Payment Mechanisms in the Healthcare Industry: An Experimental Study of Physician Incentives in a Multiple Principal Agent Setting

By

Ellen P. Green

WORKING PAPER SERIES
The views expressed in the Working Paper Series are those of the author(s) and do not necessarily reflect those of the Department of Economics or of the University of Delaware. Working Papers have not undergone any formal review and approval and are circulated for discussion purposes only and should not be quoted without permission. Your comments and suggestions are welcome and should be directed to the corresponding author. Copyright belongs to the author(s).
Current failures in the healthcare industry emphasize the need for a more fundamental understanding of how these contracts incentivize doctors. To aid this understanding, we treat the established physician-client-employer relationship as a multiple principal agent problem. We use a laboratory experiment, with a real-effort task, to test the relative performance of common payment mechanisms employed in this dual-principal agent relationship (Piece Rate, Flat Rate, Salary, Bonus, and Socialization). This study suggests, contrary to standard contract theory, that relying on extrinsic incentives to motivate physicians may be detrimental and costly for the healthcare industry.

* Green: University of Delaware, 402 Purnell Hall, Newark, DE 19711 (e-mail: epgreen@udel.edu) I would like to thank the Institute for Advanced Study at Virginia Tech for financial support. Also, I would like to thank Benjamin Congleton and Russell Osborne for their help with the program design and Ryan Donnelly for creating the essays used within the experimental task.

Currently, the increase in spending on health care in the United States continues to outpace the increase in the nation’s gross domestic product, which threatens the sustainability of the healthcare system. Approximately 52% of health care spending can be attributed to hospital care and physician or clinical services (CMS, 2011). Therefore, considerable savings would be gained if incentive systems could reduce the extent of unnecessary (and harmful) services; however, little research has measured the extent to which these unnecessary services are being provided (Korenstien et al., 2012).

Several years ago the health care industry began to move away from reimbursing physicians through simple fee-for-service and salary contracts towards pay-for-performance (P4P) contracts and other contract designs that micromanage physician behavior in order to reduce the costs of health care. However, these contract designs have failed to show any evidence of improving physician performance or reducing costs (Rosenthal et al. 2006). Moreover, these policies were implemented directly within health
care organizations and evaluated directly in the field where it is more expensive to test alternatives and hard to understand all factors at play. In short, the predictions failed and it is difficult to know why. In this study, by using laboratory experiments as a preliminary test, we are able to control variables that are impossible to control in the field and therefore reach a more definite conclusion about the consequences of alternative contract designs prior to an expensive intervention in the field (Riedl, 2009).

Here we create an environment reflective of the multiple principal agent relationship in the health care industry in which the agent (the physician) has responsibilities to two different principals, the downstream principal (i.e. the patient) and the upstream principal (i.e. the patient’s healthcare provider). In the experiment, subjects representing agents are paid to complete a real effort task in which their behavior directly impacts the downstream principal (the client) in this dual-principal agent relationship. This allows us to directly observe each payment structure’s impact on agent behavior and the resulting quality of care for the downstream principal. This experiment provides us with insight into potential improvements in contract design based on laboratory evidence. Furthermore, this experiment allows us to test the predictability of classic contract theory.

In theory the ideal agent contract must discourage the overprovision of services (i.e. services with no benefit or the harm outweighs the benefit), the misuse of services (e.g. provision of unnecessary care) and the under-provision of services (e.g. services not performed that were necessary) by creating a link between payment incentives and the contract design all while balancing risks across actors. This study shows that it is not necessary to create a direct link between the actions of agents and their compensation for services to achieve a high quality outcome.

Services provided by agents paid by the salary contract, under which agents are paid a standard rate independent of their actions, excelled in terms of quality of care provided for the downstream principal. More specifically, the agents contracted by salary provided unnecessary services only 8.5% of the time, while the agents contracted by a piece rate with bonus provided unnecessary services 72.1% of the time. If we apply this result to the average current spending per capita on health care, $7,538 per year (OECD,

---

1 Agents paid via the piece rate with bonus payment mechanism were paid per service they provided and additionally were given a bonus for providing a predetermined number of services identified as quota services.
2010), to predict the costs of unnecessary or harmful services, we find that agents paid under the piece rate with bonus payment mechanism provide approximately $4,949.93 worth of unnecessary or harmful services and only $2,588.07 of beneficial services.\(^2\) In comparison, agents paid under the salary payment mechanism are predicted to provide only $640.73 worth of harmful or unnecessary services and $6,897.27 worth of necessary services. These numbers, of course, provide only a rough indication of the money wasted based on one of the predominantly used payment structures, but they are suggestive.

The remainder of the paper is organized as follows: Section 2 provides a literature review of contract design and its implications for economic theory, experiments, and the field; Section 3 describes the experimental design and procedure; Section 4 details the results of the study, and Section 5 gives our conclusions.

II. LITERATURE REVIEW

A. AGENCY THEORY

In the healthcare industry, there is an interior optimum for the ideal number of services provided to a client, and deviating from this quantity is detrimental to the client (Korenstein et al. 2012). Because of imperfect information, selfish behavior, and costly monitoring (Harris et al. 1976, Hart et al. 1987), contracts that perfectly encourage optimal service levels are nearly impossible to design, and a solution to this problem has yet to be determined. However, there has been much discussion in theory literature about the problems with current healthcare payment mechanisms: the flat rate, the salary, and the piece rate.\(^3\)

The flat rate payment mechanism pays physicians a predetermined rate for each patient in their patient panel. Hence, the flat rate encourages physicians to treat as many patients as possible to increase their income (Robinson 2001). The flat rate also encourages “cream skimming” of patients, i.e., selecting the healthiest patients and passing on the more difficult patients. This is a result of the flat rate not adjusting for the additional time and effort required for the sickest patients (Matsaganis et al. 1994, 

---

\(^2\) This number is based only on the average total spending per capita in the United States, and all physicians are not paid under the piece rate with bonus payment structure.

\(^3\) In the health care literature the piece rate payment mechanism is commonly referred to the Fee-For-Service and the flat rate is referred to as capitation. To preserve the consistency in the text I will refer to them throughout as the flat rate mechanism and piece rate payment mechanism.
Dranove et al. 2003) and therefore, rather than bear the higher cost of sicker patients, physicians select the healthier and therefore more profitable patients.

The salary payment mechanism creates no direct link between physician effort and payment. As stated above, the salary payment mechanism pays a predetermined rate to the physician regardless of the number services provided and the number of patients seen. Therefore, in theory, physicians under this payment mechanism should conduct no services or as few as possible to remain employed. To address this, many industries that hire agents under the salary payment mechanism use other methods to encourage better care, such as social incentives (e.g. peer review by the American Board of Internal Medicine), promotions, risk of termination (Bolton 2005) and other incentives to provide better quality of services (Robinson 2001).

The piece rate mechanism, pays physicians for each service provided, which theoretically encourages physicians to do all that is necessary to improve a patient’s wellbeing. However, it also encourages physicians to treat patients for ailments that they do not have. Many economists agree that the piece rate payment mechanism leads to Physician Induced Demand, defined by Evans (1974) as a shift in a patient’s demand curve that is induced by a physician and benefits the physician (Rice 1983, Cromwell et al. 1986, Dranove et al. 1994). Physicians paid by the piece rate payment mechanism have a financial incentive to increase the demand for services by prescribing or recommending procedures irrespective of the cost or the patient’s wellbeing.

Often times the piece rate, flat rate, and salary payment mechanisms are paired with other incentives (e.g. pay for performance bonuses) to further motivate behavior. Pay for performance bonuses are frequently seen in the healthcare industry as a reward for physicians meeting quotas or target levels for specific procedures as determined by the United States Preventative Services Task Force or the National Committee for Quality Assurance. This type of bonus incentivizes agents to overinvest their time in the particular services that are rewarded by their completion. This leads to physicians placing less effort on other services (Peterson et al. 2006, Rosenthal et al. 2006).

Beyond the monetary incentives within a contract, decision-making is often dependent on other-regarding behavior, i.e. when the agent’s utility is dependent on the utility of a client. In these situations, an increase in a client’s utility will increase the
agent’s utility. Hideshi (2004), Dur et al. (2008), Kragl et al. (2009), and Englmaier et al. (2003) have each developed models in which they address the impact of the agent’s inequity aversion (a form of other-regarding behavior) on optimal contract design. These models agree that incorporating other-regarding behavior into contracts will alleviate part of the responsibility of the contract designer to promote ideal behavior through incentives. This is because the agent’s utility is directly impacted by the change in utility of the client (either negatively or positively) as a result of the agent’s behavior. Therefore, if the agent provides services that positively impact the client’s utility, the agent’s utility will increase as well, lessening the need for a contract to solve for the problems associated with imperfect information. The vast number of economic experiments that demonstrate that individuals are not purely self-interested further supports these theory models.

**B. Experimental Analysis**

As shown in many economic experiments, individuals are motivated by factors other than monetary gain. In 1972, Stahl showed that individuals in a dictator game value fairness in addition to their own monetary gain. Further studies, including Guth et al. (1982) and Roth et al. (1991), have used the ultimatum game to find that players were willing to forego a nontrivial amount of money to achieve fairness or equity in the allocation of resources. Forsythe et al. (1994) then compares the results of a dictator game and an ultimatum game and showed that people’s inclinations to distribute wealth are better explained by altruism than by inequity aversion. Although the precise motivation (e.g. altruism, inequity aversion, concern for fairness, etc.) behind this behavior is unclear, ignoring the effect of other-regarding behavior on the decision-making process results in a failure to predict human behavior accurately. These findings have encouraged several new areas of research in experimental economics, each further exploring the impact of such behavior on economic scenarios.

More recent experiments show that other-regarding behavior has a significant impact on how contracts are implemented in a principal-agent relationship. In Fehr et al. (2007) principals who acknowledged the altruistic nature of their agent when constructing their contracts achieved higher payoffs than principals who modeled their
contracts on the assumption that their agents were purely self-interested. Furthermore, Charness (2010) found that an agent’s productivity increased when the agent had information regarding her/his relative productivity in a group. In another study addressing agent position and agent effort, Hoffman et al. (1994), found that individuals behave in an increasingly selfish manner when they earn the position of the dictator in the dictator game. Their results indicate that the context of a contract alters the outcome and therefore needs to be taken into account for optimal contract design.

In light of these laboratory experiments, a field experiment was conducted in British Columbia, Canada, which explored the productivity effects of two payment mechanisms: piece rate (agents were paid a rate for each seed planted) and flat rate (agents were paid a daily wage) on workers in a tree-planting firm (Shearer, 2004). Here the author found that the piece rate mechanism resulted in a 38.11% increase in the planting of seed when compared to a flat rate payment mechanism. However, an increased rate of seed planting does not necessarily increase the quality of production. In this study, there are issues regarding the quality of the product under the piece rate payment mechanism. In particular, some locations where seedlings could be planted were desired due to ground quality and others were not. Therefore, an increase in seedlings planted did not always imply an increase in the quality of production, so that one cannot conclude that the piece rate mechanism outperformed the flat rate mechanism in quality, which is the question of interest to the contract designer.

C. EMPIRICAL STUDIES IN THE FIELD

Empirical studies in the field often provide inconsistent conclusions about the relative efficiency of alternative payment mechanisms in the dual-principal agent relationship. This is due to the difficulties that often arise in the form of exogenous variables. Demographics, timing, and procedure all have a significant impact on agent behavior and can be impossible to control in the field without interrupting business practices. However, there is still a significant amount of research on the dual-principal agent relationship in the healthcare industry, and the conclusions of this research are noteworthy.
In papers by Davidson et al. (1992) and Hohlen et al. (1990) the authors compared the behavior of physicians compensated under piece rate versus flat rate payment mechanisms. Both studies found no indication that physicians paid under a flat rate payment mechanism provided fewer services than those under the piece rate mechanism, counter to predictions by classic economic theory. However, Hellinger (1996), Hillman et al. (1992), and Gold Nelson et al. (1995) found a significant impact on physician behavior caused by payment mechanism. Their studies found that classic economic theory accurately predicts the actions of physicians induced by various payment mechanisms (i.e. piece rate physicians provided more services than salary physicians did). While there are significant implications from these articles, their authors stipulate that their own findings are subject to bias, and that in order to isolate the impact of the payment mechanisms one must control for variables that are nearly impossible to control in the field without hindering the practices’ activities.

In a literature review by Gosden et al. (2003), the authors report disparities in treatment quantity among the piece rate, flat rate, and salary payment mechanisms. However, this finding was not consistent throughout the literature as discussed above. Gosden noted that the qualifying studies that found no significant disparities in treatment quantity had failed to control for possible demographic differences among the treated populations. This failure can potentially change the conclusions of the studies and thereby is a potential explanation for the lack of differences (Kakwani et al. 1997). Therefore, the results of these studies are not robust enough to justify a significant change in payment policies before more thorough and controlled research is undertaken.

III. Data and Methodology:

Using a laboratory experiment, this paper explores [1] the behavior of agents in the dual-principal agent relationship and [2] how the success of a payment mechanism’s incentive structure is influenced by other-regarding behavior. For clarity, the following language is used to describe the actors: the Employer is the upstream principal who pays the agent and is also the contract designer, the Client is the downstream principal who receives the services of the agent, and the Agent is the actor who performs the services.
The “Dual-Principal Agent game” differs significantly from the experimental design employed by Fehr et al. in their 2002 and 2007 studies of the standard principal-agent relationship. In the Fehr et al. studies, principals earned a profit equal to the ‘work product’ created by the agent minus the agent’s salary, and the agents earned a profit equal to the salary less the cost of the effort. In their studies, the value of work produced and the cost of effort were numbers read from a table that the subjects selected, requiring no real effort to be exerted. What is different in the “Dual-Principal Agent game” is that agent effort is generated by the actions of agents, meaning that the subjects perform a real-effort task, and their actions in this task determine their final payments. Furthermore, this game does not attempt to quantify the cost of effort to the agents, which is consistent with real work situations, where the employers cannot quantify the costs to their employees (agents) of performing services for clients. However, we do assume that when an agent performs a service, the benefit to the agent outweighs the cost.4

In other respects, the framework of the “Dual-Principal Agent game” is similar to that of the traditional, principal-agent game. As in Fehr’s studies, the employer or (upstream) principal and the agent in the “Dual-Principal Agent game” interact in a one-time application of their contract. The fact that there is only a single interaction eliminates the influences of agent reputation, promotions, and/or the possibility of contract termination on the outcome of the contract. Additionally, the principals have no authority over which agents they are paired with, eliminating the possibility for the principals to employ any screening mechanisms in their efforts to improve outcomes. Furthermore, agents are not given the opportunity to select their secondary principals (employers) based on the payment mechanism; this eliminates the possibility of adverse selection. Thus this game allows us to focus exclusively on the effects of the different payment mechanisms on agent behavior.

A. EXPERIMENTAL DESIGN:

Each participant in the “Dual-Principal Agent game” represents a role in the dual-principal agent relationship. The experimenters function as the employers (upstream

---

4 Time was limited in the agent’s real effort task and therefore, we cannot conclude that the converse is true.
principals) by assigning the payment mechanisms and providing payment to the agents. Each subject acts as either the agent or the client in the relationship. Here I specifically study the relationship between the clients and the agent with respect to other-regarding behavior. I assume that there is no other-regarding behavior towards the experimenter based on the results of Frank (1998).

The “Dual-Principal Agent game” takes place in two Phases. In Phase I, each of 10 clients is endowed with $25.00 and asked to complete a proofreading task for 10 essays. Each of the 10 essays has 10 spelling errors that the clients are asked to correct. For each error that the client fails to correct a penalty of $0.25 USD is taken from the client’s endowment.\(^5\) The number of errors remaining represent represents the ideal number of services that the client can receive from an agent.

In Phase II of the game, each subject representing an agent provides proofreading assistance to the ten clients from Phase I of the game. A different subject (client) from Phase I of the game had edited each essay presented to the subjects in Phase II of the game. In this way, the Phase II subjects could potentially assist 10 distinct subjects from Phase I. The agent’s task is to correct the errors that were missed by the clients in Phase I. To assist the agents in the proofreading task, each error is highlighted in yellow, just as a practicing physician would be more capable of determining an ailment, the agents in the experiment have an advantage in finding errors. However, for each error that was missed by a client in Phase I of the game, the agent sees three yellow highlighted words within the essay. Only one highlighted word is the error missed by the client; the other two highlighted words are included to permit overprovision of services to clients. Each change by an agent to a highlighted word counts as a service provided.\(^6\) In each session the agents are assigned to a different payment mechanism, as determined by their employer (the experimenter) prior to the experimental session. At the end of their task, the agents are paid for the changes made, according to the conditions of the agent’s particular payment mechanism.

In the “Dual-Principal Agent game” the payment mechanism is the treatment variable. The treatments are piece rate, piece rate with bonus, flat rate, flat rate with

---

\(^5\) Note that if the subjects are unable to identify any of the errors, they receive no payment.

\(^6\) A change to the text did not have to be correct for an agent to be paid for services provided.
bonus, socialization, and salary. Piece rate agents are paid a set rate for each error they purported to correct, flat rate agents are paid a set rate for each client they assisted, and Salary agents are paid a set rate for participating in the experiment (refer to Table I for rates). The other three treatments employ one of the above payment mechanisms as a base and add additional incentives to create a final payment mechanism. The Socialization payment mechanism employs the flat rate payment mechanism with the added condition that throughout the proofreading task, a table with the number of changes to the text made by other agents is available on the screen. Thus we observe the effect of fellow agent behaviors in the experiment on an agent’s actions.

In flat rate with bonus, and piece rate with bonus, the agents are given a bonus if they are able to meet a predetermined service quota. In the game, the quota services are identified in the text by highlighting specific errors in the text in a unique color, green instead of yellow. In the healthcare industry physicians are often given a bonus if they provide a specified number of services such as mammograms or colonoscopies (Armour et al. 2004). In the experiment, the agents operating under the Piece Rate with Bonus and Flat Rate with Bonus mechanisms earn a bonus of $1.00 if they edit at least nine out of twenty four green errors.⁷ The agents do not earn the bonus for correcting any quantity of yellow errors. The bonus is added to the earnings as specified by the agent’s payment mechanism (i.e. piece rate or flat rate).

<table>
<thead>
<tr>
<th>Table I: Payment Mechanisms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Salary</td>
</tr>
<tr>
<td>Flat Rate</td>
</tr>
<tr>
<td>Flat Rate with Bonus</td>
</tr>
<tr>
<td>Piece Rate</td>
</tr>
<tr>
<td>Piece Rate with Bonus</td>
</tr>
<tr>
<td>Socialization (Flat Rate Base)</td>
</tr>
</tbody>
</table>

⁷ Only nine of the twenty-four quota errors were actual errors.
The agents aid the clients in the game as they would in a traditional setting of a dual-principal agent relationship. For each client error that an agent correctly identifies, the client is given $0.15 of the $0.25 lost back after Phase II is completed. To reflect the cost to a client of being overprovided with services, each instance of an agent correcting a mistake in an incorrect manner decreases the client’s earnings by $0.05.\textsuperscript{8,9}

**B. Procedure:**

The experiments were conducted in the Economics Research Laboratory at Virginia Tech. Students were recruited to participate from Principles of Economics courses offered at Virginia Tech. A total of 136 subjects participated in the experiment. Several sessions were held for both Phase I and Phase II of the “Dual-Principal Agent game.” No subject participated in more than one session.

Phase I of the experiment was completed in the first 3 experimental sessions. In this phase, recruited subjects played the role of the client in the dual-principal agent relationship. In each session, as the participants entered the laboratory lobby they were asked to complete a consent form and wait patiently for further instruction. Once all the subjects arrived, the subjects were escorted into the laboratory’s classroom and asked to edit the 10 essays within a 50-minute timeframe. The subjects were spaced so that they could not view the other participants’ actions in the task. Subjects were free to leave prior to the allotted time if they were satisfied with their effort. Once the task was completed, subjects were paid individually according to their success in completing the task. At this time, the subjects were also given further information about how some of their endowment would be earned back through assistance from participants in Phase II of the experiment. If the subjects wished to collect this reimbursement, contact information was collected at this time. Between Phase I and Phase II of the experiment, the proofread essays from the subjects in Phase I were entered into a computer database to be employed in Phase II of the experiment.

\textsuperscript{8} Agents were not penalized for making incorrect changes to errors.
\textsuperscript{9} If the case that the agent did more harm by editing the clients paper than help, the client was not required to pay additional monies.
In Phase II of the game, different subjects were recruited to represent the agents in the dual-principal agent game. This phase was conducted over 13 sessions. In each of these sessions the subjects entered the laboratory lobby and were asked to fill out a consent form and wait for further instruction. Once all the subjects were present they were escorted into the computer laboratory and asked to sit at a terminal. When seated, the subjects were given instructions for their task. The details of the payment mechanism were provided to the subjects both verbally and in text in each session. At this time subjects were given the opportunity to ask questions about their task. After the instructions were complete and all questions had been answered, the agents were asked to complete a quiz on the details of their payment mechanism to ensure that they fully understood how their actions would affect both their payoff and the client’s payoff. The experiment continued only when all subjects were able to successfully answer the quiz questions.

Next, the subjects performed the proofreading task. The subjects selected any of the 10 essays assigned to them and were able to go back and forth between essays as they wished. The interface allowed Phase II subjects to determine how many and which of the
Phase I the subjects they would provide services for; it was possible for the agent to help none or all of the Phase I subjects if they wished. Each agent in Phase II was assigned the same client panel. The agents were informed that each client from Phase I would be assisted by only one of the agents in Phase II.

To prevent proofreading out of boredom, the subjects were able to leave the proofreading task early. After the subjects were finished with their proofreading tasks or they completed a brief survey that collected data about their trustworthiness, altruism, reciprocity, socioeconomic status, and how interesting they felt the task was (Robinson et. al 1991, Deci, 1971).10 The agents were then called out to the lobby and paid individually. Once the data from Phase II was collected the total reimbursement the participants in Phase I of the game qualified for was calculated. Then the clients were contacted via email to receive their additional earnings.

<table>
<thead>
<tr>
<th>Table II: Summary of Agent's Actions by Treatment Reported as Averages</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Data</td>
</tr>
<tr>
<td>----------</td>
</tr>
<tr>
<td>Harmful</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Beneficial</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Total</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

*Standard Deviation reported in parentheses

IV. Results

A. Agent Actions

A summary of the agents’ actions in the “Dual-Principal Agent game” is provided in Table II. The agents’ actions are assigned to one of two categories: beneficial or harmful. Harmful actions are the incorrect changes the agent makes to the essays; beneficial actions are the number of correct amendments to the text the agent makes. Total services is the sum of beneficial actions and harmful actions. In Table II, they are reported as averages over all agents and organized by payment mechanism. These groupings provide us with an impression of the behavior encouraged by each payment mechanism.

10 The altruism, trustworthiness, reciprocity, socioeconomic, task-interest data collected was used in regression models to test the influence on agent behavior; however, none of these variables proved to have a significant impact on agent behavior.
The first result shown in Table II is that the relative number of total services provided by each payment mechanism agrees with the predictions of classic economic theory (Robinson 2001). Agents paid by the retrospective payment mechanisms (the piece rate with bonus and the piece rate) provided the most services when compared with the four prospective payment mechanisms (salary, flat rate, flat rate with bonus and socialization), all of which led agents to provide fewer services. 

Second, we see in Table II that agents paid by the salary payment mechanism provided the greatest number of beneficial services on average and piece rate with bonus provided the fewest number of beneficial services on average. This result can also be seen in Figure 1a. We also see that the flat rate and piece rate payment mechanisms provided significantly more beneficial services than the piece rate with bonus and socialization. The salary payment mechanism also provided significantly more services than the socialization payment mechanism.

---

11 A Mann Whitney U-test on the various payment mechanisms reveals a highly significant difference between the mechanisms that provided the greatest number of services and the least number of services (all of the relevant p values were less than 0.10).
12 A Mann Whitney U-test comparing the Piece Rate with Bonus with the Salary resulted in a p-value of 0.02.
13 A Mann Whitney U-test comparing the Flat Rate with the Piece Rate with Bonus resulted in a p-value of 0.02. A Mann Whitney U-test comparing the Piece Rate with the Piece Rate with Bonus resulted in a p-value of 0.05.
14 A Mann Whitney U-test comparing the Flat Rate with the Socialization resulted in a p-value of 0.02. A Mann Whitney U-test comparing the Piece Rate with Socialization resulted in a p-value of 0.07.
15 A Mann Whitney U-test comparing the Salary with the Socialization resulted in a p-value of 0.04.
Using the Mann Whitney U-test, we find the piece rate with bonus resulting in the most harmful services. The second most harmful payment mechanism is the piece rate. The agents paid via the piece rate payment structure provided a significant number of more harmful services to the client than agents paid via the salary, socialization, flat rate, and flat rate with bonus payment mechanisms.

![Table III: Summary of Agent's Impact on Client by Treatment](image)

<table>
<thead>
<tr>
<th>Table III: Summary of Agent's Impact on Client by Treatment</th>
<th>All Data</th>
<th>Piece Rate</th>
<th>Piece Rate with Bonus</th>
<th>Flat Rate</th>
<th>Flat Rate with Bonus</th>
<th>Salary</th>
<th>Socialization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Negative Impact</td>
<td>$0.85</td>
<td>$1.02</td>
<td>$3.45</td>
<td>$0.17</td>
<td>$0.21</td>
<td>$0.15</td>
<td>$0.18</td>
</tr>
<tr>
<td></td>
<td>$(1.69)</td>
<td>$(1.63)</td>
<td>$(2.51)</td>
<td>$(0.17)</td>
<td>$(0.14)</td>
<td>$(0.09)</td>
<td>$(0.11)</td>
</tr>
<tr>
<td>Average Positive Impact</td>
<td>$4.56</td>
<td>$4.95</td>
<td>$4.01</td>
<td>$4.91</td>
<td>$4.39</td>
<td>$4.96</td>
<td>$4.14</td>
</tr>
<tr>
<td></td>
<td>$(1.49)</td>
<td>$(1.62)</td>
<td>$(1.64)</td>
<td>$(4.91)</td>
<td>$(1.75)</td>
<td>$(1.17)</td>
<td>$(1.39)</td>
</tr>
<tr>
<td>Net impact</td>
<td>$3.71</td>
<td>$3.93</td>
<td>$0.55</td>
<td>$4.74</td>
<td>$4.17</td>
<td>$4.80</td>
<td>$3.96</td>
</tr>
<tr>
<td></td>
<td>$(2.49)</td>
<td>$(2.31)</td>
<td>$(0.55)</td>
<td>$(1.21)</td>
<td>$(4.17)</td>
<td>$(1.18)</td>
<td>$(1.42)</td>
</tr>
</tbody>
</table>

*Standard Deviation reported in parentheses

**B. IMPACT ON CLIENT**

The impact of the agent’s actions on the client is of the greatest concern from both the point of view of the client and policy makers. Ideally, each of the agent’s actions positively impacts the client’s outcome from both the client’s perspective and a cost perspective. The Agents’ behavior was tracked by whether their actions were harmful or beneficial. Impact on the client is calculated by adding +0.15 for each correctly identified error (beneficial actions) and -0.05 for each incorrectly identified error (harmful actions). Client impact measures the effect on the client’s income as a result of the agent’s actions in the game. This is reported in Table III as the average impact the agent had on the client’s payoff by treatment type. Here, we see that the salary and flat rate resulted in the highest (best) net impact rating, with averages of $4.80 and $4.74 respectively. Using the Mann-Whitney U-test, we see that the flat rate’s net impact was significantly greater than

---

16 P-values of Mann Whitney U-test when compared with Piece Rate with Bonus: Flat Rate (0.00), Flat Rate with Bonus (0.00), Piece Rate (0.00), Salary (0.00), Socialization (0.00).

17 P-Values of Mann Whitney U-test when compared with Piece rate: Salary (p=0.01), Socialization (p=0.05), Flat Rate (p=0.01), and Flat Rate with Bonus (p=0.14)
the piece rate with bonus and socialization. The salary’s net impact was significantly greater than the piece rate with bonus and the socialization. The piece rate with bonus had an impact factor significantly less than the piece rate, socialization, and flat rate with bonus.

Table IV: Costs to employer by Payment Type

<table>
<thead>
<tr>
<th></th>
<th>All Data</th>
<th>Piece Rate</th>
<th>Piece Rate with Bonus</th>
<th>Flat Rate</th>
<th>Flat Rate with Bonus</th>
<th>Salary</th>
<th>Socialization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Costs</td>
<td>$21.33</td>
<td>$10.77</td>
<td>$20.18</td>
<td>$24.35</td>
<td>$22.81</td>
<td>$25.00</td>
<td>$23.50</td>
</tr>
<tr>
<td></td>
<td>($6.87)</td>
<td>($6.95)</td>
<td>($9.56)</td>
<td>($1.72)</td>
<td>($3.94)</td>
<td>-</td>
<td>($23.50)</td>
</tr>
<tr>
<td>Benefits</td>
<td>$4.56</td>
<td>$4.95</td>
<td>$4.01</td>
<td>$4.91</td>
<td>$4.39</td>
<td>$4.96</td>
<td>$4.14</td>
</tr>
<tr>
<td></td>
<td>($1.49)</td>
<td>($1.62)</td>
<td>($1.64)</td>
<td>($1.19)</td>
<td>($1.75)</td>
<td>($1.17)</td>
<td>($4.14)</td>
</tr>
<tr>
<td>Benefits/Cost</td>
<td>0.26</td>
<td>0.49</td>
<td>0.13</td>
<td>0.19</td>
<td>0.18285</td>
<td>0.19</td>
<td>0.17</td>
</tr>
<tr>
<td></td>
<td>($0.17)</td>
<td>($0.19)</td>
<td>($0.20)</td>
<td>($0.05)</td>
<td>($0.06)</td>
<td>$0.05</td>
<td>($0.17)</td>
</tr>
</tbody>
</table>

N 126 19 21 23 21 22 20

*Standard Deviation reported in parentheses

C. IMPACT ON EMPLOYER

Consider the cost to the employer for hiring the agent. Table IV reports the average cost to the employer for hiring an agent by treatment type. The highest cost to the employer was the salary ($25.00) payment mechanism, followed closely by the flat rate payment mechanisms ($24.35). This result is surprising given that the incentives provided in the piece rate and piece rate with bonus mechanisms create a direct relationship between the income of the agent and the number of services provided; therefore, allowing the agents to maximize their income (thereby increasing their cost to the employer) by changing as many errors as possible. However, the agents paid by these mechanisms did not conduct enough services to increase their cost to the point of

---

18 P-Values of Mann Whitney-U test when compared with the Flat Rate, the Piece Rate with Bonus (p=0.00) and Socialization (p=0.03)
19 P-Values of Mann Whitney-U test when compared with the Salary, the Piece Rate with Bonus (p=0.00) and Socialization (p=0.00)
20 P-Values of Mann Whitney-U test when compared with the Piece Rate with Bonus, the Piece Rate (p=0.00), Socialization(p=0.00), and Flat Rate with Bonus(p=0.00)
21 From these observations, we can comment on the quality of services provided to the client relative to the frequency of services. While in theoretical and field studies authors discuss the decrease in the quantity of services provided when the Flat Rate or Salary mechanisms are employed; they do not fully address the relationship between frequency and beneficial services or harmful services because it is difficult to estimate this value.
22 The Salary and Flat Rate payment mechanisms cost the employer significantly more than the Piece Rate and Piece Rate with bonus; the relevant p-values in comparison were all less than 0.01.
23 One of the difficulties in creating the wages for each payment mechanism was in trying to create values that did not directly affect agent behavior but still provided incentive to reveal the incentives of each payment mechanism. Therefore, I hesitate to compare only the cost estimates and think it is more efficient to focus on the cost/benefit analysis provided in the next section.
the other mechanisms. This result could potentially be a consequence of the time limit on the game.\textsuperscript{24}

Also shown in Table IV is the relationship between quality of product and the cost of employing the payment mechanism. Here, the benefit-to-cost ratio is the net difference of actions (the number of beneficial services – the number of harmful services) to the cost of employing the client.\textsuperscript{25} Although a significant amount of money was wasted on services provided in the piece rate payment mechanism, the number of beneficial services provided outweighs the waste resulting in the greatest benefit-to-cost ratio and the benefit-to-cost ratio was significantly higher than all other payment mechanisms. The expected payoff to the agent in the piece rate mechanism was significantly lower despite the poor quality of services;\textsuperscript{26} however, the cost-benefit analysis shows that this mechanism provided the highest benefit-to-cost ratio. Conversely, the piece rate with bonus continued to fare the worst and produced results with significantly lower quality in the game when costs are compared to benefits. Although the differences are not highly significant, the results indicate the following ranking in the cost-benefit analysis: piece rate, salary, flat rate, socialization, flat rate with bonus, and piece rate with bonus.\textsuperscript{27}

\textbf{V. Conclusions and Discussion}

The behavior observed by the agents in the “Dual-Principal Agent Game” contradicts the predictions of contract theories based on self-interested models of rational choice. While the findings were consistent with classic contract theory in terms of service frequency, (i.e. the piece rate and piece rate with bonus mechanisms provided a greater number of services than the flat rate, flat rate with bonus, salary, and socialization

\textsuperscript{24} The time constraint on the proofreading task could potentially have influenced this; however, there were subjects that did change almost all of the errors available for them to change in the given time frame.

\textsuperscript{25} It is important to note that the only cost to the employer in the industry setting is not entirely collected in the payment of the agent. Employers must pay fixed costs within their firm. Additionally, the healthcare provider/employer must purchase malpractice insurance. These additional costs can increase the cost of the overprovision of services, so the cost of the piece rate mechanism in this variable is perhaps underestimated.

\textsuperscript{26} P Values of Piece Rate vs other payment mechanisms: Flat Rate (0.00), Flat Rate with Bonus (0.00), Piece Rate with Bonus (0.00), Salary (0.00), and Socialization (0.00).

\textsuperscript{27} It is important to note that with each of the cost analyses, the cost to employer for the Salary could potentially be an artifact of the experiment and further analysis should be conducted to observe the impact of a change in Salary on the behavior of the clients.
classic contract theory was unable to predict the resulting quality of services provided by the payment mechanisms.

Payment mechanisms with a direct link between agent income and the number of services provided (i.e. piece rate and piece rate with bonus) resulted in the lowest overall quality of services. In contrast, the payment mechanisms that were not directly tied to the number of services provided (i.e. salary, flat rate, and flat rate with bonus) resulted in a higher overall quality of services. There are several theories within economics and psychology can explain why monetary incentives often fail (see Kamenica 2012 for a review); among those relevant to this experiment is the crowding out effect.

The crowding out effect occurs when monetary incentives ‘crowd out’ intrinsic incentives that would otherwise motivate subjects to complete a task at higher levels of performance. Several intrinsic motivations can be at play here, i.e. self-recognition (Deci and Ryan, 1985), confidence management (Bénabou et al., 2003), other regarding behavior (Hoffman et. al 1996), etc. (see Deci et al. 1999 for a review).

The crowding out effect was most apparent in the piece rate payment mechanism in the “Dual-Principal Agent Game”. The Piece Rate Mechanism paid agents on a per service basis, which places more focus on the maximization of income relative to the other payment mechanism. This is due to the piece rate’s relationship between behavior and reward. For instance, agents paid a salary could focus on their task, rather than maximizing their income. Piece rate agents maximized their income and increased the number of services provided, thereby increasing the number of unnecessary services provided due to extrinsic or monetary incentives. The ‘hidden cost’ of the extrinsic incentive is that it can override intrinsic motivations that lead to higher quality services. The retrospective monetary incentives muddle or crowd out these intrinsic incentives, whereas those under the prospective payment mechanisms were able to focus on their task.

Due to the constraints of the experimental design, we are unable to determine which of the intrinsic factors are at play; however, the end result is very interesting. The impact that these non-monetary incentives had on agent behavior supports the use of prospective payment mechanisms in a multiple principal agent relationship. The “Dual-Principal Agent Game” demonstrates that agent behavior is indeed impacted by payment
mechanisms, but not as straightforward as classic theory predicts. In some cases, intrinsic motivations were able to improve the quality of services provided by the agents when the payment incentives did not crowd them out. If we assume that the healthcare industry successfully recruits and produces altruistic physicians, as recommended by the Medical School Objectives Project (MSOP 1999), the influence of these intrinsic motivations should be escalated in the healthcare industry. Therefore motivating physicians with intrinsic motivations by removing complex payment structures, such as the piece rate, should be explored in future healthcare policy research.
REFERENCES:

Armour, Brain S.; Carol Friedman; M. Melinda Pitts; Jennifer Wike; Linda Alley and Jeff Etchason. 2004. “The Influence of Year-End Bonuses on Colorectal Cancer Screening.” American Journal of Managed Care 10(9), 617-24.


Charness, Gary; David Masclet; and Marie-Claire Villeval. 2010. “Competitive Preferences and Status as an Incentive: Experimental Evidence.” IZA Discussion Paper No. 5034


Evans, Robert. 1974. “Supplier-Induced Demand: Some Empirical Evidence And Implications.” The Economics of Health and Medical Care 162-173


The Medical School Objectives Writing Group. 1999. “Learning Objectives For Medical Student Education—Guidelines For Medical Schools: Report I Of The Medical School Objectives Project.” Academic Medicine 74: 13-8
APPENDIX I:

Introduction

This is an experiment in decision-making. You will have an opportunity to earn a considerable amount of cash through your participation in this experiment. You will be required to complete the experiment individually. You are not allowed to communicate with any other participant at any point during the course of the experiment. If you complete the experiment, you will be paid for the decisions you make in the task. You may withdraw from the experiment at any time. If you withdraw from the experiment before its completion, you will only be paid for the portion of the experiment that you have completed. If you do remain in the experiment, then you should feel free to try to make as much money as you can. All responses and decisions will be anonymous and the only piece of experimental material that will contain your identity will be your receipt of payment at the end of the experiment.

Before we begin, please set your cell phones to silent. We ask that you not make calls or send text messages until the experiment is complete. We also ask that you not talk to other participants in the experiment until after the experiment is complete.

Instructions

Basic Overview:
In this experiment you will be asked to proofread 10 short essays. The number of essays that you proofread and the number of corrections per essay that you complete is up to you. Your income earned in this experiment will be based on your performance in the task.

Proofreading Task:
In a previous session of this experiment we asked participants to proofread short essays for typographical and spelling errors. They were given 10 essays with 10 errors in each essay. These participants were initially given $25.00 for completing the proofreading task, but had $0.25 taken away for each error they were not able to correct.

For instance:
- If in Essay 1, the participant corrected 5 out of 10 errors, $0.25*5=$1.25 was taken away from their initial endowment.
- If in Essay 2 the participant corrected 0 out of 10 errors, $0.25*10=$2.50 was taken away from their initial endowment.
- A total of $3.75 was taken away from their initial endowment in this example.

Note that if the participant was unable to correct any errors, they received $0.00 for the experiment.

Your task in the experiment will be to proofread the essays that the first participants have already edited. The changes that you make to their essays will allow them to earn back a portion of their $25.00 that they lost. For each error that you correctly identify, we will reimburse the participant in the first group $0.15. However, if you make an incorrect change the participant in the first group will lose $0.05.

The mistakes that the first participants identified will no longer be in the essay and therefore you will not need to correct them.

The mistakes that the first group missed will still be in the essay. Your task is to proofread these essays for the errors that the first group missed. You will be given assistance with your editing task. For each error in the essay, we have highlighted 3 words, only one of which is the actual error. The first group was not given this assistance in their proofreading task. For example, if the person who initially proofread the essay
missed 3 errors you will see 9 highlighted words, only 3 of which contain actual errors. The total number of actual errors will be reported to you at the top of each essay.

This proofreading task allows each participant from the previous session to obtain help from one person in this part of the experiment for his or her essays.

In total you will be presented with 10 essays to edit, each from a different participant. You will have a total of 10 minutes to complete this task. At the end of the 10-minute period you will no longer be able to proofread essays.

**Proofreading Payment:**
You will be paid $0.20 for each of the highlighted words that you attempt to correct within the 10-minute time frame. You do not have to correctly change the typo to be paid, but the person whose essay you are editing will only be reimbursed for correct editing and will lose money if incorrect editing.

For instance if you make 9 changes to Essay 1 your payment will be calculated as follows:
- If you correctly identify 3 of the highlighted errors and incorrectly identify 6 of the highlighted errors:
  - You will receive $0.20 for all 9 changes and your total payment for Essay 1 will be $1.80
  - The participant whose essay you are editing will receive $0.15 for each of the 3 correctly identified errors (a total of $0.45) and lose $0.05 for the 6 incorrectly changed words (a total of $0.30), so their total reimbursement payment will be $0.45-$0.30=$0.15.

Note: Your total payment will be calculated by summing your changes over all 10 essays.

**Completion of Task:**
You will be given 10 minutes to complete the proofreading task, however if you are satisfied with the number of corrections that you have made you can leave the experiment by clicking the “Finish” button on the lower left-hand side of your screen before the 10 minutes are up. At this time the instructor will provide you with a brief survey to complete before you are paid and are free to leave.
APPENDIX II:
Quiz: You will now complete a quiz to see if you understand how your earnings will be calculated.

1. If you correctly identify 5 of the highlighted errors and incorrectly identify 1 of the highlighted errors in Essay 1:
   For your corrections in Essay 1 you receive _____
   The participant whose essay you are editing will receive _____

2. If you correctly identify 1 of the highlighted errors and incorrectly identify 5 of the highlighted errors in Essay 2:
   For your corrections in Essay 2 you receive _____
   The participant whose essay you are editing will receive _____

3. If you correctly identify 7 of the highlighted errors and incorrectly identify 2 of the highlighted errors in Essay 3:
   For your corrections in Essay 3 you receive _____
   The participant whose essay you are editing will receive _____

4. If you correctly identify 4 of the highlighted errors and incorrectly identify 4 of the highlighted errors in Essay 4:
   For your corrections in Essay 4 you receive _____
   The participant whose essay you are editing will receive _____

5. If you correctly identify 4 of the highlighted errors and incorrectly identify 0 of the highlighted errors in Essay 5:
   For your corrections in Essay 5 you receive _____
   The participant whose essay you are editing will receive _____

6. If you correctly identify 0 of the highlighted errors and incorrectly identify 0 of the highlighted errors in Essay 6:
   For your corrections in Essay 6 you receive _____
   The participant whose essay you are editing will receive _____

7. If you correctly identify 9 of the highlighted errors and incorrectly identify 5 of the highlighted errors in Essay 7:
   For your corrections in Essay 7 you receive _____
   The participant whose essay you are editing will receive _____

8. If you correctly identify 1 of the highlighted errors and incorrectly identify 6 of the highlighted errors in Essay 8:
   For your corrections in Essay 8 you receive _____
   The participant whose essay you are editing will receive _____

9. If you correctly identify 3 of the highlighted errors and incorrectly identify 1 of the highlighted errors in Essay 9:
   For your corrections in Essay 9 you receive _____
   The participant whose essay you are editing will receive _____

10. If you correctly identify 0 of the highlighted errors and incorrectly identify 6 of the highlighted errors in Essay 10:
    For your corrections in Essay 10 you receive _____
    The participant whose essay you are editing will receive _____

11. Please calculate your total earnings from correcting these ten essays (questions #1-10)
    Total Earnings _____
APPENDIX III:

Instructions
This is an experiment in decision-making. You will have an opportunity to earn a considerable amount of cash through your participation in this experiment. You will be required to complete the experiment individually. You are not allowed to communicate with any other participant at any point during the course of the experiment. If you complete the experiment, you will be paid for the decisions you make in the task. You may withdraw from the experiment at any time. If you withdraw from the experiment before its completion, you will only be paid for the portion of the experiment that you have completed. If you do remain in the experiment, then you should feel free to try to make as much money as you can. All responses and decisions will be anonymous and the only piece of experimental material that will contain your identity will be your receipt of payment at the end of the experiment.

Basic Overview
In this experiment, you are asked to proofread several short essays for spelling and typographical errors. At the end of the experiment you will be paid for your performance in this proofreading task.

Earnings Task:
Each person starts the experiment with $25. You will be given 10 essays to proofread. Each essay contains 10 spelling and typographical errors. For each error you correctly identify there is no change to your earnings, however for each error in the essays that you fail to correct you will have $0.25 taken away from your $25.00. You will not be penalized for changes made to portions of the text that did not originally contain an error.

You will have a total of 50 minutes to complete the proofreading task. At the end of the 50-minute period, the instructor will collect the essays to check for correctness and calculate your earnings.

For example:
• If in Essay 1, if you correct 5 out of 10 errors, $1.25 was taken away from your $25.00.
• If in Essay 2 if you correct 0 out of 10 errors, $2.50 was taken away from your $25.00.
• A total of $3.75 would be taken away from your $25.00 in this example.

Note: total earnings will be calculated over all 10 essays.

Future Earnings:
In future sessions of this experiment, other participants will proofread your corrected essays for the spelling and typographical errors that you were unable to correct, these participants will not know whose essay they are proofreading and you will not know who proofread your essay. This will allow you to earn some of your money back. For each error that they find that you failed to correct, we will award you $0.15.

For instance:
• If you missed 10 errors and the participants in the future sessions found 2 of them, you will receive $0.30.
• For this reason it is important that we have an email address to contact you so that you collect these future earnings.