However, the chosen signal structure can modify other moments, e.g., the variance, and still be credible. If the receiver's utility is a function of the variance of the expected posterior, then sending information can both decrease utility even if it is informative in the sense that it is Bayesian rational for the receiver to update in resposne to it. For example, suppose an individual has a payoff like the prosecutor in Kamenica &

[7] In the case of strategic effects, there is also an indirect effect on receiver i from receiver j obtaining better information, but this is zero by the Envelope Theorem. Readers familiar with industrial organizational will see the connection to the reasoning in, for example, Tirole (1988).

[8] The result in the strategic context is intuitively similar to a number of classic results in industrial organization. For instance, Grossman (1981) shows that when two firms compete in a potentially differentiated product market and can also advertise, an increase in advertising costs can actually make firms better off. Again, this is the strategic effect outweighing the direct effect. The direct effect of an increase in advertising costs is bad–higher costs mean lower profits. But the strategic effect is that it leads to less advertising and hence more product differentiation, which can lead to an overall increase in profits.

Gentzkow (2011), i.e., gets a payoff if she can show a judge – based on the realization from the government's signal structure – that the defendant is guity with probability 0.95, but not otherwise. Then adding variance to the signal structure increases the prosector's payoff, because, for any prior that the defendant is guilty with probability less than 0.95, it increases the chances of the prosecutor prevailing. In this context, if the government's signal structure actually reduces the variance of the prosecutor's expected posterior, then it will reduce the prosecutor's utility. It may be better for the government not to send the prosecutor a signal. If this is true for both receivers, it might be better for the government to send no signal to any receiver.

Note that we are not claiming that a more accurate signal structure is *always* bad, merely that it can be. Again, understanding if there are economically meaningful restrictions on the environment that cause more information to be necessarily better would be valuable, but seems challenging given the richness of the setting.

5 Discussion

Our model illustrates that the government may want to target information regardless of spillovers. Allocation of information has distributive effects that track the initial allocation unless there are positive spillovers, which can offset the initial allocation. Finally, even if accuracy is costless, the government may not want to send more accurate signals if there are negative spillovers or the signal shifts the expected shape of the posterior in a manner that reduces utility. In that case, the government should consider the trade-off between the direct positive value of accuracy to the receiver and exacerbation of the negative spillover effects or deleterious shift in the shape of the posterior. Optimal accuracy will equate marginal effects on these two fronts. In the extreme case where the negative spillovers or shape changes from the very first increment of accuracy exceed the positive direct benefits of that accuracy, the government may not want to provide any information to any receivers. Because the model does take a position on the source of the spillovers, it captures previous theories that suggest additional information may lower welfare, including the market for lemons (Akerlof 1970) , adverse selection in insurance markets (Rothschild and

Stiglitz 1976), and the Hirshleifer effect in trading of contingent claims (Hirschleifer 1971). Because the model does not take a position on the realtionship between the shape of the posterior and payoffs, it captures Bayesian persuasion (Kamenica and Gentzkow 2011).

In this section we consider the robustness and scope of our model. This is best done by considering how sensitive the results are to modification of the various assumptions made in the model. Some modifications are innocuous. For example, the model assumed only two types of receivers. In reality there is something closer to a continuum of types of receivers. In that context, our basic findings continue to hold. Targeting is still appropriate. Spillovers may be more complicated, but distribution and accuracy depend on those spillovers in the same manner as the binary model.

Other assumptions do more substantive work. For example, the model assumed that the government and the two receivers had common priors μ_0. Suppose that, prior to the government sending a signal, however, each receiver obtains a private signal that causes them to have different priors μ_0^i. It is useful to consider two extreme cases here. The first is where the private signal is equal to a government signal with accuracy ζ_i targeted at receiver i. Thus a targeted signal from the government does not change the expected state but rather the variance across states or confidence. In this case, the model would still recommend targeting, though the returns to targeting are keyed to the utility of increasing confidence, rather than moving beliefs about the mean state. The second case is where the private signal is inaccurate in a way that makes the receiver's beliefs about the their states diverge from each other in the sense that the expected state of each receiver under his prior is farther from the actual state of the other receiver: $E_{\mu_0^i}[\omega_i] - \omega_{-i} > \omega_i - \omega_{-i}$. If the divergence is large enough, it could be that a mixed strategy by the government could bring each parties posterior closer to the truth: $E_{\mu_0^i}[\omega_i] - \omega_i > E_{\mu_s^i}[\omega_i] - \omega_i$. In words, a mixed government signal may correct private signals that are erroneous, but only in one special case.

Another substantive assumption is that the receivers can back out the government's optimal behavior and thus know when to ignore the government's signal. An alternative assumption would be to presume that receivers make mistakes in solving the government's problem and thus, if the government is targeting i with weight γ, the receivers think the government is targeting i with weight $\alpha(1/2) + (1 - \alpha)\,\gamma$. This sort of mistake should not affect the targeting result so long as $\alpha < 1/2$, that is, the receiver is not more likely to believe the government targeted i when it actually targeted $-i$. The reason is that mistakes by the receiver function as blending by the government. Moreover, any actual blending by the government functions to further move the de facto blending closer to $1/2$. For the same reasons that blending lowers welfare generally, blending further than implied by receivers' errors in solving the government's problem also lowers welfare.

6 Concluding Remarks

Our model concerns the allocation of scarce information by a social planner. If information were like any other good, the model would establish a benchmark against which alternative methods of allocation, e.g., a Walrasian equilibrium, could be judged under criteria such as Pareto optimality. However, the information we consider is unlike any other good, even in the absence of spillovers. We study information that is a local public good and is thus undersupplied by the private agents. Moreover, the government produces information that is useful only to certain locales, and thus receivers. For each of these reasons, there is also no scope for *ex post* trade across locales and thus a Walrasian equilibrium. This rules out the possibility of analogues to the first or second welfare theorems for information. It also makes it vital that the government hit the constrained optimal allocation; *ex post* trade cannot correct initial misallocations. A corollary is that considerations of distribution cannot generally be divorced from considerations of allocative efficiency.

Because ex post trade is not possible, allocation of information has non-pecuniary externalities, even in the absence of spillovers of the sort we discussed above. Thus information suffers the same problems as behavior with externalities, such as playing loud music or polluting. To the latter problem there are two solutions that involve non-governmental allocations and can replicate the social planner's allocation: Pigouvian taxes without trade or Coasian trade in the right to engage in behavior.

There exist analogues to each in the information allocation problem. For example, the government may sell the signal. If it sets prices (Pigouvian taxes) appropriately, it can replicate the results of direct allocation. Alternatively, the government can allocate the right to obtain the signal and receivers can engage in *ex post* trade in the right (assuming no transactions costs).[9] Neither solution achieves the first best because information is a local public good. The reason information is underproduced locally is the reason it would be undervalued in trade among local agents. If we ignore this, however, both mechanisms can achieve the first best allocation and disentangle allocative efficiency from distribution. We stress, again, that all these results obtain even without spillovers that do not stem from the impossibility of ex post trade in information. Such spillovers just compound the allocation problem.

If we continue to ignore that information is a local public good, the possibility of *ex ante* trade in rights to information raises the question whether fully state contingent contracts over actions negotiated prior to allocation of a signal can yield the same welfare as ex ante trade in rights. For example, regardless to whom the government sends a signal, can a set of privately negotiated contracts that specify the action each receiver is to take conditional on realization of signals sent to that receiver achieve the same as ex ante trade in signals? Presently, we believe the answer is uncertain. Welfare is maximized by specifying both who receives the signal and the actions parties take conditional on that signal. Contingent contracts do not specify the allocation of a signal, only the actions conditional on allocation. (The government is not a party to the contracts.) If the government sends a signal to the wrong receiver, contingent contracts cannot achieve maximal utility. Contingent contracts can only make the best of a bad allocation. Conversely, *ex ante* trade in rights can change the government's target for a signal, but the target is selected assuming no specification of ex post actions. This achieves maximal utility only where there are no spillovers after signals are sent, in which receivers take optimal actions (and contingent contracts would not bind). With spillovers, however, ex

[9]Interestingly, one rarely observes either the government selling the rights to a signal or assigning rights to a signal and allowing trade in that right all prior to sending a signal. We leave that puzzle for future research.

ante trade only achieves the best allocation assuming privately optimal actions that have spillovers. It is ambiguous whether this yields greater utility than contingent contracts. The superior strategy depends roughly on whether the initial allocation is more important or spillovers are.

References

Akerlof, George, "The market for lemmons," *Quarterly Journal of Economics*, 1970, *84*, 488–500.

Grossman, Sanford J., "The Informational Role of Warranties and Private Disclosure about Product Quality," *Journal of Law and Economics*, 1981, *24* (3), 461–483.

Hirschleifer, Jack, "The private and social value of information and the reward of inventive activity," *American Economic Review*, 1971, *61* (4), 561–574.

Kamenica, Emir and Matthew Gentzkow, "Bayesian Persuasion," *American Economic Review*, 2011, *101* (6), 2590–2615.

Mirrlees, James A., "An Exploration in the Theory of Optimal Income Taxation," *Review of Economic Studies*, 1971, *38* (2), 175–208.

Rothschild, Michael and Joseph E. Stiglitz, "Equilibrium in competitive insurance markets: An essay on the economics of imperfect information," *Quarterly Journal of Economics*, 1976, *90*, 629–649.

Thaler, Richard H. and Cass R. Sunstein, *Nudge*, Yale University Press, 2008.

Tirole, Jean, *The Theory of Industrial Organization*, MIT Press, 1988.