DEPARTMENT OF ECONOMICS

WORKING PAPER SERIES

2017-10

McMASTER UNIVERSITY

Department of Economics
Kenneth Taylor Hall 426
1280 Main Street West
Hamilton, Ontario, Canada
L8S 4M4

<http://www.mcmaster.ca/economics/>

Moral Hazard and Horizontal Information

Anastasios Papanastasiou*

McMaster University

June, 2017

Abstract

This paper examines a moral hazard problem where the principal observes a location signal of the agent's unobserved effort. When the agent incurs costlier effort the demand for the agent's services is *ex-ante* more spread out over a location measure of demand. If location outcomes are informative about the agent's action and verifiable, the principal can improve efficiency by contracting on such information. The introduction of location information suggests that the optimal contract that maximizes expected market size can be locally decreasing in sales when small-market location outcomes are a strong signal of high effort (JEL D82, D86).

Keywords: Moral Hazard, Market Position, Market Size

*Address: KTH 434, Department of Economics, McMaster University, 1280 Main St W, Hamilton, ON L8S 4L8
e-mail: papanasa@mcmaster.ca. I would like to thank René Kirkegaard for many valuable comments. I also thank Brian Ferguson, Johanna Goertz, Mike Hoy, Pau Pujolas, Thanasis Stengos and Mark Tremblay for their comments on earlier versions of this work. All errors are mine.

1 Introduction

In the standard moral hazard problem a principal delegates the task of sales to an agent and in order to induce the agent to incur the level of effort that maximizes expected sales, the principal offers a contract that specifies transfers which are conditional on the volume of sales. If Milgrom's (1981) monotone likelihood ratio property (MLRP) holds, so that higher effort in sales always increases the expected volume of sales, the optimal transfers to the agent are monotonic in market size. But when the outcome of the agent's actions contains verifiable information on the relative location of the market, should the principal that seeks to maximize sales still provide transfers that are monotone in sales? It turns out that introducing a location measure of information implies that the optimal contract that induces the agent to exert the level of effort that maximizes expected sales, has transfers that can be locally decreasing in the volume of sales, even when costlier effort always increases expected sales without any indication that MLRP is violated. The intuition behind this result is that, in certain cases, small-market outcomes can be a strong signal of high effort relatively to high-volume outcomes that are less likely to result from high effort.

It is not unusual for the outcome of the agent's effort to contain information on *who* buys from the agent, on top of *how many*. For instance, a branding agency can be hired by a firm to design a new advertisement that appeals *ex-post* more so to a particular audience in the spectrum of potential consumers, while other consumers may find the ad unattractive, depending on their intrinsic characteristics, such as aesthetic preferences, personal beliefs or cultural background, et cetera. If the firm intends to promote sales within a certain geographic region, it may specify *ex-ante* greater rewards for advertisements that will *ex-post* attract that population of consumers. Similarly, a pharmaceutical firm that is after developing a drug that works against a particular serotype of a bacterial species which is resistant to standard methods of treatment, may offer more generous payments to researchers who develop drugs that appear more promising against that strain. In general, if the principal attaches more value to a specific segment of demand she can encourage the agent to work harder towards capturing that particular segment by offering payments conditional on market location.

If location outcomes are of direct value to the principal, location information can be in principle mapped to a set of cardinal numbers, with more valuable location outcomes translating to a higher payoff for the principal. But even when the principal has no preferences about location outcomes, there is an indirect value to verifiable location information when it's a signal of agent's unobserved effort as it can be utilized to provide incentives for the intended level of effort at a lower expected cost (Holmström, 1979). I consider this baseline case, where location outcomes are informative about the agent's action but of no direct value to the principal who is indifferent about who consumes the product and only seeks to maximize the volume of consumer demand. Under standard assumptions, the optimal contract has U-shaped transfers in location outcomes, and in some cases the agent's payment is greater when his sales efforts capture a niche subset of consumers.

2 Model

A continuum of potential consumers with unit-demands is uniformly distributed in the interval $\Omega = [\underline{\omega}, \bar{\omega}]$ whose length is normalized to one. A principal seeks to hire an agent to attract the most consumers in Ω without overcompensating him.¹ The principal announces the terms of the contract to the agent which he either rejects and no further events occur, or accepts, chooses his research effort, and is compensated according to the contract once uncertainty is resolved.

The agent's level of effort determines his ideas for his sales proposal. We say that the agent's final idea I *appeals* to consumer $x \in \Omega$, or equivalently x is *attracted* by I , if the ex-post demand of x equals to one.²

Assumption A1: (*Convexity in consumer attraction*) If the agent's idea I appeals to two distinct consumers x' and x'' then it also appeals to all consumers located *between* x' and x'' .

From the agent's point of view, A1 implies that if he gets an idea for a sales proposition that appeals to x' and another idea that appeals to x'' he can combine the two ideas to attract all consumers located between x' and x'' . For the consumers, A1 requires that ones within smaller neighborhoods in Ω have similar predispositions with respect to the sales proposal of the agent.

Let x' and x'' denote the two *extreme* consumers that are known to be attracted by the agent's sales efforts. By A1 the sales proposal will appeal to every $x \in [x', x'']$, given that $x' < x''$, and its market size is $d = |x' - x''|$. In the standard model, the principal only observes d and the optimal transfers are constant for all values of x' and x'' as long as d is constant.

I assume that the agent only has two ideas for attracting consumers, drawn by a distribution that is conditional on how hard he works on research $F(x|e)$, where e is his level of effort in sales. By A1, the agent's final idea I can be characterized by the pair (y_1, y_2) of the two extreme consumers that are attracted by the agent's sales pitch, where $y_1 = \max\{x', x''\}$ and $y_2 = \min\{x', x''\}$. The principal is assumed to observe directly the pair of extreme consumers (y_1, y_2) which is a signal of the relative *position* of the market.

Moreover, let $e = \{e^l, e^h\}$ so the agent can only exert either low (e^l) or high (e^h) effort in sales. Since the location of consumers cannot be directly compared in terms of their magnitude, it is invalid to use the standard assumption of first-order stochastic dominance to rank the conditional distributions as there is no direct way of identifying better outcomes in terms of the magnitude of

¹For expositional purposes one may think of the principal as the government who seeks an agent to develop a non-rival public good, and welfare maximization requires maximizing the expected number of consumers. Location preferences for the principal can be introduced as a density over the set of consumers representing the number of consumers that fall within that location.

²This definition implies that whether a consumer is attracted by the agent's idea or not, does not depend on the per-unit price of the product and to this extent the term "unit-demands" is used loosely. One way to think of this is to assume that per-unit prices are fixed exogenously. For example, the entry fee in a given theatre is usually constant regardless of the featured motion picture. Alternatively, if x is not attracted by the outcome of the agent's efforts, his consumer's surplus would be zero even when the good was given away for free. In the pharmaceuticals example, if a drug does not work for a given patient then there it must be of zero value to that patient since re-selling is not possible.

x .³ Spreads of densities are a better candidate since they imply different probabilities of attracting consumers in given neighborhoods when the agent changes his level of effort.

Assumption A2: (*Thinking outside-of-the-box as median preserving spreads of ideas*) An increase in e results in a clockwise rotation of F around its median.

Johnson and Myatt (2006) use a similar framework to analyze the effect of marketing strategies, in this scenario from the part of the agent, on consumer demand. By assumption, there is a critical point in $m \in \Omega$, the idea to attract the median consumer. When the agent increases his effort it is equally likely to attract consumers to the left or right of m , and the ideas become more spread out in Ω , making it more likely to attract consumers that are located in intervals further from the median consumer.

Figure 1: The location effect of an increase in sales effort, example.

The neighborhoods in Ω that become more likely to be attracted under high effort can be considered as the more "difficult" segments of demand, in the sense that the agent has to go beyond conventional sales methods to better the chances that his proposal will attract those particular consumers.

Proposition 1: If consumer attraction is convex and if higher effort in sales results in median preserving spreads of ideas, the high level of effort maximizes expected market size.

³Restricting the problem further, and focusing on market size, one can at this point impose the assumption that $E[d] = E[y_2 - y_1]$ is increasing in e in an FOSD sense.

Proof. Let $h_{1,2}$ denote the joint density of y_1 and y_2 under e^l . We have

$$E[(y_1 - y_2)|e^l] = \int_{y_2}^{\bar{\omega}} \left(\int_{\underline{\omega}}^{y_1} (y_1 - y_2) h_{1,2}(y_1, y_2) dy_2 \right) dy_1 \quad (1)$$

$$\begin{aligned} &= \int_{\underline{\omega}}^{y_1} \int_{y_2}^{\bar{\omega}} y_1 h_{1,2}(y_1, y_2) dy_1 dy_2 - \int_{\underline{\omega}}^{y_1} \int_{y_2}^{\bar{\omega}} y_2 h_{1,2}(y_1, y_2) dy_1 dy_2 \\ &= \int_{\underline{\omega}}^{\bar{\omega}} y_1 h_1(y_1) dy_1 - \int_{\underline{\omega}}^{\bar{\omega}} y_2 h_2(y_2) dy_2 \end{aligned} \quad (2)$$

where h_1 and h_2 are the probability density functions of the highest and lowest order statistics of x respectively, conditional on e^l .

Since $h_1(y) = 2H(y)h(y)$ and $h_2(y) = 2h(y)(1 - H(y))$ for $y_1 \geq y_2$, equation (2) can be written as

$$E[(y_1 - y_2)|e^l] = \int_{\underline{\omega}}^{\bar{\omega}} y_1 2H(y_1)h(y_1) dy_1 - \int_{\underline{\omega}}^{\bar{\omega}} y_2 [2h(y_2)(1 - H(y_2))] dy_2 \quad (3)$$

Integrating each term in the RHS of (3) by parts

$$1 - \int_{\underline{\omega}}^{\bar{\omega}} H(y_1)^2 dy_1 - \int_{\underline{\omega}}^{\bar{\omega}} (1 - H(y_2))^2 dy_2 \quad (4)$$

$$= 2 \int_{\underline{\omega}}^{\bar{\omega}} H(t)(1 - H(t)) dt \quad (5)$$

By symmetry, $E[(Y_1 - Y_2)|e^h] = 2 \int_{\underline{\omega}}^{\bar{\omega}} G(t)(1 - G(t)) dt$. Let $V(H(t)) = H(t)(1 - H(t))$. $V(H(t))$ attains a global maximum at $F(m) = \frac{1}{2}$ where m is the median of $H(t)$. That is, $V(H(t))$ decreases as $|H(t) - 0.5|$, or equivalently as $|t - m|$, increases. Since G and H have the same median, the same result holds for $V(G(t))$. Single crossing implies $|H(t) - 0.5| > |G(t) - 0.5|$ for all $t \neq m \in (0, 1)$; the direction of the inequality follows from the assumption that G crosses H from above. Thus $V(G(t)) > V(H(t))$ for all $t \neq m \in (0, 1)$ which implies that $E[(y_1 - y_2)|e^h] - E[(y_1 - y_2)|e^l] > 0$. Alternatively, $H(t)[1 - H(t)] = 0.25 - [H(t) - 0.5]^2$ so $H(t)[1 - H(t)] - G(t)[1 - G(t)] = [G(t) - 0.5]^2 - [H(t) - 0.5]^2$. Again, the last step follows from the single-crossing and common median assumptions. ■

By the above result, the expected volume of sales is increasing in the agent's effort, in consistence with the standard moral hazard assumptions on effort and sales. In order for the principal to maximize expected sales, she will offer a contract that induces the agent to put more than minimum effort in salesmanship. One way of addressing the moral hazard problem is to link the ex-post compensation of the agent to the volume of sales. If the agent's expected payoff under high effort is sufficiently high to compensate for the extra disutility that he receives from working harder, he will find it rational to do so. Indeed, the principal could simply offer payments that are increasing

in market size $d = |y_1 - y_2|$ to encourage the agent to put high effort, as the agent knows that d is greater under e^h . But since y_1 and y_2 contain additional information on market location, the principal could achieve a more efficient outcome by taking advantage of this information and specifying more elaborate payment plans that do not depend on market size exclusively.

Thus, instead of offering a contract $w(d)$ I assume that the principal offers $w(y_1, y_2)$ to the agent in exchange for coming up with the idea $I(y_1, y_2)$. I adopt the standard assumption that the agent has additively separable preferences over wealth and effort $v(e, w) = u(w) - c(e)$, where w is the payment that the agent receives and u and c are strictly increasing real-valued functions. The agent is assumed risk-averse with $u'' < 0$ and the principal is risk neutral.

Incentive compatibility requires that the agent's expected payoff is greater under the high level of effort,

$$\begin{aligned} & \int_{y_2}^{\bar{\omega}} \int_{\underline{\omega}}^{y_1} u(w(y_1, y_2)) g_{1,2}(y_1, y_2) dy_2 dy_1 - c(e^h) \\ & \geq \int_{y_2}^{\bar{\omega}} \int_{\underline{\omega}}^{y_1} u(w(y_1, y_2)) h_{1,2}(y_1, y_2) dy_2 dy_1 - c(e^l) \end{aligned} \quad (6)$$

and the participation constraint ensures that the agent's expected payoff from accepting the contract is at least as much as the value of his exogenous outside option r :

$$\int_{y_2}^{\bar{\omega}} \int_{\underline{\omega}}^{y_1} u(w(y_1, y_2)) g_{1,2}(y_1, y_2) dy_2 dy_1 - c(e^h) \geq r \quad (7)$$

The problem of the risk-neutral principal is to determine a payment mechanism $w(y_1, y_2)$ that minimizes her expected costs without violating the participation and incentive compatibility constraints, that is,

$$\min_{w(y_1, y_2)} \int_{y_2}^{\bar{\omega}} \int_{\underline{\omega}}^{y_1} w(y_1, y_2) g_{1,2}(y_1, y_2) dy_2 dy_1 \quad (8)$$

subject to (6), and (7).

Assuming transfers in the interior of the solution set, the necessary conditions to problem (8) are:

$$\frac{1}{u'(w(y_1, y_2))} = \lambda + \gamma \left[1 - \frac{h_{1,2}(y_1, y_2)}{g_{1,2}(y_1, y_2)} \right] \quad (9)$$

and must hold for all (y_1, y_2) , where γ, λ , are the respective multipliers to constraints (6) and (7).

The last set of optimality conditions suggests that the principal must introduce some variability in the compensation scheme that induces e^h , otherwise the agent will exert low effort. As in the

standard problem, the agent will choose not to receive the extra disutility from working harder when he receives a fixed payment.

The payment scheme $w(y_1, y_2)$ that solves the principal's problem depends on the likelihood ratio $z_{1,2}(y_1, y_2) = \frac{h_{1,2}(y_1, y_2)}{g_{1,2}(y_1, y_2)}$. Since the conditional joint densities are given by $g_{1,2}(y_1, y_2) = 2g(y_1)g(y_2)$ and $h_{1,2}(y_1, y_2) = 2h(y_1)h(y_2)$, $z_{1,2}(y_1, y_2) = z(y_1)z(y_2)$, where $z(x) = \frac{h(x)}{g(x)}$ and equation (9) can be restated as

$$\frac{1}{u'(w(y_1, y_2))} = \lambda + \gamma [1 - z(y_1)z(y_2)] \quad (10)$$

Intuitively, z is a measure of the probability that y_i occurs as a result of high effort relatively to the probability that the same outcome is a result of low effort, for $i = \{1, 2\}$. In order to induce high effort, the principal takes advantage of this information by writing payments conditional on z . By assumption 2, $z(x)$ is non-monotonic and has at least one peak.⁴ For the rest of the analysis, I consider the simplest case in which z is single-peaked.

Definition Two random variables with densities $h(x)$, $g(x)$ and support Ω satisfy the Unimodal Likelihood Ratio (ULR) order, and we write $H \succ_{ULR} G$, if the likelihood ratio $z(x) = \frac{h(x)}{g(x)}$ is single-peaked at $x = \hat{x}$.⁵

Example 1 (*Spreads of Normal Distributions*): Suppose $A \sim N(\mu_A, \sigma_A^2)$ and $B \sim N(\mu_B, \sigma_B^2)$ with $\sigma_B^2 > \sigma_A^2$ and respective CDF's G and H . Then $H \succ_{ULR} G$.

Proof: See appendix.

Proposition 2 If the likelihood ratio satisfies the ULR order, the optimal transfers $w(y_1, y_2)$ are U-shaped in (y_1, y_2) and minimized at $(\hat{x}, \hat{x})$.

The proposition follows directly from (10) and the ULR assumption.

Intuitively, the agent is penalized when the extreme consumers attracted by the invention are close to $\hat{x}$ in order to deter low effort. On the other hand, the agent receives more generous payments when attracting extreme consumers located further from $\hat{x}$ since such outcomes serve as stronger evidence of high effort.

Corollary (*Locally decreasing transfers in market size*): The optimal contract that induces the level of effort that maximizes expected market size specifies transfers that not are always increasing in market size $d = y_1 - y_2$. In particular if $y_2 > \hat{x}$ then $\frac{\partial w}{\partial y_2} < 0$ and if $y_1 < \hat{x}$, $\frac{\partial w}{\partial y_1} < 0$.

⁴If $z(x)$ is monotonic then FOSD would hold, violating A2.

⁵See also Hopkins and Kornienko, 2003.

Proof: The proof of the first part follows by contradiction. Suppose that transfers are always increasing in market size and consider an idea $I(y'_1, y'_2)$ that attracts a subset of the population relatively to the idea $I(y_1, y_2)$. That is, $y'_1 \leq y_1$ and $y_2 \leq y'_2$ with at least one inequality being strict. Suppose that $\hat{x} < y_2 < y'_2$ and $y_1 = y'_1$. By assumption $w(y'_1, y'_2) < w(y_1, y_2)$ since $|y'_1 - y'_2| < |y_1 - y_2|$ but by proposition 2 $w(y_1, y_2) < w(y'_1, y'_2)$, a contradiction. Therefore transfers are not always increasing in market size. The second part follows by showing that the above claim holds for all $\hat{x} < y_2 < y'_2$ and $y_1 = y'_1$, and that in the symmetric case where $y'_1 < y_1 < \hat{x}$ and $y_2 = y'_2$ we have $w(y_1, y_2) < w(y'_1, y'_2)$ for all $y'_1 < y_1 < \hat{x}$. ■

Figure 2 illustrates an example. Note that while $I(y'_1, \bar{y}_2)$ attracts a subset of the consumers attracted by $I(y''_1, \bar{y}_2)$, the former idea receives a higher payment from the principal.

Figure 2: Locally decreasing transfers in market size, example.

As in the standard moral hazard problem, the optimal transfers are monotonic in the distance $d(x) = g(x) - h(x)$, but by A2 $d(x)$ is non-monotonic in x . In this scenario, market size and the principal's profits are still increasing in the agents effort and there is no indication that MLRP is violated with respect to the conditional distribution of market size. Suppose that the principal can't observe y_1 or y_2 but market size d is verifiable *ex-post*. If casual observation has led to the common understanding that sales are increasing in effort, it is natural to provide transfers increasing in sales to induce high effort. Now suppose that the principal can observe both y_1 and y_2 . If casual observation suggests that given extreme consumers are more likely to be attracted under high effort then the principal will start adjusting the contract by allowing it to depend on the values of y_1 and y_2 , in addition to d . While both mechanisms induce high effort, the second one is Pareto improving. The welfare gains can be thought in the following way: First, the principal now has the ability

punish the agent when the extreme consumers are close to the median consumer as it is a signal of low effort, and therefore can induce high effort at a lower expected cost. In addition, the agent can be seen as receiving some insurance for when small size market outcomes occur under high effort, and he can be made as least as well-off as before accepting a cut in expected payments, by receiving insurance against unfavorable outcomes (i.e. small market size) when he exerts high effort.

References

- [1] Hölmstrom, Bengt. "Moral hazard and observability." The Bell journal of economics (1979): 74-91.
- [2] Hopkins, Ed, and Tatiana Kornienko. "Ratio orderings and comparative statics." Working paper, University of Edinburgh (2003).
- [3] Johnson, Justin, P., and David P. Myatt. "On the Simple Economics of Advertising, Marketing, and Product Design." American Economic Review (2006), 96(3): 756-784.
- [4] Milgrom, Paul R. "Good news and bad news: Representation theorems and applications." The Bell Journal of Economics (1981): 380-391.

3 Appendix

Proof of Example 1: It suffices to show that $\log z(x)$ is concave in Ω . We have

$$\log z(x) = \log \left(\frac{h(x)}{g(x)} \right) = \log h(x) - \log g(x) \quad (11)$$

$$= \log \left(\frac{e^{-\frac{(x-\mu_A)^2}{2\sigma_A^2}}}{\sigma_A \sqrt{2\pi}} \right) - \log \left(\frac{e^{-\frac{(x-\mu_B)^2}{2\sigma_B^2}}}{\sigma_B \sqrt{2\pi}} \right) \quad (12)$$

and

$$\frac{d^2}{dx^2} \log z(x) = \frac{1}{\sigma_B^2} - \frac{1}{\sigma_A^2}. \quad (13)$$

Since $\sigma_B^2 > \sigma_A^2$, expression (13) is negative and $\log z(x)$ is concave in Ω .