

## **The More Closely we are Watched, the Better we Behave?**

Andrea Prat  
London School of Economics

November 2005

Transparency is a complex issue, which can be approached from many angles and with various conceptual tools. This contribution provides a brief survey of the economic literature on transparency. The conceptual tool that economists use is the *principal-agent model*, a game-theoretic setting where a principal (the citizens, shareholders, etc) wants an agent (the government, the CEO, etc.) to perform a certain task. In this setting, transparency corresponds to the ability of the principal to observe what the agent does. We can then ask what happens when transparency improves.

The survey begins with the key theoretical result in this area. Holmstrom (1982) provides a powerful and general rationale for full transparency. More information about the agent's behavior makes the agent more accountable and more likely to work for the common good. The rest of the contribution discusses a long list of objections to this result, from both a conceptual and a practical viewpoint. Given that Holmstrom asserts that transparency improves accountability, possible criticism belongs to two broad classes. First, one can argue that the increase in accountability is not sufficient to offset other drawbacks such as the violation of privacy, the direct cost of disclosure, or the revelation of sensitive information. Alternatively, one can attack the link between transparency and accountability: it is not necessarily true that more disclosure makes the agent behave better. We will discuss both classes of objections, with a particular emphasis on the latter because it can be analysed within the principal-agent setting. The survey concludes with a tentative appraisal of the validity of the Holmstrom benchmark and its relevance for politics and corporate governance.

### **The Principal-Agent Model**

The economic model that is most useful to analyse transparency comes from game theory. There is a *principal* (conventionally, she), who requires a certain service but does not have the time or the ability to take care of it directly. She therefore enters a contractual relation with an *agent* (he), who can potentially provide the service. First, the principal offers a contract to the agent, which the agent accepts or rejects. The contract is designed by the principal and it specifies a payment from the principal to the agent, which may be contingent on observable outcomes. Once the agent knows the contract, he takes an action (or a series of actions) on behalf of the principal. The action may involve effort or other forms of disutility on the part of the agent, and it produces, perhaps stochastically, an observable outcome which will then determine the agent's compensation on the basis of the contract in place.

To make things more concrete, let us look at a simple principal-agent relation. The principal is a retail investor who is looking to invest her savings in the stock market, but does not have the time and the inclination to do it personally. The agent is a mutual fund manager who can make investment decisions on behalf of the investor. In principle, the investor can design a variety of contracts. For instance, she could make the payment to the fund manager contingent on the return of the fund. One should expect that the willingness of the fund manager to spend effort and resources in trying to find the right investment strategy depends on the kind of monetary incentives that he is offered.

As we can see from the table below, the principal-agent model applies to a number of interesting economic relations.

<i>Principal</i>	<i>Agent</i>	<i>Action</i>
Car owner	Mechanic	Repair car
Patient	Doctor	Perform tests, choose therapy
Employer	Worker	Perform job
Shareholders	CEO	Manage company
Investors	Fund manager	Choose portfolio
Citizens	Government	Choose policy

Typically, the principal does not observe the details of the agent's action. Holmstrom (1979) assumes that the principal can observe a list of variables, and that the contractual payment can be contingent on those variables. The transparency result is then immediate: *adding one more observable variable can never hurt the principal*. The principal is free to make the payment contingent on that additional variable, and she will do so if and only if it increases her net expected utility (expected benefit from the agent's action less expected payment to the agent).

Holmstrom makes this result more precise. He provides a mathematical condition on the information provided by the additional variable such that the principal is strictly better off if and only if such condition is satisfied. Basically, an additional variable is useful if it provides more information on the action that the agent takes.

The principal benefits from more transparency, but can the agent be hurt? In general, the agent benefits from the presence of an informational asymmetry between himself and the principal. His ability to keep his action hidden forces the principal to offer high monetary payments for positive outcomes. If the informational asymmetry reduces, the principal may be able to reduce payments. In the limit case, where the agent's action is perfectly observable, the agent will only be paid a wage that compensates for the disutility of effort.

However, the expected cost to the agent is never larger than the expected benefit to the principal. For every improvement in transparency, there exists a monetary amount such that, if the improvement is accompanied by the payment of that amount from the principal to the agent, then both the principal and the agent are better off.

### **The Full-Transparency Principle in Practice**

In real life, we observe systematic deviations from full transparency in several areas. Take the fund management example discussed above. Investors are typically supplied with limited information on the composition of the fund they own. Currently, the US Securities and Exchange Commission requires disclosure every six months, which consists of a portfolio snapshot at a particular point in time and can easily be manipulated by re-adjusting the composition just before and after the snapshot is taken – a well-documented practice known as *window dressing*. It would be easy and almost costless to have more frequent disclosure by requiring mutual funds to publicize their portfolio composition on the internet. Yet there is strong resistance from the industry to proposals in the direction of more frequent disclosure (Tyle, 2001).

In corporate governance, violations to the transparency principle are so widespread that some legal scholars argue that secrecy is the norm rather than the exception in the relation between shareholders and managers (Stevenson, 1980): “Corporations – even the largest among them – have always been treated by the legal system as ‘private’ institutions. When questions about the availability of corporate information have arisen, the inquiry has typically begun from the premise that corporations, like individuals, are entitled to keep secret all information they are able to secure physically unless some particular reason for disclosure [...] could be adduced in support of a contrary rule. So deeply embedded in our world view is this principle that it is not at all uncommon to hear serious discussions of a corporate ‘right to privacy’.

In politics, the principle of open government has made great inroads in the last decades, but there are still important areas in which public decision-making is, by law, protected by secrecy. For instance, all Freedom of Information Acts include circumstances in which the government can legally withhold information from citizens. Similar provisions exist in other countries. As recent events related to the Iraq war have shown, democratic governments can be expected to claim a need for secrecy in extremely controversial circumstances.

Why is there not more transparency in practice? The Holmstrom Principle may be a bad predictor of actual policy for two orders or reasons. First, full disclosure may indeed be the right thing to do from a normative viewpoint, but there may be strong entrenched interests that prevent more transparency. As we saw earlier, the agent stands to lose some of his rents from an increase in disclosure and he should be expected to fight against a change in policy. There is no doubt that such a struggle exists in all the examples seen above. Principals (investors, shareholders, citizens) have been asking for more transparency for decades and have encountered varying degrees of resistance from their respective agents. Typically, the devil is in the detail: agents pay lip service to the principle of openness but then make it onerous or impossible for the principal to get hold of the most sensitive information.

However, there is a second class of explanations. Full disclosure may not be the optimal policy, at least in certain cases. One can think of a number of reasons:

1. A certain degree of privacy is a basic human right. While this argument is important for certain principal-agent relations, it does not seem to be central to the points debated in finance, politics and governance.

2. There may be a direct cost of information disclosure. While in the past, disclosure occurred through paper publication, the internet has greatly reduced the direct cost of making information public. It is difficult to imagine that direct cost plays a determinant role in most cases of interest.
3. The agent may be unable to communicate a certain piece of information to the principal without other hostile parties learning of it as well. The leading example of this “third-party” rationale for secrecy is of course military intelligence. The British government cannot keep the British people fully informed about security operations without divulging these details to potentially hostile organisations. Similarly, a company that follows a full-disclosure policy may end up revealing non-patented information to competitors, or a fund manager who publicises his trading strategy may be “front-run” by other funds.

This is potentially a strong rationale for secrecy. However, in many of the instances under debate there is a simple way around it. Information should be fully disclosed, but with a delay that is sufficiently long to make it of little value to hostile parties (and sufficiently short to allow the principal to keep the agent accountable). For instance, the SEC has proposed that mutual funds make the daily composition of their portfolios public, but with a 60-day lag.

4. Sometimes agents refuse to disclose information to principals on the ground that the principals would not evaluate it “in the right way”. This argument is usually supported by the observation that investors/shareholders/citizens are not fully rational. CEOs, especially European ones, keep complaining about investors’ short-termism. For instance, the recent ousting of Werner Seifert from Deutsche Borse has prompted an avalanche of accusations against allegedly myopic investors.

This line of defence of secrecy may have some truth in it, but the argument that the principal is not fully rational is insufficient. The agent needs to argue that the errors that the principal could make are more damaging (to the principal) than the errors (or voluntary damage) that the agent himself can make. There is no evidence to support the thesis that principals mis-use information in a way that is very costly to themselves. Instead, there is a large body of evidence that insufficiently monitored agents can ruin their principals. With more transparent accounting systems, disasters like Equitable Life, Enron and Parmalat could have been avoided. Viewed in this light, the “principal’s irrationality” argument lacks plausibility. While corporate leaders appear to stick to this line of defence, most democratic governments seem to have abandoned it, at least because a government that accuses its citizens of irrationality would face a serious re-election challenge.

5. Finally, excessive transparency may create incentives for the agent to behave in ways that damage the principals. If that is the case, the principal would want to commit not to observe certain variables in order to avoid the damaging behaviour. This line of argument is the most subtle, and potentially the most convincing. We devote the next section to it.

## Can Excessive Transparency Make the Agent Behave Badly?

We now go back to the principal-agent model to make one important modification. In Holmstrom the principal can make her payment contingent on every observable piece of information. In other words, contracts are *complete*. Every observable variable can become part of the contractual arrangement. In practice, this is not the case: contracts are far from complete. In politics, if contracts were complete, citizens could offer the prime minister a monetary payment contingent on all kinds of observable variables, like the GDP growth rate or health indicators. This seldom happens in practice. A typical high-ranking official is simply offered a fixed salary independent of any outcome. Similarly, in the US most mutual fund managers are offered a percentage of funds under management, independent of the return they generate for their investors (this is due to SEC restrictions, which explain why hedge funds are usually incorporated outside the US).

The fact that contracts are not directly contingent on outcomes means that the agent must be kept accountable in a different, and more complex, way. The principal has an implicit threat: if she is unhappy with the agent, she can replace him with another agent. Citizens can vote for another party; shareholders can oust a CEO; and investors can take their money to another fund. This creates a dynamic game between the principal and the agent, often referred to as the *career concerns model*, and was first discussed by Holmstrom (1999).

We assume that the contract between the principal and the agent involves a flat fee for a given term (one could complicate the model without changing the results substantially). At the end of the term, the principal can either keep the agent (and offer a flat fee again) or replace him with a new agent (and offer a flat fee to him). For simplicity, suppose that the world ends after two terms (but see Holmstrom (1999) for an infinite-horizon formulation). The agent is characterized by a *type*, that is, a set of personal characteristics that determine his ability, vision, willingness to put effort, intrinsic honesty, political preferences, etc.

In the second term, the agent is unaccountable and his behaviour is dictated purely by his type. A lazy agent would put no effort, a dishonest one would steal, etc. The principal cannot observe the agent's type directly, but she tries to infer it from the agent's action in the first period. If she then feels that the agent's type is "bad", she will replace him before he can do further damage in the second term. The agent wants to keep his job and he realizes that the principal is watching him. He has an incentive to behave in the first term in a way that will convince the principal that he is a "good" type. For instance, a lazy agent may want to look active in the first term in order to hide his laziness from the principal. In equilibrium (a *perfect Bayesian equilibrium*), the principal has expectations on how the agent behaves, and such expectations are correct.

Career concerns models generate rich and interesting equilibrium behaviour on the part of the agent. From our viewpoint, they can be used to study the role of transparency. The question is: what is the effect of letting the principal observe an additional variable at the end of the first term, before she decides to keep or replace the agent?

The answer depends on what the type of the agent is. As we shall see, two cases yield starkly different results. If the agent's type is his ability to exert effort, then more transparency benefits the principal, because it induces the agent to exert more effort in the first period and it allows the principal to identify and fire bad agents. If instead the type is the agent's ability to make the right decision, then more transparency can create an incentive for the agent to behave in a conformist way, which is highly damaging to the principal.

Let us begin with the case where the type is the ability to exert effort, developed in Holmstrom (1999). With small modifications, this case also encompasses situations where the type represents laziness (the lazy type is the one who experiences more disutility from work) and dishonesty (the dishonest type is the one who experiences more utility from stealing). At the end of the first period, the principal receives a signal on the output produced by the agent. The output is given by a sum of the agent's innate ability, the effort exerted by the agent, and a stochastic component. Holmstrom shows that, when the signal becomes more accurate (the variance of the stochastic component increases), the agent puts more effort in the first period and the principal's information on the agent's type becomes more precise.

Holmstrom's result is intuitive. I know that my boss is watching my output to understand whether I am an able worker. The quality of my output depends on how able I am, how hard I work and some factors that are not under my control, but the boss cannot distinguish between these three components. If she observes a high-quality output, she will ascribe part of this success to my innate ability and she will be less likely to fire me. However, if the factors not under my control play an important role, a high-quality output is more likely to be due to luck rather than my effort, and the boss will not give me a lot of credit for it. The more important the luck component, the less incentive I have to work hard to impress my boss. Conversely, if the factors that are not under my control are negligible, the channel that links my personal characteristics to the final output is less noisy and I have a strong incentive to work hard. Moreover, the principal will be able to form a more precise view on my ability after she observes my output and she will be more likely to fire (or reward) me when I am a bad (good) worker.

Holmstrom's result is not universal, even when the agent's type is his ability to exert effort. There are cases where an increase in transparency leads to less, rather than more, effort on the part of the agent. An example is provided by Dewatripont, Jewitt and Tirole (1999). Suppose a teacher is trying to maximize the effort that her students put in coursework. At the end of the course, all students take an exam which provides a mark between 0 and 100. The score is observed by potential employers who use it to evaluate the expected on-the-job productivity of students (which is assumed to be correlated to their scholastic productivity). Hence, the higher the score, the higher the wage that the student can expect to receive. Dewatripont and his colleagues show that, under certain circumstances, the teacher may elicit more effort from her students if she commits to revealing a pass/fail mark rather than the complete score. That is, the principal sets a threshold score and she only communicates to employers whether a student is above or below the threshold. If the threshold is chosen carefully, a pass/fail mark acts a strong incentive for the bulk of students to exert high effort in order not to be lumped with the fail group. However, the authors also show that the counter-

example is somewhat special. With mild regularity conditions on the information structure, such a situation does not arise and more disclosure induces the agent to put more effort.

Let us now turn to the other case, when the agent's type is his ability to make the right decision. This assumption may be a more accurate representation of agents who are entrusted with high-level tasks, such as prime ministers, CEO's, and fund managers, while the ability to exert effort is more important for jobs that involve less responsibility. Specifically, we assume that there is a stochastic *state of the world*, which the principal does not observe directly (What is the effect of different policies? What is the effect of different business strategies? What stocks are going to go up next year?). The agent's task is to observe the state of the world and take the appropriate decision (Adopt the right policy or business strategy; buy promising stocks). The ability of the agent consists in the precision with which he observes the state of the world: a *smart* agent is one who receives an accurate signal.

In this set-up there are two classes of variables that the principal can observe: decisions (policies, business strategies, portfolios) and outcomes (statistics, profits, returns). Prat (2005) studies the effect of increased disclosure over these two classes of variables. Transparency over outcomes is uncontroversial: it plays the same role as transparency in Holmstrom (1999), inducing the agent to use his information efficiently and enabling the principal to screen agents more accurately. However, transparency over decisions may have detrimental effects. If the agent knows that his decision is scrutinized, he may choose to disregard his private information about the state of the world in order to look smart.

Consider mutual funds. Suppose that a smart fund manager has private information about future stock returns (that the rest of the market does not possess), while a dumb one has no private information. Under standard assumptions, the optimal decision for a dumb manager is to select a neutral portfolio (basically, one that reflects the composition of the market). Instead a smart manager should deviate from the neutral portfolio to weight his holdings in favour of stocks he deems promising. This can be the equilibrium behaviour if returns are observed but portfolio decisions are not. Instead, if portfolio composition is observed as well, we have a contradiction. A dumb fund manager who chooses a neutral portfolio is immediately revealed for what he is, and will lose his investors. Hence, he has an incentive to pretend that he is a smart manager by deviating from the neutral portfolio. In equilibrium, all fund managers – dumb and smart – deviate from the neutral portfolio. Smart managers are more likely to obtain high returns, but the dumb ones too may be lucky. In this case, the investor may gain from committing to observing the portfolio return only, in order to ensure that dumb managers behave optimally. Dasgupta, Prat and Verardo (2005) analyse data on US institutional investors from 1984 to 2002 and provide evidence of widespread conformist behaviour on the part of institutions: institutions tend to buy (sell) stocks that other institutions have bought (sold) in the recent past and such stocks do worse (better) than the rest of the market in the medium term.

The logic of the example can be extended to politics and business. Suppose that, faced with a crisis, a leader can adopt a radical solution or keep the status quo. A smart leader is more likely than a dumb leader to understand the cause of the crisis and to identify the right radical solution. If leaders behaved optimally, smart ones would

adopt radical policies more often than dumb ones would. But from the argument seen above, this is not an equilibrium. Leaders who keep the status quo would be perceived as more likely to be dumb and would be replaced. Hence, every leader, dumb or smart, tends to adopt radical policies even when they are not necessarily optimal.

Prat (2005) also shows that transparency over decisions goes hand in hand with transparency over outcomes. If the principal can observe the output in an accurate and timely manner, then disclosing the agent's decision has less negative effects because his desire to behave in a conformist manner is tempered by the risk that his decision turns out to be wrong. If instead final outcomes are difficult to observe, the principal runs the risk of facing extremely conformist agents and will want to reduce or delay disclosure of the agent's decision. This rationale is enshrined in one of the most important exceptions to the open government principle: the exemption of pre-decision information. These are documents that are prepared at an early stage, before a policy is formulated and hence implemented. If such documents were made public, citizens could observe nothing about outcomes and there would be a serious risk that advice is less candid or forthcoming.

We conclude this section by discussing another reason why in career concerns models transparency may have detrimental effects. This happens when there are multiple principals and multiple agents. The leading example is the European Central Bank (ECB). Members of the Governing Council are selected by national governments. While nominally members should only be concerned by the interests of the whole Eurozone, in practice the future of their careers depends on their ability to further national interests. There has been an intense debate on whether the minutes of the meetings of the Council should be made public (see, for instance, the exchange between Buiters (1999) and Issing (1999)). At this stage, they are still secret, and there is a strong rationale behind such a policy. If the discussions at meetings were public, it is feared that national members would have an incentive to pander to their home audiences by taking adversarial stances, which would make the decision-making process slow and cumbersome. The ECB disclosure policy is in contrast with that of other central banks, like the Bank of England and the Federal Reserve Bank, who do not have a trans-national composition and therefore do not face the same risk. Both those institutions have chosen a much higher level of transparency on their monetary policy decision process.

The multiple-agent rationale for secrecy can be extended to other trans-national organisations. The minutes of the Council of the European Union are usually kept secret. This is in sharp contrast with the legislative bodies of most of the member countries which hold open-door meetings. Multilateral organisations like the World Bank and the International Monetary Fund have been criticised for the lack of transparency in their decision-making process. While the multiple-agent argument is applicable to all these cases, it is clear that the policy on transparency is only a small part of the more general, and extremely controversial, issue of what the mandate of such multilateral organisations should be. It is impossible to form a view on the optimal level of transparency of a particular body before we decide to whom and in what forms the body should be accountable.

## **Conclusions**

On the issue of transparency, economics provides a strong argument in favour of complete disclosure and a list of potential objections to that argument. As Holmstrom showed, in a world of complete contracts, the more the principal knows about the agent, the better the agent behaves. Some objections to Holmstrom – the right to privacy, the direct cost of disclosure, the risk that hostile parties learn sensitive information – are perfectly valid but they find limited application in politics, corporate governance and other important areas. Others object that information should not be disclosed because there is a risk that people may misunderstand it or misinterpret it, but there is no sign that such risk is comparable to the well-documented damages that secrecy has created in the past.

A potentially more relevant class of objections arises when we move to a world of incomplete contracts. A career concerns model applied to agents whose job is to take decisions on behalf of their principals shows that a certain kind of transparency creates perverse incentives. An agent who knows that his decisions are closely scrutinised before the consequences of the decisions are fully known has an incentive to behave in the way that the principal expects him to behave rather than in the optimal way. Another interesting case arises when there are multiple principals and multiple agents. However, the scope of this kind of objections appears to be limited as well. The only two (somewhat) documented cases where transparency may be detrimental to the principal are the disclosure of portfolios of mutual funds and the publication of the minutes of the ECB Governing Council.

In sum, there are a number of potential objections to the full-transparency principle, but at this stage they all appear to have a limited scope of application. Available economics research supports the idea that full transparency should be the default option. In politics, this principle has been accepted with the passage of the Freedom of Information act. In corporate governance, we are extremely far from it: the default option is secrecy. The law allows companies to withhold all information from stakeholders, except a small class of aggregated accounts. The same can be said about international organizations as well as most of the non-profit sector. Now that the principle of open government has been accepted, the next frontier is open governance.

## References

Buiter, W. Alice in Euroland. *Journal of Common Market Studies* 37(2): 181-209, 1999.

Dasgupta, Amil; Prat, Andrea and Verardo, Michela. The Price of Conformism. Working Paper, London School of Economics, 2005.

Dewatripont, Mathias; Jewitt, Ian and Tirole, Jean. The Economics of Career Concerns, Part I: Comparing Information Structures. *Review of Economic Studies*, January 1999, 66(1), pp. 183-198.

Issing, O. The Eurosystem: Transparent and Accountable, or Willem in Euroland, *Journal of Common Market Studies* 37(3): 503-519, 1999.

Holmstrom, Bengt. Moral Hazard and Observability. *Bell Journal of Economics*, Spring 1979, 10, pp. 74-91.

Holmstrom, Bengt. Managerial Incentive Problems: A Dynamic Perspective. *Review of Economic Studies*, January 1999, 66(1), pp. 169-182 (a working version of this article was first circulated in 1982).

Prat, Andrea. The Wrong Kind of Transparency. *American Economic Review*, forthcoming, 2005.

Stevenson, Russell B. Jr. *Corporations and Information: Secrecy, Access, and Disclosure*. Baltimore: Johns Hopkins University Press, 1980.

Tyle, Craig S. Letter to the SEC on the Frequency of Mutual Fund Portfolio Holdings Disclosure. Investment Company Institute, 2001.