# Bureaucratic Responses

Canice Prendergast

November 2014

**Abstract**

This paper's interest is in understanding how bureaucracies should respond to their clients. I claim that because many services are not priced, bureaucratic responses to their clients will often be the opposite of the reactions of "normal" firms. Specifically, they ignore the most credible complaints from clients, provide the poorest service to clients to value the service most, and require most red tape for clients who already know their needs.

University of Chicago Booth School of Business. Email: canice.prendergast@chicagobooth.edu.
I am grateful to the editor and seminar participants at Stanford University and the Applied Theory Conference at the University of Chicago for helpful comments. The author declares that he has no relevant material or financial interests that relate to the research described in this paper.

People routinely tell exasperated stories of their interactions with bureaucrats. The source of such discontent often lies in their apparently non-sensical behavior: intransigence, digging their heels in, asking for ridiculous amounts of paperwork, and so on. This paper argues that this behavior may arise not for the usual reasons - poor performance measures and weak monetary incentives for bureaucrats[1] - but rather because many services offered by bureaucracies to their clients are not priced.[2] Specifically, I argue that because public agencies cannot charge for services in the way that other firms do, they often respond in unusual ways to their clients' needs and information.

I consider a setting where a bureaucracy provides a service to a client (or consumer). In the public sector, these are known as service agencies. The backdrop to this work is the role that consumers play in such settings. Consumers are needed for many reasons. First, they have to show up: the take-up of many public services is very low for example. Second, consumers play an important role in solving agency problems by, among other things, pointing out errors. Input from consumers is ubiquitous in the private sector: restaurant diners unhappy with their meals complain to management, people whose Fedex packages go astray alert the company, and so on. Yet they similarly play a key role in the allocation of public services. They provide documentation, advocate for their needs, and complain when decisions do not go their way. Indeed, much of the discussion on improving public agencies revolves around client advocacy. For example, the head of rationing for health care in Britain recently argued that "individual patients should become more knowledgeable about their health conditions, and tell their doctors if they believed they were missing out on treatment which could help them." and they should "be more pushy with their doctors about drugs to which they are entitled" (Daily Telegraph, 2014).

The focus of the paper is on how a bureaucracy responds to its clients. I focus on two "reasonable" predictions about such responses:

1. The more a client marginally values the bureaucracy's inputs, the more of those inputs are provided.

2. Bureaucracies should be more responsive to concerns of better informed clients than to those more likely to be erroneous.

These will both hold when monetary transfers are unrestricted. However, when transfers are

[1] See Dixit, 2002, for an excellent review of these issues.

[2] I do not address the issue of why many public services are not priced in the way that a profit maximizing firm would. The most natural reason why underpricing arises is that perhaps some clients "don't have the money" and that pricing excludes some deserving clients.

restricted - the distinguishing feature of a bureaucracy here - the opposite outcome often arises.

I consider two roles for consumer pricing. The first is that unrestricted transfers imply that client services are not distorted by the need to ensure that clients take up those services. Public agencies are often accused of treating their clients poorly. A simple answer for why is "because the can": client rents are so great that they will show up even with poor service. Poor service at agencies such as immigration departments likely reflects this issue.

Consider a setting where a client values higher quality service. When he pays for such quality, marginal improvements in quality are reflected in higher prices paid to the service provider. In that case, better service is provided to those who value it most on the margin. When service is not priced on the margin, what matters is whether the client is willing to show up for benefits at the firm's preferred choice of service. When participation is not a constraint, service is (intuitively) independent of the client's marginal valuation. However, when marginal quality choice affects participation, the outcome above is inverted. Clients who value service most need little of such service to participate, whereas those who care little need to be offered more. Hence the absence of pricing inverts prediction 1. above.

The second part of the paper addresses the effectiveness of using clients to solve agency problems, extending Prendergast, 2002, 2003. Consider a setting where a bureaucrat is charged with providing service to a client. Based on her information, she assigns a treatment. (For example, a doctor makes a diagnosis on a patient.) The bureaucrat's superiors typically learn little about these interactions without the arrival of further information. Such information often comes from a complaint by the client, which may overturn the decision or trigger further investigation.

However, investigations are a double edged sword. On the one hand, they fix error, yet on the other they may expose a bureaucrat's performance in a way that affects her career. I study this in a setting where the bureaucrat's wage can change based on the outcome of an investigation: it may decline if she made an error (perhaps she is fired), and may increase if she identified the correct course of action (she is promoted). I initially impose few assumptions on these wages. This allows for agency issues where either the bureaucrat wishes to avoid investigation, or alternatively craves oversight as it gives her a chance to show her talent.

There is potential misalignment of both the bureaucrat's and the client's interests here. Despite this, if the principal can freely choose a single price charged to the client for the service, the first best arises if efficient oversight is symmetric and the wage from being found wrong at investigation is no higher than the other wages. In this equilibrium, information is efficiently aggregated and investigations only launched when needed. Hence, with a very simple mechanism - a single price - efficiency is achieved. Two further intuitive outcomes arise here. First, surplus rises as clients become better informed. Second, better informed clients are more influential in outcomes. As such, a natural solution to many agency problems is to allow end users to have a voice.

The consumer's role here is to point out mistakes, and appropriate pricing makes his information credible. However, many public agencies do not charge for their services. This has direct and indirect effects. First, clients stay quiet when incorrectly given benefits. This in itself has the unsurprising outcome that the first best is no longer achievable as mistakes that benefit the client are corrected less often. However, our focus is on how this asymmetry interacts with bureaucrat incentives. As in Prendergast, 2002, the interaction derives from the tension that using clients both correct errors and threatens bureaucrats who worry about their errors being identified.[3] This latter concern implies that the bureaucrat may capitulate to the client by always giving him what he wants, even when she should not. By doing so, she may avoid investigation.

Overcoming this temptation to capitulate is called the "bureaucratic constraint" below. Unlike the minimal conditions on wages for the first best when prices are unrestricted, I show that for either the bureaucrat or the client to have *any* role in the allocative process when oversight is efficient, it must be that wages are weakly convex in the probability of making correct decisions. Inverted bureaucratic responses arise when the bureaucrat capitulates at efficient levels of oversight, i.e, when wages are sub-linear in probabilities. Specifically, oversight is distorted to assure the bureaucrat that her errors are unlikely to be punished. The distortions involve either denying legitimate complaints or biasing oversight against clients by over-monitoring. Remember that better informed clients both increase surplus and become more influential when prices are appropriately set. The opposite is true here: well informed clients are most threatening to a bureaucrat, and so their complaints are most likely to be ignored. Less credible complaints may garner a response, but sufficiently credible ones are ignored. Compared to the first best, the role of clients is inverted here. Clients who are sufficiently well informed play no role in the allocation, and so surplus can decrease as consumers become better informed. For reasonable compensation of bureaucrats, client complaints play no role. Hence bureaucratic responses look very different to the supposedly

[3] For example, see Skolnik, 1966, on police officers: "it is doubtful that norms are needed because policemen are lazy" but rather "springs from the reluctance of policemen to expose themselves to what they know to be public hostility" (p.55).

reasonable predictions above.

This part of this paper extends Prendergast, 2002, 2003, in a number of directions. First, the modeling here considers much more general preferences and strategies for all parties, which allows a better understanding of both the ease with which efficiency can be achieved with flexible pricing, and the difficulty in using clients when prices are restricted. These issues were masked in special assumptions in the previous literature. Second, the focus here is squarely on how better informed clients affect the alleviation of agency problems. The advent of the internet has surely led to better informed consumers. When consumers can be charged for services, the paper suggests these technological changes have alleviated agency problems. By contrast, bureaucratic efficiency may be harmed by better informed clients. Hence, this paper points to a possible divergence in efficiency across the two arenas.

I begin by considering participation issues in Section 1. Section 2 describes the role of prices as a means of revealing client information, and its impact on the incentives of bureaucrats. Section 3 addresses possible bureaucratic responses, and considers a number of extensions to the model. Section 4 concludes.

# 1 Participation and Prices

In this section, I consider the role that prices play in alleviating the need to distort service to ensure service take-up. Consider an agency required to provide a service to a client, where the client values higher levels of service. The focus here is on how service varies with how much the client marginally values it. I show that when prices are unrestricted, those who value service more will receive greater service, but with restrictions on transfers, the opposite will often be true.

An agency or principal is required to provide a good (or service) to a client. The quality of that good, $e$, is chosen by the principal at a cost of $C(e)$, where $C'(e) > 0$, $C''(e) > 0$, and $C(0) = 0$. Both the principal and the client potentially derive intrinsic value from the good, where the principal values the good at $ve$ and the client values it at $\beta e$, where both $v$ and $\beta$ are non-negative. The principal can commit to $e$, and the client has a reservation utility of $U^* \geq 0$ (think of this as the cost of navigating the application process).[4] The principal

[4]This has been couched in terms of the principal's constraint being to induce the client to "show up". More generally, the key ingredient is that the principal cares about the net utility that the client receives from the transaction. What is needed is that the principal feels he can reduce service when the client greatly values the good. One alternative would be where the client is reluctant to complain about service when rents are large in case he is cut off from that service as a result.

chooses a price $p(e)$ paid by the client for a good of quality $e$, where the principal's utility is $ve + p(e) - C(e)$ and the client's utility is given by $\beta e - p(e)$.

However, monetary transfers are assumed to be constrained: for example, many countries impose limits on how much can be charged for medical care. Specifically, there is a maximum transfer $x \geq 0$ that can be transferred from the client to the principal and a maximum transfer $y \geq 0$ that can be transferred from the principal to the client. The program of the principal is then to choose $p(e)$ and $e$ to maximize

$$ve + p(e) - C(e) \tag{1}$$

subject to $p(e) \in (-y, x)$, and $\beta e - p(e) \geq U^*$, where the participation condition for the client is the implication of the principal being required to provide service.

The timing of the model is as follows: the principal makes a take-it-or-leave it offer to the client, where he offers a quality $e$ at price $p(e)$, subject to $p(e) \in (-y, x)$. If the client accepts, that quality is produced: otherwise, the game ends with no production.

The issue of interest here is simple: how does the quality of service $e$ depend on the client's marginal valuation $\beta$? There are two useful benchmarks. The first maximizes the joint welfare of the principal and the client $(v + \beta)e - C(e)$, with optimum $e^*$ defined by $v + \beta = C'(e^*)$. In this case, $\frac{de^*}{d\beta} = \frac{1}{C''} > 0$. This result, of course, depends on the complementarity between $\beta$ and $e$. Let $p^* = \beta e^* - U^*$ be the transfer required for the client's participation constraint to bind at that level of service. The second only takes the principal's welfare into account, and is given by $e^{**}$ where $v = C'(e^{**})$. Not surprisingly, $e^* \geq e^{**}$ and $\frac{de^{**}}{d\beta} = 0$. With these two benchmarks, the outcome is described in Proposition 1.

**Proposition 1** *Assume that the principal is required to offer the service to the client. Then*

- *If $p^* \in (-x, y)$, quality is $e^*$ and increases in $\beta$.*
- *If $p^* \notin (-x, y)$, either quality is independent of $\beta$ and $e = e^{**}$, or is declining in $\beta$. A necessary and sufficient condition for quality to be declining in $\beta$ is if either*
    1. *$\beta e^{**} - x < U^* < \beta e^* - x$ and $\beta < \tilde{\beta} = \frac{U^* + x}{e^*}$, in which case $e \in (e^{**}, e^*)$, or*
    2. *$\beta e^* + y < U^*$, in which case $e > e^*$.*

First, when prices are not constrained, the principal can charge for marginal improvements in quality, and charges a price of $p = \beta e - U^*$. Hence, he maximizes $(v + \beta)e - C(e)$, with optimum $e^*$. Intuitively, $\frac{de^*}{d\beta} > 0$.

This outcome is feasible if $p^* \in (-y, x)$. If it is not feasible, then one of two conditions must hold: either the client is required to transfer too much money (for service $e^*$) to the principal $\beta e^* - x > U^*$, or the principal is required to transfer too much money to the client so that $\beta e^* + y < U^*$.

The second case is the simplest to see. Consider a public agency that has little private interest in providing a service ($v$ low) but is mandated to do so. Transfers are limited ($y$ low). Its incentive is then merely to provide sufficient service that satisfies their mandate. When transfers are maximized and $\beta e^* + y < U^*$, the only instrument available to induce participation is better service than $e^*$. However, the more the client values the good, the less needs to be given to satisfy the service uptake or participation constraint.[5] In words, the more the client wants the good, the less he receives.

There is one other case with $\frac{de^*}{d\beta} < 0$, where $\beta e^{**} - x < U^* < \beta e^* - x$. In words, this is when - after all transfers are extracted from the client - he will not participate at the principal's preferred point ($e^{**}$) but will participate at the joint maximum ($e^*$). In this range, either the principal has to increase service from $e^{**}$ and/or reduce the price from $x$ in order to satisfy participation. This has a simple outcome: the principal continues to charge $x$ and only increases service from $e^{**}$ until he reaches $e^*$ in order to induce participation. At that point, he retains service at $e^*$ and only changes the price to induce participation until he hits the maximum transfer to the client, $y$. In the first of these two situations, $\frac{de^*}{d\beta} < 0$ for the same reason as above.

**Comments**

1. Many public agencies likely do not care much about monetary transfers, but rather have non-monetary objectives. The results above to not rely on the assumption that the principal cares about monetary transfers. A natural interpretation of $v$ is that the agency internalizes some client benefits: for example, social work agencies likely value any improvements in the lives of their clients. Consider a case where the preferences of the principal are $ve + \gamma\beta e$ but transfers are not feasible. The optimal choice in the absence of participation issues is given by $\tilde{e}$, given by $v + \gamma\beta = C'(\tilde{e})$. This has the natural outcome that service is increasing in the consumer's marginal preferences, but here the internalizing incurs through caring about the client rather than about money.

[5]The principal chooses $\tilde{e}$ such that $\beta\tilde{e} + y = U^*$, and so $\tilde{e}$ is decreasing in $\beta$. An empirically relevant case for many public bureaucracies is where transfers are negligible. Consider the case where no monetary transfers are possible. Then by necessity $p^* \notin (-y, x)$ and if $\beta e^* < U^*$, increases in client's return to quality reduce the quality offered.

Yet if $\beta\tilde{e} < U^*$, the client does not participate at these levels of service, and so the same influences arise as above: the principal satisfies his mandate at $\overline{e} = \frac{U^*}{\beta} > \tilde{e}$, which is decreasing in $\beta$. Hence the results carry over to the case where the principal cares about social surplus rather than money. Yet even in this case, allowing monetary transfers *from* the principal can relax this need, as the principal may prefer to bribe the client into participating through money rather than excessive service.[6]

2. Note above that I have assumed that the principal is mandated to provide service by requiring that the client participates, independent of $v$. The principal here is the bureaucratic agency. One interpretation is that $v$ reflects the value that the agency sees in satisfying client needs, not necessarily the view of society that mandates the service. From this perspective, $v$ could reflect agency problems for the bureaucratic agency itself. However, for completeness consider the case where the mandate for service only holds if the principal earns non-negative returns. Then the effects above arise until a level of service $\tilde{e}$ where $C(\tilde{e}) = v\tilde{e} + \tilde{p}$, where $\tilde{p}$ is the associated chosen price. This is the range up to which the principal earns rents, and so these effects continue to arise.

3. A simple way to understand the result above is through expenditure functions. Let the equilibrium amount paid be $p(e, \beta)$. Then the objective of the principal is to maximize $ve + p(e, \beta) - C(e)$. This has the feature that $\frac{de}{d\beta} = \frac{p_{e\beta}}{C''}$. When transfers are unconstrained, expenditures are convex in $\beta$ so $p_{e\beta} > 0$ and quality increases in $\beta$. However, when the budget constraint $p \in (-y, x)$ binds, $p_{e\beta} \leq 0$ at a lower level of $e$, as the budget constraint binds earlier and hence the result reverses.

4. In the model above, the principal makes a take it or leave it offer to the client. When distortion-free transfers are not possible, bargaining protocols can potentially matter. I show in the Appendix that the inverted outcomes also extend to a Nash bargaining setting when transfers are small.

[6] There is one case where the inverted effect does not occur. This is where the principal's preferences are identical to the net preferences of the client, where the principal also maximizes $\beta e - U^*$. Then the principal's preferences are those of the client's and so it makes no sense to mandate service if the client does not also want service.

# 2 Prices and Information Revelation

One of the central lessons of the literature on agency theory is the difficulty of using compensation to align incentives. As a result, much recent work has addressed other ways of providing incentives - giving workers control rights, leveraging their intrinsic motivation, career concerns, and so on. See Dixit, 2002, for an excellent review in the context of public agencies. A final instrument that agencies control is through deciding when their performance is scrutinized. This paper extends Prendergast, 2002, 2003, by addressing the role that clients play in that scrutiny. I model the mechanics of a typical service agency in which a bureaucrat makes a decision on service for a client. However, the bureaucrat sometimes errs, and the client can complain. Conditional on a complaint, the agency decides whether to scrutinize the case or overturn the original decision.

An assignment $A \in \{0,1\}$ is made to a client. Its depends on a state of nature $\alpha \in \{0,1\}$. The state is initially unknown but it is assumed that each state occurs with equal probability. The client's return when $A = i$ and $\alpha = j$ is given by $\beta_{ij}$. I assume that the client gains 0 utility from $A = 0$ (being denied the good): $\beta_{0j} = 0$, but $\beta_{11} = \overline{\beta} > \underline{\beta} = \beta_{10} > 0$. In words, the client benefits from $A = 1$ (receiving the good) more when $\alpha = 1$ than when $\alpha = 0$. The client's utility also depends on the price $p$ charged (meaning a price for assignment $A = 1$), where her utility from receiving the good is either $\overline{\beta} - p$ or $\underline{\beta} - p$. The client can opt out of consuming the good by choosing $A = 0$ at no cost.[7]

The principal's return when $A = i$ and $\alpha = j$ is $v_{ij}$. I assume that $v_{ii} > v_{ij}$, $j \neq i$, so that the principal wishes to match the assignment to the underlying environment, $\alpha$. I treat the $v_{ij}$ as parametric: they can be interpreted as including the benefits received by the client, or potentially prices (though see the discussion below), so long as the ordering above continues to hold.[8] This defines the central misalignment of interests between the client and the principal: the principal wants to deny the benefit to the client when the state is $A = 0$ but the client may want it anyway.[9]

[7]There are some cases where clients can be compelled to consume: for example, suspects cannot opt out of being arrested. The results below also hold for that case.

[8]For the ordering to hold when the principal cares about prices received from the client, he must still prefers allocation $A = 0$ when the state is $\alpha = 0$. As the maximum that the client would pay is $\overline{\beta}$, a sufficient condition for this ordering is that $v_{00} - v_{01} > \overline{\beta}$.

[9]There are of course cases where interests are aligned even without prices, where $\underline{\beta} < 0$. For example, few patients want invasive surgery if they do not need it. In these cases, there is no problem of aligning client interests. Equally, there are cases where benefits are independent of type, such as when monetary fines are imposed based on the supposed guilt of the client. In these cases, prices cannot be used to reveal client

There are three sources of information on $\alpha$. First, an agent or bureaucrat receives a signal $\alpha_a \in \{0,1\}$, where the bureaucrat's signal is correct with probability $e \geq \frac{1}{2}$. On the basis of his signal, the bureaucrat makes a recommendation $\hat{\alpha}$. The second source of information comes from the client. After the bureaucrat makes her recommendation, the client receives a signal $\alpha_c \in \{0,1\}$. If the bureaucrat's recommendation is incorrect, the client's signal always differs from $\hat{\alpha}$: she can identify error. However, if the bureaucrat's recommendation is correct, then the client's signal agrees ($\alpha_c = \alpha_a$) with probability $s > 0$, but with probability $1-s$, she incorrectly believes a mistake has been made and $\alpha_c \neq \alpha_a$.[10] The value of the client's information is parameterized by $s$: $s = 0$ is completely uninformative whereas $s = 1$ always identifies the true state. The final source of information is that the principal can investigate at fixed cost $\kappa$. An investigation identifies $\alpha$ and both signals.[11]

If prices are unrestricted, the principal can choose a price $p$ for the good. The only action taken by the bureaucrat is a recommended course of action $\hat{\alpha}$, which he makes after seeing his signal $\alpha_a$. After the client receives his signal, he also takes a single action: he sends a costless message, where he either complains or does not complain. A complaint is a message that his signal differs from the recommendation of the bureaucrat.[12] On the basis of these two actions, the principal then (i) investigates or not, and (ii) makes an allocation. Let $m \in \{0,1\}$ be the probability of an investigation ($m$ for "monitoring"). I assume that the principal holds control rights over $A$, and can commit to both when he investigates and to the allocation based on available information. These assumptions are similar to Prendergast, 2002, 2003. However, unlike that previous work we assume that the allocation by the principal must satisfy the client's desire to consume, as the client can costlessly opt out of $A = 1$.

So far there is no misalignment of interests for the bureaucrat. I allow a possible agency issue for her. The model is meant to reflect the reality of bureaucratic life, where investigations focus attention on her. Accordingly, there are assumed to be three possible wages: (i) with no investigation, she earns a wage of $w_0$, (i) if investigated and her signal and recommendation are correct, her wage is $w_1$, but (iii) if investigated and either her signal or her

information.

[10]There is an asymmetry in the client's information, where he mistakes correct decision but perfectly identifies bad ones. The reason for this asymmetry is to avoid frivolous complaints: where the client complains even though she does not believe that a mistake has been made. When the client believes that no error has occurred, there is no chance of the decision being overturned here. There are other ways to eliminate frivolous complaints, such as charging for a complaint or making the appeal process arduous. I use this asymmetry in information as one simple way to implement this.

[11]Whether an investigation identifies the client's signal is irrelevant.

[12]Messages are interpreted literally in this case.

recommendation are incorrect, her wage is normalized to 0.[13] I impose no assumptions so far on these wages other than $w_i \geq 0$: being investigated and found to be wrong (or to have been found to have lied) does not improve the bureaucrat's career.[14]

The timing of the game is as follows. First, the principal is endowed with a pricing regime for the good: either transfers are feasible or not. Given this, he then commits to a price $p$ (if feasible) and when he investigates. He commits to $m$, which can depend on the recommendation of the bureaucrat and whether the client complains. He also commits to the allocation based on all available information (the recommendation, complaint, and investigation outcome). Second, the bureaucrat observes her signal $\alpha_a$, and makes a recommendation $\hat{\alpha}$.[15] Following this, the client observes $\alpha_c$, and decides whether to complain. The principal then investigates according to the agreed contract, and the assignment is made by the principal. After this, the bureaucrat is paid and the game ends. I consider pure strategy Perfect Bayesian equilibria of the game.

It is useful to begin by describing the first best. This requires that the good is allocated efficiently based on all available information, and that the principal investigates when the return to doing so exceeds $\kappa$. Use the notation "agree" for the case where the signals concur, and "disagree" when the signals are different. If the signals agree, there is no value to investigating, and the principal implements the action suggested by the signals. When the signals disagree, the probability of an error in the bureaucrat's signal is $\rho$ where

$$\rho = \frac{1-e}{(1-e)+e(1-s)} = \frac{1-e}{1-es}. \tag{2}$$

This illustrates the role of the client: with $s = 0$ the probability of an error conditional on disagreement is $1-e$ while if $s = 1$, then the client perfectly identifies mistakes. The optimal assignment without an investigation depends on the signal, the payoffs, and the quality of information. If the bureaucrat's signal is $\alpha_a = 0$ and the client's signal disagrees, the return to implementing the bureaucrat's signal is $\rho v_{01} + (1-\rho)v_{00}$ while the return to overturning it is $\rho v_{11} + (1-\rho)v_{10}$. The bureaucrat's signal is overturned if $\rho \geq \frac{v_{00}-v_{10}}{v_{11}+v_{00}-v_{10}-v_{01}}$, with surplus $max\{\rho v_{01} + (1-\rho)v_{00}, \rho v_{11} + (1-\rho)v_{10}\}$. However, the principal may investigate. Let

[13] This captures the idea that the bureaucrat is penalized either if she was wrong or she lied. As lying does not occur on the equilibrium path, in equilibrium it is merely a reward for "getting it right". In Section 3.1 I show that the results are unchanged when lying is not directly penalized.

[14] One interpretation could entail a worker who is rewarded only for keeping her job $w_1 = w_0 > 0$, while another could be a convex return where workers are looking for their chance to show how good they are $w_1 > w_0 = 0$. In the former case, the bureaucrat shuns investigations whereas the latter wants an investigation, all else equal.

[15] This timing is discussed in the conclusion.

$m_d^*(i) \in \{0,1\}$ be the optimal probability of an investigation when the signals disagree and the bureaucrat observes $\alpha_a = i$. Investigation gives rise to efficient allocations so $m_d^*(0) = 1$ if $\kappa \leq \rho v_{11} + (1-\rho)v_{00} - max\{\rho v_{01} + (1-\rho)v_{00}, \rho v_{11} + (1-\rho)v_{10}\}$. When the bureaucrat's signal is $\alpha_a = 1$, a similar logic applies and $m_d^*(1) = 1$ if $\kappa \leq \rho v_{00} + (1-\rho)v_{11} - max\{\rho v_{10} + (1-\rho)v_{11}, \rho v_{00} + (1-\rho)v_{01}\}$. This describes the first best.[16]

## 2.1 When Prices Reveal Information

The scenario studied below is one where clients cause asymmetric oversight through their complaints: when a client complains about certain suggested allocations, oversight of those recommendations becomes unusually high. As a result, I consider the case where the first best involves symmetric oversight in Assumption 1 (I consider the asymmetric case below).[17]

**Assumption 1:** $m_d^*(1) = m_d^*(0)$.

There is no necessary reason to think that the client's and bureaucrat's interests are aligned with those of the principal. The bureaucrat may either wish to avoid oversight or may desire it, based on the $w_i$. Similarly, the client would like to consume the good when he should not if the price is low, and not consume it when he should if the price is too high. First consider the case where there is no constraint on the choice of $p$ by the principal. Remember that I have imposed no restrictions on wages other than $w_i \geq 0$. Despite this, Proposition 2 shows how clients eliminate agency problems, and can do so with a very simple mechanism, a single price (proofs are in the Appendix).[18]

**Proposition 2** *Assume Assumption 1 holds and that prices are unconstrained. Then the first best arises as an equilibrium if* $w_i \geq 0$.

By appropriate choice of a price, the principal can align the client's interest with his, by persuading him to reveal cases where his signal disagrees with the bureaucrat's recommen-

[16] An simple example is where the payoffs are symmetric and the returns to a correct allocation normalized to 1, where $v_{ii} = 1, v_{ij} = 0, i \neq j$. Then after a disagreement, the bureaucrat's decision is overturned if $\rho \geq \frac{1}{2}$ with surplus $es + (1-e+e(1-s))max\{\rho, 1-\rho, (1-e)(1-\kappa) - e(1-s)\kappa\}$.

[17] Symmetry requires either $\kappa \leq max\{\rho v_{11} + (1-\rho)v_{00} - max\{\rho v_{01} + (1-\rho)v_{00}, \rho v_{11} + (1-\rho)v_{10}\}, \rho v_{00} + (1-\rho)v_{11} - max\{\rho v_{10} + (1-\rho)v_{11}, \rho v_{00} + (1-\rho)v_{01}\}\}$ or $\kappa \geq min\{\rho v_{11} + (1-\rho)v_{00} - max\{\rho v_{01} + (1-\rho)v_{00}, \rho v_{11} + (1-\rho)v_{10}\}, \rho v_{00} + (1-\rho)v_{11} - max\{\rho v_{10} + (1-\rho)v_{11}, \rho v_{00} + (1-\rho)v_{01}\}\}$.

[18] The proposition describes this as an equilibrium. There is another pure strategy equilibrium here involving no complaints. If the client believes that complaints are never responded to, no complaint is sent, and this is an equilibrium.

dation. If the client's incentives are aligned, so also are the bureaucrat's if $w_i \geq 0$. Hence the first best is attained. As such, this Proposition shows the important role that end users can play in resolving agency problems, and it can be implemented through the simplest of instruments, a single price.

The first best maximizes the welfare of the principal. This need not be divorced from the client's welfare. A natural case is where the principal's preference is the welfare of the client minus a cost $\gamma$ of producing the good. Then if $\overline{\beta} > \gamma > \underline{\beta}$, the results carry over as the required ordering holds.

Two other intuitive outcomes arise in the first best. First, better informed clients weakly increase surplus. Second, better informed clients are more influential in that the likelihood of the bureaucrat's recommendation being changed is weakly increasing in $s$. This arises in either of two ways: (i) overturning the decision without an investigation, or (ii) causing an investigation.[19] Both are (weakly) more likely as $s$ rises, and so the intuitive implication of better informed clients described in the introduction arises.

**Monopoly Distortions:** The first best result above shows how the principal can design a single price that reveals the information of the client. It has been assumed above is that marginal changes in $p$ only affects the principal's utility through its effect on allocations. In many settings, public agencies are funded through budgetary allocations rather than through payments from clients, and their objectives may be more closely approximated by non-monetary missions ("reducing crime", "housing the poor", "securing our borders", etc.).[20] But if $p$ affects the principal's returns directly (such as where the return to the principal is $v_{ij} - p$), then a standard monopoly distortion can arise where the surplus from the allocation may not be maximized if it involves too small a transfer to the principal relative to some alternative. As a result, monopoly distortions can potentially upend the first best result. It is worth emphasizing, however, that because the principal can commit to allocations, this problem can be resolved by allowing ex ante transfers between the client and the principal before information is known by any party. If such transfers are allowed, this can be retransformed into a surplus maximization problem in the now standard way.

[19] The former occurs after a recommendation of $\hat{\alpha} = i$ if $\rho \geq \frac{v_{ii} - v_{ji}}{v_{11} + v_{00} - v_{10} - v_{01}}$ when $\kappa > \rho v_{jj} + (1-\rho)v_{ii} - max\{\rho v_{ij} + (1-\rho)v_{ii}, \rho v_{jj} + (1-\rho)v_{ji}\}$, and the latter arises if $\kappa \leq \rho v_{jj} + (1-\rho)v_{ii} - max\{\rho v_{ij} + (1-\rho)v_{ii}, \rho v_{jj} + (1-\rho)v_{ji}\}$. Both are weakly increasing in $s$.

[20] Additionally, transfers may not be internalized by the agency because they may not be reflected in higher budgets.

**State-Contingent Prices:** I have shown that the use of prices can generate first best outcomes. This was implemented here through a single non-contingent price for receiving the good. There are many pricing arrangements that can potentially replicate this outcome, and no claim is being made that pricing has to be implemented this way.[21] The central point here is that very simple pricing can induce first best outcomes.

## 2.2 Non-Priced Goods

Now turn to the case where the good cannot be priced, as many bureaucracies are severely constrained in their ability to charge. Accordingly, consider the case where $p = 0$ (what matters with a single price is that he cannot charge at least $\underline{\beta}$).

Our interest is in how clients affect the incentives of bureaucrats by inducing focus on their performance. For this to matter, both the client and the bureaucrat must be relevant to the allocation, which needs two ingredients. First, the bureaucrat must influence the outcome in Assumption 2.

**Assumption 2:** $e[v_{ii} - v_{ji}] > (1-e)[v_{jj} - v_{ij}]$ for $i = 0, 1$, and $i \neq j$.

This assumption implies that if the bureaucrat's signal is $i$ and there is no other information, the principal earns higher surplus from following that signal than from overturning it. In the symmetric case where $v_{11} - v_{01} = v_{00} - v_{10}$, Assumption 1 always holds as $e > \frac{1}{2}$.

Second, I assume that oversight is symmetric in the absence of the client: this allows the client to be potentially marginal for focused oversight. Consider a counterfactual where the client's information is ignored. Let $m_i, i = 0, 1$, be oversight if the bureaucrat has $\alpha_a = i$, with $m_i^*$ its ex post optimal level. Her signal is correct with probability $e$, and so $m_i^* = 1$ iff $\kappa \leq (1-e)[v_{jj} - v_{ij}]$. Assumption 3 implies symmetry in the absence of client intervention.[22]

**Assumption 3:** $m_0^* = m_1^*$.

As with Assumption 1, this is always satisfied with symmetric payoffs for the principal. Now consider the client's role in the absence of pricing. Two issues arise. First, when the bureaucrat recommends $\hat{\alpha} = 1$, the client never complains if complaints increase oversight.

[21] One possibility is to make prices state contingent. For example, a client who legitimately appeals a decision could be offered a lower price than one who receives the good without appealing. In this way, complaints could be subsidized.

[22] This assumption requires that $\kappa \leq (1-e)min\{v_{11} - v_{01}, v_{00} - v_{10}\}$, or $\kappa \geq (1-e)max\{v_{11} - v_{01}, v_{00} - v_{10}\}$.

This is because $\underline{\beta} > 0$ and the client has nothing to gain from complaining. As a result, the outcome is independent of the client when $\hat{\alpha} = 1$.[23]

Second, unlike the case where client interests are aligned, a complaint cannot overturn a recommendation without an investigation. If it did, a client who is denied and whose signal agrees would also complain. As a result of these two issues, the principal either implements the bureaucrat's recommendation (because of Assumption 1) or investigates. Given this, in any setting where complaints affect outcomes, there are only two cases where oversight may occur: when $\hat{\alpha} = 0$ and there is a complaint, and when $\hat{\alpha} = 1$.

To understand the agency problem, consider a bureaucrat with $\alpha_a = 0$. Is there an equilibrium where the client is denied? Let $m_d \in \{0, 1\}$ be the probability of an investigation after the client complains, with $m_d^*$ its ex post optimal level. Then $m_t = [e(1-s)+(1-e)]m_d$ is the probability of an investigation if the bureaucrat denies the client, with expected wage $w_1 e(1-s)m_d + w_0(1-m_t)$. The (out-of-equilibrium) return to capitulating to the client by choosing $\hat{\alpha} = 1$ when $\alpha_a = 0$ is $w_0(1-m_1)$. The bureaucrat denies only if

$$w_1 e(1-s)m_d \geq w_0(m_t - m_1) \tag{3}$$

I refer to this as the "bureaucratic constraint".

The tension here is that while client information corrects error, it also potentially threatens the bureaucrat. There are two ways in which the client affects the bureaucrat. First, a complaint may be informative enough to launch an investigation. When $m_0^* = 0$, it is not worthwhile to investigate without a complaint. However when the client complains, the probability of error is $\rho = \frac{1-e}{1-es} > 1-e$. Then for $s \geq \tilde{s}$, where

$$\tilde{s} = \frac{1}{e} - \frac{(1-e)[v_{11} - v_{01}]}{e\kappa} \tag{4}$$

a complaint causes an efficient investigation with $m_d^* = 1$.

The ordering of $m_d^*$ and $m_1^*$ is important below. If a complaint does not affect efficient oversight ($m_d^* = m_1^* = 0$) there is no incentive problem for the bureaucrat. The focus here is where $m_1^* = 0$ but $s > \tilde{s}$, in which case $m_d^* = 1$. (The case where $m_1^* = 1$ is discussed below.) The bureaucratic constraint when a complaint is needed to trigger an efficient investigation is

$$w_1 \frac{e(1-s)}{1-es} \geq w_0 \tag{5}$$

The bureaucratic responses below arise from the comparative statics of (5). This incarnation of the bureaucratic constraint offers a second way in which the client can affect the

[23] Technically, oversight can be no higher with a complaint than without.

bureaucrat: conditional on an investigation, she is more likely to be wrong when the client is better informed. The critical value of $s$ above which she capitulates to the client is found by solving (5) with equality and is given by

$$s^* = 1 - \left(\frac{1+e}{e}\right) \frac{w_0}{w_1 - w_0}. \tag{6}$$

Before considering how oversight is distorted to ensure the bureaucratic constraint holds, note that (5) offers a simple interpretation of wages in terms of the probability of making the correct decision. Much of the recent literature on public officials stresses the importance of career concerns. From that perspective, bureaucrats are rewarded for their ability, a natural metric for which is the belief that the bureaucrat made the correct allocation. Now consider (5). Conditional on an investigation, the bureaucrat is correct with probability $\frac{e(1-s)}{1-es}$. If wages are linear in that probability (so $\frac{w_1}{w_0} = \frac{1-es}{e(1-s)}$), the bureaucrat is indifferent about telling the truth or concealing her information. This is because an investigation is simply a mean preserving spread on the prior $e$. Said another way, the bureaucratic constraint fails whenever wages are sub-linear in these probabilities: $\frac{w_0}{w_1} > \frac{e(1-s)}{1-es}$.

Note that neither the client nor bureaucrat plays a role unless the bureaucratic constraint is satisfied, as the bureaucrat simply gives the good to the client, and the client never complains. Hence, in stark contrast to the case where prices are flexible and the first best can be achieved with any $w_i \geq 0$, here a necessary condition for the bureaucrat or client to have any role at efficient oversight is sub-linearity in wages.

# 3 Bureaucratic Responses

This section's interest is in how agencies respond to their clients. From the previous section, (i) if $s < \tilde{s}$, the client's information is not valuable enough to cause an investigation, and (ii) if $s \in [\tilde{s}, s^*]$, the principal investigates after a complaint ($m_d^* = 1$, and $m_1^* = 0$) and the client will be denied when appropriate. The remaining case is how the principal responds if $s > s^*$. To see this, note that the principal never designs oversight such that the bureaucratic constraint is violated.[24] Satisfying the constraint is accomplished here in one of two ways: (i) by ignoring complaints ($m_d^* > m_d = 0$), or (ii) by always investigating when the client is given the good ($m_1 = 1 > m_1^*$).

[24]This is shown in the Appendix, but intuitively a violated bureaucratic constraint loses the information of both the client and bureaucrat, and the principal can do weakly better than this by having the constraint hold.

Ignoring complaints fails to correct allocative error, and generates surplus $\frac{ev_{00}+(1-e)v_{10}}{2}+\frac{ev_{11}+(1-e)v_{01}}{2}$ (as if the bureaucrat is the only party involved in the assignment).[25] The alternative is to investigate when the client is given the good, so $m_1 = 1$. (In effect, a second pair of eyes is necessary before giving a treatment to the client.) This has the benefit of efficient allocation at the cost of over-monitoring, with surplus $\frac{v_{11}}{2}+\frac{v_{00}}{2}-\frac{\kappa}{2}-[(1-e)+e(1-s)]\frac{\kappa}{2}$. Oversight in that case is biased against well informed clients as decisions that benefit them are always scrutinized whereas denials are not.[26] The principal then optimally ignores complaints if $\kappa \geq \frac{1-e}{2-e+e(1-s)}[v_{11}-v_{10}+v_{00}-v_{01}]$. In words, when random oversight is sufficiently costly, the optimum is to ignore complaints from the most informed clients. The outcome is described in Proposition 3.

**Proposition 3** *Assume that $m_1^* = 0$ and prices are not feasible. Then (i) if $s < \tilde{s}$, $m_d = m_d^* = 0$, (ii) if $s \in [\tilde{s}, s^*]$, $m_d = m_d^* = 1$, and (iii) for $s \geq s^*$,*

- *If $\kappa \geq \frac{1-e}{2-e+e(1-s)}[v_{11}-v_{10}+v_{00}-v_{01}]$, the principal ignores valuable client complaints: $m_d = 0 \neq m_d^*$.*
- *If $\kappa < \frac{1-e}{2-e+e(1-s)}[v_{11}-v_{10}+v_{00}-v_{01}]$, the principal investigates when the client is allocated the good: $m_1 = 1 \neq m_1^*$.*

These results shows how the absence of prices distorts bureaucratic responses. The first part shows how the usual comparative statics get overturned, where a sufficiently well informed client is cut out of the allocation process, even though his input is most valuable. The alternative is to always investigate when more informed clients are given services, even though those clients have no input.

More than simply showing that the bureaucracy becomes distorted, it also illustrates when, by noting the determinants of $s^*$ in (6). First, a more informed bureaucrat ($e$ higher) increases $s^*$: intuitively, better informed bureaucrats worry less about the information of their clients. Second, $s^*$ depends on how the bureaucrat is compensated, with $s^*$ increasing in $w_1$: those who get more upside benefits from being proved right worry less about denying a client. However, in most public agencies, wages are relatively insensitive on the upside. Instead, the only real danger is the possibility of being fired for some perceived incompetence. When $w_1$ is close to $w_0$, $s^* = 0$, and so the bureaucrat always capitulates to the client.

[25]By doing so, the bureaucrat earns $w_0$ for sure and so is willing to deny the client when she feels it is appropriate.

[26]The bureaucratic constraint here is $e(1-s)w_1 + w_0es \geq 0$, which always holds.

Note from Proposition 3 that unless $\tilde{s} < s^*$, the client can never induce efficient investigation. However, in reality, it is unlikely that this condition holds for most public servants. A necessary condition for $\tilde{s} < s^*$ is that the bureaucrat strictly benefits from random oversight. Random oversight - meaning an investigation triggered with no input from the client - requires that $w_1 e > w_0$. Such convexity in wages is unlikely for most public servants. Yet for $\tilde{s} < s^*$ to hold, there must be a range of $s > 0$ where $w_1 \frac{e(1-s)}{e(1-s)+(1-e)} > w_0$: the bureaucrat wishes to be investigated, even though complaints are partially informative of their errors. As this seems implausible, the empirically relevant effect of client information on the bureaucrat is likely through his ability to trigger an investigation.[27]

Of interest here is how more informed clients affect efficiency. With flexible prices, better information weakly increases surplus. When the bureaucratic constraint binds with constrained prices, this is no longer true. When the principal ignores complaints as the optimal response, surplus increases in $s$ up to $s^*$ but then discretely falls to the level that would arise with $s = 0$ and remains at that level for all $s > s^*$. In the case where $s^* < \tilde{s}$, clients play no role in the allocation.[28]

The advent of the internet has resulted in better informed consumers. For example, the ability to provide reviews of sellers on Amazon.com or Yelp.com has surely improved agency problems, and Thomas and Stanton, 2013, show that companies such as ODesk have reduced adverse selection in some kinds of hiring. The paper proposes that this has different implications for bureaucracies than for other firms. When prices are flexible, better informed clients improve efficiency, both directly through correcting error but also through economizing on the need for other oversight. However, without appropriate prices, the opposite arises. Here as clients become better informed, bureaucratic agencies may respond by ignoring them: they are ultimately cut out of the allocation process. Alternatively, the agency incurs costs of over-monitoring cases.

This work revisits Prendergast, 2002, 2003, in its interest in how clients affect bureaucracies. It extends that work in a number of directions. First, the prior work (i) restricted attention to strong symmetry in the payoffs for the principal and (ii) imposed a strong as-

[27] The final conceivable way of solving the problem is by always investigating when the client is denied, even when the client does not lodge a complain. However, the bureaucratic constraint here is $w_1 e \geq w_0$ which is violated with wages sub-linear in probabilities. Hence this does not solve the problem.

[28] When the principal responds by choosing $m_1 = 1$, surplus discretely falls at $max\{\tilde{s}, s^*\}$ as the principal now incurs the (inefficient) cost of monitoring. However, after that initial drop better informed client continue to focus complaints better and surplus rises from that lower level. In that case, the clients who minimize efficiency are those at $max\{\tilde{s}, s^*\}$.

sumption on wages (where $w_1 = w_0 > 0$). By considering more general valuations and wages, I can show both the generality of the benefits of using client when goods are priced, and the relatively strong requirements for avoiding capitulation to clients without them. Furthermore, this does not rely on a complicated mechanism that could be hard to implement. Rather, all it requires is a single non-contingent price.

The more general setting also shows that what matters for bureaucratic problems is when wages are sub-linear. As a necessary condition to overcome the bureaucratic constraint at efficient oversight is (weak) convexity of wages, such convexity seems implausible. As such, it illustrates the central role of pricing in harnessing the value of client information: with pricing, the conditions on wages are minimal, yet without it, the required wages seem implausible in reality. Considering more general wages also shows that unless random oversight is beneficial to bureaucrats, consumers are always cut out of the allocation process when random oversight is costly.

The final contribution of the paper is its focus on how bureaucracies respond to better informed clients,. Prendergast, 2002, introduces the idea that oversight may be distorted by the need to induce truth-telling, by over-monitoring in some cases and by ignoring complaints in others. Prendergast, 2003, focuses largely on how variation in the quality of client information can crowd out effort exertion by the bureaucrat.[29] This work combines the endogenous distortions required for truth telling with variations in the informativeness of client signals to induce inverted bureaucratic responses that depend on consumer pricing. Finally, this work also extends the previous literature by (i) allowing client complaint to directly overturn recommendations, and (ii) allowing the client to have an exit option, thereby extending the set of relevant cases beyond those where the bureaucrat is sovereign.

## 3.1 Extensions

In this section, extensions to the set-up above are considered.

**Probabilistic Oversight:** So far, I have assumed that the oversight choice is binary. The results are easily extended to the case where oversight is probabilistic. As returns to oversight are linear in probabilities, the first best is not changed by this option. The only change is in the responses needed when the bureaucratic constraint is violated. Consider a case where

[29]This was done in a setting where the principal can charge a price for complaints as a simple reduced form way of making complaints credible. I make no such assumption here: instead the credibility of complaints derives from the information structure.

the principal can commit to a probability $\mu_d$ of oversight when a complaint occurs, and a probability $\mu_1$ when the client is assigned the good. Then the bureaucratic constraint is given by

$$w_1 e(1-s)\mu_d \geq w_0(\mu_t - \mu_1), \tag{7}$$

where $\mu_t = [e(1-s)+(1-e)]\mu_d$. Consider each case above: either changing $m_d$ from $m_d^* = 1$ or changing $m_1$ from $m_1^* = 0$. First note that when (5) is violated and $\mu_1 = 0$, the only way that the principal can satisfy this amended bureaucratic constraint is by choosing $\mu_d = 0$: all complaints are ignored. As a result, allowing probabilistic oversight does not change this outcome. However, when the principal chooses to distort $m_1$, he need only monitor with probability $\tilde{\mu}_1 = (1 - \frac{w_1}{w_0})e(1-s) + (1-e)$, as this causes (7) to bind. However, in keeping with the inverted outcomes above, note that $\tilde{\mu}_1$ is increasing in $s$ if $w_1 > w_0$, so that clients who know their own needs best face the greatest amount of oversight.

**When $m_1^* = 1$:** I have so far assumed that it is ex post efficient to investigate only after a complaint. There are many cases where it is difficult to identify what bureaucrats should have done in the typical case. Police officers may see hundreds of people a day, and those at immigration may see thousands: random investigation is like searching for a needle in a haystack. The bureaucratic responses of the paper are specific to the case where random oversight is costly. To see this, consider the case of $m_1^* = 1$, which arises if $(1-e)[v_{00} - v_{01}] > \kappa$. The logic of the bureaucratic incentive problem is that she can misallocate to avoid oversight. When $m_1^* = 1$, this is no longer true. The bureaucratic constraint now always holds with sub-linear wages, and so the inverted outcomes above do not arise. As such, the focus of this work is on agencies where signals of error are needed for oversight.

**Red Tape and Delay:** One outcome of the previous section is the "bang bang" nature of responses to complaints: at some point, the client is discretely cut out of the process. Here I consider an alternative instrument that has a more continuous outcome: where the principal cannot commit to $m_i$, but to the quality of the bureaucrat's information. Agencies often choose the level of certainty that they need before moving forward. This can be done either by vast amounts of paperwork (red tape) or by delaying decisions until all relevant information is available.

Accordingly, assume that the principal has an additional instrument $t \geq 0$ (for "time") that increases the precision of the bureaucrat's information to $\tilde{e} = e + t$. This occurs at a cost $C(t) \geq 0$, where $C(t)$ is increasing in $t$ with increasing marginal cost. (The time interpretation here would be that the value of any benefits are discounted.)

The principal now commits to $t$ rather than $m_i$. More formally, the timing is changed to "The principal commits to a price $p$ (if feasible) and chooses $t$". Oversight is at its ex post optimal level here, and I consider the case where $m_d^* = 1$ and $m_1^* = 0$. The bureaucratic constraint (5) is amended here to

$$w_1 \frac{\tilde{e}(1-s)}{1-\tilde{e}s} \geq w_0 \tag{8}$$

The objective of the principal is then to choose $\tilde{e}$ to maximize surplus subject to (8). If the bureaucratic constraint is violated at the technologically efficient level of $t$, the principal chooses the lowest level of $\tilde{e}$ to satisfy it. But setting (8) to hold with equality and totally differentiating yields

$$\frac{d\tilde{e}}{ds} = \frac{(w_1 - w_0)\tilde{e}}{w_1(1-s) + w_0} \geq 0 \tag{9}$$

if $w_1 \geq w_0$. In words, $e$ and $s$ are complements, and the relationship is continuous. Once again, a violated bureaucratic constraint distorts the bureaucracy, though here it arises by the agency being excessively certain before making a decision. Yet because $e$ and $s$ are positively related, those who know their needs best must wait longest for service, whereas those who (locally) know less receive faster service (see also Banerjee, 1997, for another theory of red tape). Hence the agency is less responsive to the best informed clients.

**Endogenous Wages:** So far, I have treated wages as parametric and the bureaucratic constraint has been interpreted through the lens of sub-linearity with respect to probabilities of being correct. In this section, I provide an interpretation of such sub-linearity through a career concerns lens.

Career concerns reward agents for their perceived ability. Accordingly, assume that wages are linear in the ex post likelihood that the bureaucrat made the correct decision: if the bureaucrat made the correct decision with probability $z$, then the wage is normalized to be $z$. In the model above, there are three wages. Its interpretation is that the market only observes true outcomes if an investigation occurs. Otherwise the market observes nothing beyond the absence of an investigation.[30]

Consider the baseline case of $m_d^* = 1$ and $m_1^* = 0$. At efficient oversight, this wage generating process gives rise to a desire to avoid investigation. In words, those who are not

[30] It is unrealistic to assume that the labor market observes all outcomes for bureaucrats: most happenings inside firms are unknown to outside parties. In the case where the market observes all recommendations and all responses by the bureaucrat, there is no agency problem. The reason is that whatever action is taken by the bureaucrat results in a mean preserving spread of wages, and as the bureaucrat is risk neutral, she is indifferent about this. Hence, some limited observability is necessary for the agency problem.

investigated are on average better than those who are. The wage earned if correct is $w_1 = 1$ and the wage earned if incorrect is 0. The wage earned when not investigated, $w_0$, is the probability of being correct conditional on no investigation, and is given by

$$w_0 = \frac{e + es}{1 + es} = \frac{e(1+s)}{1+es} \geq e. \tag{10}$$

The bureaucratic constraint then becomes

$$\frac{e(1-s)}{1-es} \geq \frac{e(1+s)}{1+es}. \tag{11}$$

This constraint is violated for any $s > 0$, as wages are now endogenously sub-linear in probabilities as required.

**More Informative Wages:** There are two pieces of information on the merits of the bureaucrat's recommendation: (i) whether the client complains, and (ii) whether an investigation arises. I have assumed that wages can only be conditioned on the second of these in keeping with the public nature of investigations. For completeness, consider the case where wages can be conditioned on both.

There are now four relevant wages that the bureaucrat can earn: $w_1$ if investigated and correct, 0 if investigated and incorrect, $w_c$ if a complaint is made and there is no investigation, and $w_n$ with no complaint and no investigation.[31] In the benchmark case where $m_d^* = 1$ and $m_1^* = 0$, all complaints are investigated and so the bureaucratic constraint remains at (5). Hence this extension does not change the basic result.

**Asymmetric Efficient Oversight:** Proposition 2 shows that the first best is possible if prices are unrestricted and efficient oversight is symmetric. The symmetry of efficient oversight is important. Consider the case where the client's incentives are aligned with those of the principal through choice of prices, yet efficient oversight is biased with $m_d^*(i) = 1$ and $m_d^*(j) = 0$.[32] If (5) does not hold at these efficient levels of oversight, the bureaucrat always offers allocation $j$ to the client independent of his signal, even though the client's

[31] As investigations reveal the truth about the bureaucrat's decision, it is redundant to condition the wage if investigated on whether a complaint arose.

[32] Many services to clients have asymmetric payoffs such that the cost of failing to give service incorrectly exceeds the cost of giving it when not needed. For example, in many medical settings, it is better to do excess testing than too little. And in a legal setting, it is a bedrock of the legal system that the cost of convicting an innocent man exceeds the cost of letting a guilty one go free. In our model's language, this implies that $v_{11} - v_{10} \geq v_{00} - v_{01}$.

interests are aligned. If the principal chooses to have the bureaucratic constraint hold, then the same insights arise as above: either complaints are not investigated or there is over-monitoring for recommendation $j$ (though here those investigations can be targeted at cases where complaints are made). As a result of this insight, what matters for distorted oversight is that monitoring is asymmetric, not necessarily that the biased incentives of clients caused such asymmetric oversight.[33]

The same issue is true with constrained prices if Assumption 3 is violated. For example, consider the case where $m_0^* = 1$ but $m_1^* = 0$. Then the bureaucratic constraint is (5) even in the absence of the client being involved in the allocation. If so, then the principal distorts oversight in the same way as above such that $m_d$ and $m_1$ are equated.

**The Role of Investigations:** Investigations play two roles above: (i) they correct error, but (ii) also directly penalize the bureaucrat for lying. However, in some settings, it could be that the bureaucrat is only penalized if he got it wrong, even if he lied. (For example, in a career concerns setting, outside observers may only observe that the bureaucrat's recommendation was vindicated even if he lied in that recommendation.) To see the implication of this, consider an alternative where an investigation offers a reward $w_1$ if the recommendation is correct and $w_0$ if the recommendation is incorrect, independent of his signal.

This assumption is relevant for truth telling only when lying involves an investigation. As a result, it has no effect on the baseline case of $m_d^* = 1$, $m_1^* = 0$ as lying (by giving the client the good) never causes investigation. However, it could be relevant when prices are unrestricted or when prices are restricted and $m_d^* = 1$ and $m_1^* = 1$.

**Proposition 4** *Assume that prices are unconstrained and that after an investigation the wage is $w_1$ if $\hat{\alpha} = \alpha$ and 0 otherwise. Then the first best arises as an equilibrium if $w_i \geq 0$, and $m_d^*(0) = m_d^*(1)$.*

In the case where wages are sub-linear in probabilities, the outcomes with constrained prices from the previous section also remain unchanged (this is shown in the Appendix). As a

[33] There is one difference in this case compared to the bureaucratic responses above. Specifically, the principal may choose to allow the bureaucratic constraint to be violated. In the baseline model above, when the bureaucrat capitulates to the client, the principal gains the information of neither party. As a result, it is never optimal to allow the bureaucratic constraint to be violated. This is not necessarily true here. As the client's information is not distorted, the principal may find it optimal to allow the bureaucrat to allocate $\hat{\alpha} = j$ and respond to the client's complaint. Said another way, the principal now has a choice between whether to use the bureaucrat's information (by having the bureaucratic constraint hold through distorted oversight) or only the client's information.

result, the outcomes are robust to this change in assumptions.

# 4 Conclusion

This is not the first paper to address how bureaucracies provide incentives. Much of this literature has been concerned with identifying instruments other than pay to affect actions. What levers can these agencies pull to induce better performance? The literature has highlighted a number of possibilities: the use of focus on a small number of activities (Dewatripont et al, 1999, Prendergast, forthcoming), intrinsic motivation (Besley and Ghatak, 2005, Prendergast, 2008), and delegation of control (Aghion and Tirole, 1997). This paper fits into this line of research by arguing that the choice of when to oversee performance may be a useful instrument to affect outcomes.

Bureaucracies come in many shapes and sizes, and vary in their ability to monitor performance. In this context, a useful taxonomy is Wilson's, 1989, categorization of public agencies based on the ability to monitor inputs to performance and their outputs. Oversight is most difficult in settings where direct measures of both outputs and inputs are poor, what Wilson calls "coping organizations". Wilson's prescription for appropriate agency actions here feels little more than throwing his hands in the air: "they can hire the best people (without having much knowledge about what the best person looks like), they can create an atmosphere that is conducive to good work (without being certain what "good work" is) and they can step in when complaints are heard or crises erupt (without knowing whether a complaint is justified)" (p.169). These institutions are the focus of this work, which argues that the need to use interested parties to reveal error potentially gives rise to unusual bureaucratic responses.

One of the main implications of this work is that valuable information by clients may be ignored. It is appropriate to be somewhat more nuanced in interpreting this. The timing of the model is that the client observes information that the bureaucrat does not have at the point of her recommendation. Realistically, the allocative process is more sophisticated than this, where clients show some information at the point where the recommendation is to be made. Yet bureaucrats make mistakes, perhaps by claiming that the client's information is not pertinent. The model in reality likely best deals with cases where the client's complaints address his information being ignored in the recommendation. Hence the timing assumption need not be taken literally to mean that the client has no role in the allocation of public services, rather that the equilibrium appeals process chosen by the agency likely involves excessive obstacles.

To summarize, because prices rarely clear bureaucratic markets, their goods are allocated differently to other firms. The central point of this paper is that this causes strange but constrained efficient bureaucratic responses. This is meant to contribute to the literature in a number of ways. First, and most straightforward, public agencies should not be expected to act on the margin as others do. Second, it is well known that pay for performance works poorly in bureaucratic settings. This paper makes the point that a natural alternative - letting end users reveal error - may also backfire, especially when those end users are well informed. An extension of this is that the rise of consumer information - largely through online searches - is likely to exacerbate the differences between “normal” firms and bureaucracies, in that such information likely economizes on the firm’s time and resources in such normal firms, but may impose greater costs on public agencies.

# References

**Aghion, Philippe, and Jean Tirole.** 1997. Formal and Real Authority in Organizations. *Journal of Political Economy*, 105 (1): 1-29.

**Banerjee, Abhijit.** 1997. A Theory of Misgovernance. *Quarterly Journal of Economics*, 112(4): 1289-1332.

**Besley, Tim, and Maitreesh Ghatak.** 2005. Competition and Incentives with Motivated Agents. *American Economic Review*, 95(3): 616-636.

**Daily Telegraph.** 2014. www.telegraph.co.uk/health/nhs/10595806/NHS-patients-should-be-pushy-with-GPs-about-treatment-and-drugs-says-health-chief.html.

**Dewatripont, Matthias, Ian Jewitt, and Jean Tirole.** 1999. Missions and the Accountability of Government Agencies. *Review of Economic Studies*, vol. 66(1): 199-217.

**Dixit, Avinah.** 2002. Incentives and organizations in the public sector: an interpretative review. *Journal of Human Resources* 37, no 4: 696-727.

**Prendergast, Canice.** 2002. Consumers and Agency Problems. *Economic Journal*: 112: C34-C51.

**Prendergast, Canice.** 2003. The Limits of Bureaucratic Efficiency. *Journal of Political Economy*. 111(5): 929-958.

**Prendergast, Canice.** 2007. The Motivation and Bias of Bureaucrats. *American Economic Review*, 97(1): 180-96.

**Prendergast, Canice.** 2014. Professionalism and Contracts in Organizations. *Journal of Labor Economics*, forthcoming.

**Skolnick, Jerome.** 1966. *Justice without Trial*, John Wiley, New York.

**Thomas, Catherine and Christopher Stanton.** 2013. Landing the First Job: The Value of Intermediaries in Online Hiring, mimeo.

**Wilson, James** 1989. *Bureaucracy*. Basic Books, New York, .

**Appendix:**

**Proof of Proposition 1:** The objective of the principal is to choose $p(e)$ and $e$ to maximize $ve + p(e) - C(e)$ subject to $p^* \in (-y, x)$, and $\beta e - p \geq U^*$.

First consider the case where the $p^* \in (-y, x)$ constraint is ignored. Then $p$ is chosen such that $p = \beta e - U^*$. Substituting this into the objective function and differentiating yields $C'^{-1}(v + \beta) = e^*$ and $\frac{de^*}{d\beta} = \frac{1}{C''} > 0$ so quality is increasing in $\beta$.

The first best transfer is $p^* = \beta e^* - U^*$. This outcome is feasible if $p^* \in (-y, x)$. If it is not feasible, then one of two conditions must hold: either the client is required to transfer too much money to the principal $\beta e^* - x > U^*$, or the principal is required to transfer too much money to the client so that $\beta e^* + y < U^*$.

First consider the case where $\beta e^* - x > U^*$. In words, the client cannot pay for jointly efficient service. There are two sub-cases. The first is where $\beta e^{**} - x > U^*$, so that he is willing to pay $x$ for the principal's preferred level of service. As the principal can extract no more than $x$ and is at his own private optimum, the equilibrium contract is $e^{**}$ with $p = x$. In this case, quality is independent of $\beta$ on the margin, as $\frac{de^{**}}{d\beta} = 0$.

However, if $\beta e^{**} - x < U^* < \beta e^* - x$, the client's IR constraint is violated when all cash is taken at service $e^{**}$ but not at service $e^*$. The first part means that the principal has to increase $e$ from $e^{**}$ and/or reduce $p$ from $x$ to satisfy participation. In this region, the principal chooses $e$ and $p$ to maximize profits subject to $\beta e + p = U^*$, as the IR constraint always binds here.

To see which instrument will be used by the principal, note that the client's indifference curve is characterized by $\beta e + p = U^*$. As a result, by increasing $e$, the principal can reduce monetary transfers by $\beta$. But remember that the principal additionally directly benefits from increasing $e$ by $v - C'(e)$ so that the net return to increasing $e$ on the client's participation constraint is $v + \beta - C'(e)$. If $v + \beta > C'(e)$, the optimal strategy is to increase $e$ rather than reduce $p$ to satisfy participation, and vice versa is if it negative. But $v + \beta - C'(e) > 0$ for $e < e^*$, so that the principal satisfies the IR constraint by first only changing $e$ until $e^*$ is reached, and beyond that by only changing $p$ until $p = -y$. As a result, for $e \in (e^{**}, e^*)$ he chooses

$$\overline{e} = \frac{U^* + x}{\beta}. \tag{12}$$

This continues until $\overline{e}$ reaches $e^*$ or $\beta = \tilde{\beta} = \frac{U^*+x}{e^*}$. In this region $e$ is inversely proportional to $\beta$ as required. If $\beta e^* + y < U^* < \beta e^* - x$, the firm then reduces the price from $x$ to the level required to satisfy participation while holding $e$ at $e^*$. This holds as long as $\beta e^* + y > U^*$.

Finally, consider the case where $\beta e^* + y < U^*$. The only way the principal can induce the client is by changing $e$. Specifically, he chooses $\tilde{e} = \frac{U^*-y}{\beta}$, which is decreasing in $\beta$, as required.

Hence $e$ is increasing in $\beta$ if $p^* \in (-y, x)$, is independent of $\beta$ if either $\beta e^{**} - x > U^*$ or $\beta e^* + y < U^* < \beta e^* - x$, and is decreasing in $\beta$ if either $\beta e^{**} - x < U^* < \beta e^* - x$ and $\beta \leq \tilde{\beta}$, or $\beta e^* + y > U^*$.

**Nash bargaining:** Proposition 1 also holds for small enough allowed payments under Nash Bargaining with non-transferrable utility. Consider the case where the monetary transfer constraint is binding with transfer $\tau \in \{-y, x\}$. Under Nash bargaining, when the transfer constraint is binding, $e$ is chosen to maximize

$$(\beta e - \tau - U^*)(ve - C(e) + \tau) \tag{13}$$

If $U_i$, $i = k, f$, is the net utility of the client and principal, then the optimum is characterized by $\frac{U_k'}{U_k} = -\frac{U_f'}{U_f}$ (where the derivatives are with respect to $e$), which implies that

$$\frac{\beta e - \tau - U^*}{\beta} = \frac{ve - C(e) + \tau}{C'(e) - v}, \tag{14}$$

or

$$e - \frac{ve - C(e) + \tau}{C'(e) - v} = \frac{U^* + \tau}{\beta}. \tag{15}$$

For $\tau$ small, the right hand side is decreasing in $\beta$ so the left hand side must decline in response to equilibrate. The left hand side is increasing in $e$ as $\frac{d(e - \frac{ve-C(e)+\tau}{C'(e)-v})}{de} = 1 + \frac{C''(ve-C(e)+\tau)+(v-C')^2}{(C'-v)^2} > 0$, for $\tau$ small as $v < C'(e)$. Hence $e$ falls with $\beta$ to restore equilibrium as required.

**Proof of Proposition 2:** Actions are taken by three parties: the client, the bureaucrat, and the principal. Consider a possible equilibrium where (i) $\hat{\alpha} = \alpha_a$, (ii) the client only complains when $\alpha_c \neq \hat{\alpha}$, and (ii) the principal investigates with a report $\hat{\alpha} = i$ only if $\kappa \leq \rho v_{jj} + (1-\rho)v_{ii} - max\{\rho v_{ij} + (1-\rho)v_{ii}, \rho v_{jj} + (1-\rho)v_{ji}\}$ and the client complains. This describes the first best.

First consider consider the principal's actions conditional on the equilibrium actions of the other parties. First, the principal never investigates when the two parties agree. If the bureaucrat's signal is $\alpha_a = 0$ and the client disagrees, the return to implementing the bureaucrat's recommendation is $\rho v_{01} + (1-\rho)v_{00}$ while the return to overturning the recommendation is $\rho v_{11} + (1-\rho)v_{10}$. Therefore the bureaucrat's recommendation is overturned if

$\rho \geq \frac{v_{00}-v_{10}}{v_{11}+v_{00}-v_{10}-v_{01}}$. Surplus in that state without an investigation is then $max\{\rho v_{01} + (1-\rho)v_{00}, \rho v_{11}+(1-\rho)v_{10}\}$. Investigation gives rise to efficient allocations so that the principal investigates after a disagreement if $\kappa \leq \rho v_{11}+(1-\rho)v_{00}-max\{\rho v_{01}+(1-\rho)v_{00}, \rho v_{11}+(1-\rho)v_{10}\}$. When the bureaucrat's signal is $\alpha_a = 1$, a similar logic applies and the surplus without an investigation is $max\{\rho v_{10} + (1-\rho)v_{11}, \rho v_{00} + (1-\rho)v_{01}\}$. The principal investigates if $\kappa \leq \rho v_{00} + (1-\rho)v_{11} - max\{\rho v_{10} + (1-\rho)v_{11}, \rho v_{00} + (1-\rho)v_{01}\}$. Hence the principal's actions are a best response.

Then consider the client's action given the other strategies. The relevant issue here is whether there exists a $p$ that guarantees truth telling by the client and also satisfies his participation constraint when $A = 1$. As he has no additional information when his signal agrees, the recommended action will not change even with an investigation, and so it is an equilibrium to not complain when his signal agrees with that of the bureaucrat if $p \in (\underline{\beta}, \overline{\beta})$ (all prices below will be in this interval). Now consider the case where the signals disagree. Here there are a number of cases:

**When a complaint always causes an investigation:** For $\kappa \leq min\{\rho v_{00} + (1-\rho)v_{11} - max\{\rho v_{10} + (1-\rho)v_{11}, \rho v_{00} + (1-\rho)v_{01}\}, \rho v_{11} + (1-\rho)v_{00} - max\{\rho v_{01} + (1-\rho)v_{00}, \rho v_{11} + (1-\rho)v_{10}\}\}$, a complaint always induces an investigation. Then for any $p \in (\underline{\beta}, \overline{\beta})$ truth telling is satisfied as the expected return to the client from complaining is $\rho(p - \underline{\beta}) > 0$ if a recommendation of $\hat{a} = 1$ is made, and $\rho(\overline{\beta} - p) > 0$ if a recommendation of $\hat{\alpha} = 0$ is made.

**When a complaint is always ignored:** For $\kappa \geq max\{\rho v_{00} + (1-\rho)v_{11} - max\{\rho v_{10} + (1-\rho)v_{11}, \rho v_{00} + (1-\rho)v_{01}\}, \rho v_{11} + (1-\rho)v_{00} - max\{\rho v_{01} + (1-\rho)v_{00}, \rho v_{11} + (1-\rho)v_{10}\}\}$, and for a report of $i$, $\rho v_{ij} + (1-\rho)v_{ii} > \rho v_{jj} + (1-\rho)v_{ji}$, the principal always implements the bureaucrat's recommendation. Then truth-telling by the client is weakly a best response for any $p \in (\underline{\beta}, \overline{\beta})$.

**When one complaint is overturned without an investigation:** There are two possible recommendations: begin with the case where for one recommendation, the decision is overturned, but for the other, it is investigated. Take the case of $\hat{\alpha} = 1$ being overturned but $\hat{\alpha} = 0$ being investigated. If a recommendation of $\hat{\alpha} = 1$ is made, the expected return to the client from the decision being overturned is $-[\rho(\underline{\beta} - p) + (1-\rho)(\overline{\beta} - p)]$. To ensure that the client complains here, it must be that $\rho(\underline{\beta} - p) + (1-\rho)(\overline{\beta} - p) \leq 0$. The firm can satisfy

this through appropriate choice of $p \in (\underline{\beta}, \overline{\beta})$ and so truth telling will arise by the client by such choice. This price also guarantees truth telling when $\hat{\alpha} = 0$ is investigated. Now consider the other case where a recommendation of $\hat{\alpha} = 0$ is overturned without investigation (but the other is investigated). The expected return to the client from the decision being overturned when the signals disagree is $\rho(\overline{\beta} - p) + (1 - \rho)(\underline{\beta} - p)$. To ensure that the client complains here, it must be that $\rho(\overline{\beta} - p) + (1 - \rho)(\underline{\beta} - p) \geq 0$, which again can be ensured for a $p \in (\underline{\beta}, \overline{\beta})$. Hence a single price can guarantee truth-telling when one complaint leads to an investigation and the other is overturned.
The other case is where a complaint overturns one decision, but the other is ignored. The same prices as above satisfy truthtelling.

**When any complaint overturns $\hat{\alpha}$ without an investigation:** Finally consider the case where there are never investigations, but both decisions are overturned by a complaint. Then from the above cases, for a single price to satisfy the client's constraint to complain only when his signal disagrees, the price needs to satisfy $p \in (\rho\underline{\beta} + (1 - \rho)\overline{\beta}, \rho\overline{\beta} + (1 - \rho)\underline{\beta})$. If $\rho \geq \frac{1}{2}$ a single price can be found to satisfy both constraints. However, if $\rho < \frac{1}{2}$, there is no price that can do so as $\rho\underline{\beta} + (1 - \rho)\overline{\beta} > \rho\overline{\beta} + (1 - \rho)\underline{\beta}$. However, it can never be the case that the principal overturns both recommendations without an investigation if $\rho < \frac{1}{2}$. To see this, note that an $\hat{\alpha} = 0$ recommendation is overturned without an investigation if $\rho \geq \frac{v_{00} - v_{10}}{v_{11} + v_{00} - v_{10} - v_{01}}$ and an $\hat{\alpha} = 1$ recommendation is overturned without an investigation if $\rho \geq \frac{v_{11} - v_{01}}{v_{11} + v_{00} - v_{10} - v_{01}}$. One of these critical values must exceed $\frac{1}{2}$ and so it can never be the case that the principal overturns both recommendations without an investigation if $\rho < \frac{1}{2}$. As a result, there always exists a price that can induce truth telling when decisions are automatically overturned by a complaint.

Having shown that there exist a price for any of these cases that satisfies truth telling, information revelation by the client can be achieved by a single price $p$.

However, so far I have ignored the client's ability to opt out of consuming the good. This is only relevant when the principal wishes to implement $A = 1$. In any state where there is an investigation, interests are aligned for any $p \in (\underline{\beta}, \overline{\beta})$. As all the prices required above lie in these bounds, this implies that the exit option imposes no additional constraint. They are also aligned with the signals agree. When the signals disagree and there is no investigation, the principal could wish to implement $A = 1$ for either recommendation with a disagreement. When $\hat{\alpha} = 1$, this would require that the principal ignore the client's complaint. The client is only willing to consume $A = 1$ if $\rho(\underline{\beta} - p) + (1 - \rho)(\overline{\beta} - p) \geq 0$. To satisfy this constraint, he chooses a price to satisfy this. (As the complaint is ignored here, there is no need to offer an incentive through price choice for the client to complain.) The final case is where $\hat{\alpha} = 0$ and he disagrees and the decision is overturned. This requires $\rho(\overline{\beta} - p) + (1-\rho)(\underline{\beta} - p) \geq 0$. This is the same constraint as that required to induce complaints and so automatically holds. Finally, if both conditions above need to hold - in the case where the $\hat{\alpha} = 1$ ignores complaints, and the $\hat{\alpha} = 0$ recommendation is overturned without an investigation, the principal chooses a low enough price such that $p \leq min\{\rho\overline{\beta} + (1-\rho)\underline{\beta}, \rho(\overline{\beta} + (1-\rho)\underline{\beta}\}$. Hence a single price can be designed to reveal the client's information as desired.

Finally consider the incentives of the bureaucrat given the strategies above with signal $\alpha_a = i$. If $m_d^*(0) = m_d(^*1) = m_d$ then if the bureaucrat reveals his information truthfully, his utility is

$$U_t = e(1-s)m_d w_1 + esw_0 + e(1-s)(1-m_d)w_0 + (1-e)(1-m_d)w_0. \tag{16}$$

If the bureaucrat deviates, her utility is

$$U_d = em_d 0 + e(1-m_d)w_0 + (1-e)(1-s)m_d w_1 + (1-e)sw_0 + (1-e)(1-s)(1-m_d)w_0. \tag{17}$$

Then

$$U_t - U_d = (2e-1)m_d[(1-s)w_1 + sw_0] + (1-e)(1-s)m_d w_1 \geq 0, \tag{18}$$

as $m_d \geq 0, e \geq \frac{1}{2}$, and $s \geq 0$. Hence the first best is an equilibrium.

Surplus when $\hat{\alpha} = 0$ is then given by $esv_{00} + max\{\rho v_{11} + (1-\rho)v_{00} - \kappa, \rho v_{01} + (1-\rho)v_{00}, \rho v_{11} + (1-\rho)v_{10}\} = V_0$ while surplus when $\hat{\alpha} = 1$ is $esv_{11} + max\{\rho v_{00} + (1-\rho)v_{11} - \kappa, \rho v_{10} + (1-\rho)v_{11}, \rho v_{00} + (1-\rho)v_{01}\} = V_1$. In the first best, surplus $V^* = es\left(\frac{v_{00}+v_{00}}{2}\right) + (1-e+e(1-s))\frac{V_0+V_1}{2}$, which is weakly increasing in $s$.

**Proof of Proposition 3:** Actions are taken by the client, the bureaucrat, and the principal. Consider the incentives of each party.

The principal makes two choices: $m$, and $A$. First note that if an investigation occurs ($m = 1$), the principal chooses $A = \alpha$, as the allocation does not affect the incentives of the bureaucrat so $[A|m = 1] = \alpha$. The remaining issue is $[A|m = 0]$. In an equilibrium where clients are informative, it must be the case that $[A|m = 0] = \hat{\alpha}$ so the bureaucrat's recommendation is not overturned without an investigation. When $\hat{\alpha} = 1$, this is immediate from Assumption 1 as there is no complaint. However, when $\hat{\alpha} = 0$ and client complains only when his signal disagrees, the probability of an error is $\rho$. The ex post optimal allocation is to overturn the recommendation if $\rho \geq \frac{v_{00}-v_{10}}{v_{11}+v_{00}-v_{10}-v_{01}}$. However, if the principal chooses to overturn without an investigation, all clients complain as $\underline{\beta} > 0$, in which case the probability of an error conditional on a complaint is $1-e$. By Assumption 1 it then cannot be optimal to overturn without an investigation, and so $[A|m=0]=\hat{\alpha}$.

Now consider the choice of $m$ given that $[A|m=0]=\hat{\alpha}$ and $[A|m=1]=\alpha$. Begin by identifying optimal ex post oversight. In any equilibrium where complaints are informative, if the client is denied and he does not complaint, the principal never investigates. If the client complains when he is denied and a complaint is informative, $m_d^*=1$ if $\kappa \leq \rho[v_{11}-v_{10}]$. When $\hat{\alpha}=1$, the probability of error is $1-e$, and $m_1^*=1$ if $\kappa \leq (1-e)[v_{00}-v_{01}]$. This defines optimal ex post oversight given the allocative rules outlined above.

Now consider the client. Note that the client's participation constraint always holds here as he gets rents from $A=1$ and so it can be ignored. Consider the client faced with $[A|m=0]=\hat{\alpha}$ and $[A|m=1]=\alpha$, $m_d \geq m_1$. As $m_d \geq m_1$, he will complain if denied and his signal disagrees, and as there is zero probability of the recommendation being overturned if they agree, it is an equilibrium for the client to complain only when the signals disagree.

Finally consider the bureaucrat. Consider a bureaucrat with $\alpha_a=0$, faced with $[A|m=0]=\hat{\alpha}$ and $[A|m=1]=\alpha$, oversight $m_d \geq m_1$, and a client who only complains when his signal disagrees and $\hat{\alpha}=0$. Then $m_t=[e(1-s)+(1-e)]m_d$ is the probability of an investigation if she denies the client. Her expected wage from denying is then $w_1e(1-s)m_d+w_0(1-m_t)$. The return to capitulating to the client by choosing $\hat{\alpha}=1$ is $w_0(1-m_1)$. The bureaucrat only denies if $w_1e(1-s)m_d \geq w_0(m_t-m_1)$. Then an equilibrium exists with optimal ex post oversight where the client is denied in equilibrium only if $w_1e(1-s)m_d^* \geq w_0(m_t^*-m_1^*)$, where $m_t^*=[e(1-s)+(1-e)]m_d^*$.

Now consider the case where $w_1e(1-s)m_d^* < w_0(m_t^*-m_1^*)$. The principal has two choices: to let the constraint be violated or to change oversight to satisfy it. Allowing the constraint to be violated cannot be the optimal choice. To see this, note that if it is violated, the bureaucrat always recommends $\hat{\alpha}=1$ and the client never complaints. As a result, the principal chooses the optimal assignment based on his information alone His three choices are: (i) allocate $A=1$ without an investigation, (ii) allocate $A=0$ without an investigation, or (iii) always investigate. This has surplus $max\{\frac{v_{00}+v_{10}}{2}, \frac{v_{11}+v_{01}}{2}, \frac{v_{11}+v_{00}}{2}-\kappa\}$. But the principal can achieve at least as high surplus by having the bureaucratic constraint bind. To show this, I need only show one case where it holds and achieves higher surplus. Hence, consider the case where the principal ignores complaints, and chooses $m_0^*=m_1^*$. At $m_0^*=m_1^*$, the bureaucratic constraint always holds, and the principal now gains the information of the bureaucrat. His choices are - from Assumption 1 - to allocate according to the recommendation of the bureaucrat or to investigate. This yields surplus of $max\{\frac{ev_{00}+(1-e)v_{10}}{2}+\frac{ev_{11}+(1-e)v_{01}}{2}, \frac{v_{11}+v_{00}}{2}-\kappa\}$, which (weakly) exceed that from (5) being violated.

Given this, the principal chooses the cheapest way to satisfy the bureaucratic constraint. In the baseline case of $m_d^* = 1$, and $m_1^* = 0$, he has two instruments to do so. First, he could ignore complaints by the client and choose $m_0 = m_1^* = 0$. If he does, no information of the client is used, and surplus is $\frac{ev_{00}+(1-e)v_{10}}{2} + \frac{ev_{11}+(1-e)v_{01}}{2}$. Alternatively, he can choose $m_1 = m_d^* = 1$ in which case the bureaucratic constraint is $e(1-s)w_1 + w_0 es \geq 0$, which always holds. The surplus from choosing $m_1 = 1$ is given by $\frac{v_{11}}{2} + \frac{v_{00}}{2} - \frac{\kappa}{2} - [(1-e)+e(1-s)]\frac{\kappa}{2}$. The principal then optimally ignores complaints if $\kappa \geq \frac{1-e}{1-e+e(1-s)}[v_{11}-v_{10}+v_{00}-v_{01}]$. Proposition 3 follows.

**Proof of Proposition 4:** The actions of the client and the principal are exactly as in Proposition 2. Consider the incentives of the bureaucrat. If the bureaucrat reveals his information truthfully, his utility is

$$U_t = e(1-s)m_d w_1 + esw_0 + e(1-s)(1-m_d)w_0 + (1-e)(1-m_d)w_0. \tag{19}$$

If the bureaucrat deviates, her utility is

$$U_d = e(1-m_d)w_0 + (1-e)(1-s)m_d w_1 + (1-e)sw_0 + (1-e)(1-s)(1-m_d)w_0. \tag{20}$$

Then

$$U_t - U_d = (2e-1)m_d[(1-s)w_1 + sw_0] \geq 0, \tag{21}$$

as $m_d \geq 0, e \geq \frac{1}{2}$, and $s \geq 0$. Hence the first best is an equilibrium.

Now consider the case where prices are constrained and whether the bureaucrat with a signal $\alpha_a = 0$ denies the client. Her expected wage from denying is $w_1 e(1-s)m_d + w_0(1-m_t)$ as above. The return to capitulating to the client by choosing $\hat{\alpha} = 1$ is now higher at $(1-e)m_1 w_1 + w_0(1-m_1)$ as he is now rewarded if his lie turned out to identify the correct state. The bureaucrat therefore only denies if

$$w_1(e(1-s)m_d - (1-e)m_1) \geq w_0(m_t - m_1) \tag{22}$$

In the case where $m_1 = 0$ this has no effect on the analysis above. However, when $m_1 = 1$, the bureaucratic constraint is now given by

$$e(1-s)w_1 + w_0 es \geq w_1(1-e). \tag{23}$$

Unlike the case above, this needs not always hold. However, when wages are sub-linear in probabilities, this amended constraint always holds.

There is one other deviation that must be considered here: where the bureaucrat denies the client even through she believes that the client should be given this good. This is is only relevant when $m_d = 1$ and $m_1 = 0$, and the constraint that the bureaucrat honestly gives the good to the client is given by $(1-e)(1-s)w_1 \leq w_0$, which holds if wages are sub-linear in probabilities. Hence similar insights arise when investigations cannot directly penalize lying.