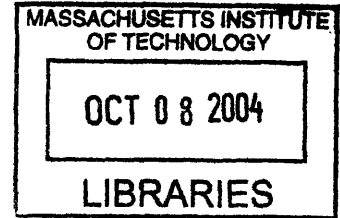


Essays in Macroeconomics, Corporate Finance,
and Social Learning

by

Andrew C. P. Hertzberg

B.Ec. (Hons), University of Sydney (1997)
M.A., Queen's University (1998)



Submitted to the Department of Economics
in partial fulfillment of the requirements for the degree of

ARCHIVES

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2004

© 2004 Andrew C. P. Hertzberg. All Rights Reserved.

The author hereby grants to MASSACHUSETTS INSTITUTE OF TECHNOLOGY
permission to reproduce and
to distribute copies of this thesis document in whole or in part.

Signature of Author

Department of Economics
August 15 2004

Certified by

Ricardo J. Caballero
Ford International Professor of Economics
Thesis Supervisor

Certified by

K. Daron Acemoglu
Charles P. Kindleberger Professor of Applied Economics
Thesis Supervisor

Accepted by

Peter Temin
Elisha Gray II Professor of Economics
Chairperson, Department Committee on Graduate Students



**Essays in Macroeconomics, Corporate Finance,
and Social Learning**

by

Andrew C. P. Hertzberg

Submitted to the Department of Economics
on August 15 2004, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

This thesis consists of three essays at the intersection of macroeconomics, corporate finance, and social learning. The underlying theme which links these three chapters is the study of private incentives to gather and release information and their impact upon the way society learns as a whole. Chapters 1 and 2 focus upon the incentives within a firm to distort the way a firm's performance is perceived publicly. Chapter 1 focuses upon the case of a manager whose incentives encourage her to take actions which distort the observed performance of the firm. In that context the firm's owners tolerate distortions so as to gain from high powered incentives. In chapter 2 I argue that the owners of a firm may have it in their interest to distort the information a firm releases so as to create uncertainty and thereby sponsor speculation on the value of the firm. Chapter 3 considers a scenario where an intelligence agency relies upon the information provided by citizens not under their direct control. A common feature of each chapter is that private decisions about information release are made without considering their full social value.

Turning to specifics, Chapter 1 presents a theory of gradual booms and rapid recessions that is motivated by the experience of the US economy over the last decade. The dynamics of the economy are driven by the speed at which agents learn about the unobserved aggregate state. Agents learn from the observed performance of firms. Each firm is controlled by a manager whose compensation is based upon the observed performance of their firm. If a manager is given short-term incentives, she will try to hide weak short-term performance. This makes the short-term performance of a firm less informative for the aggregate state, delaying the release of information until the long-term results of the firm are realized. In equilibrium, when the belief about the aggregate state is high, managers will be given short-term incentives, delaying the release of information. When the belief about the aggregate state is low, long-term incentives will be prevalent and information will be released without delay. This produces asymmetric learning dynamics for the economy, with gradual booms and rapid recessions. In a boom the belief about the aggregate state increases, information is pushed off into the future, and learning is slow. In a recession the belief is falling, triggering a switch to long-term incentives, that brings forward the release of information and accelerates learning.

Chapter 2 presents a model of corporate misreporting in an environment where investors have heterogeneous beliefs and short sale constraints. The disagreement between investors provides a motive for agents who start a firm to limit the amount of information which it releases

to the public so as to sponsor speculation over its value. This incentive to limit information is stronger when the heterogeneity of beliefs among investors is stronger. Investors also learn about a firm's expected profitability from the information released by other firms in the industry. I show that this creates a strategic complementarity in the precision of information released by each firm. This can give rise to multiple equilibria: one in which all firms release precise reports and one in which their reports are inaccurate.

Chapter 3 (joint work with Professor Christophe Chamley from Boston University) is motivated by the fact that information gathering agencies often rely upon reports made to them by agents not directly under their control. The police, for example, rely upon sightings made by members of the public in the course of investigating a crime. If agents have some choice as to what type of information they can look-out for then this gives rise to a principal-agent problem. Chapter 3 considers an environment in which agents choose to look where they think they are most likely to make a positive sighting. The police can influence the beliefs of the agent by the information they release. We show that when the police have weak information (flat prior) then it is optimal to hide this from the public so as to encourage random search. Conversely when the police have strong information which leaves them to choose between a few likely suspects then they will release their findings so as to direct search away from areas they have already ruled out.

Thesis Supervisor: Ricardo J. Caballero
Title: Ford International Professor of Economics

Thesis Supervisor: K. Daron Acemoglu
Title: Charles P. Kindleberger Professor of Applied Economics

Acknowledgements

I would like to thank my advisors Ricardo Caballero and Daron Acemoglu for their generosity, patience, understanding and guidance throughout this thesis. I have learnt so much from them. I cannot express how grateful I am to you both. Every time I was sure my work had gone off the rails you would have me back on the tracks and heading in the right direction.

Special thanks are also due to Bengt Holmstrom whose support and advice have helped me over the course of writing this thesis. I also thank Christophe Chamley my coauthor on the third chapter in this thesis. His insight and enthusiasm have been an invaluable source of inspiration. I have also been privileged to benefit from many of the faculty at MIT. In particular Paul Joskow, Glenn Ellison, Marios Angeletos, Ivan Werning, Olivier Blanchard, Peter Temin, Jean Tirole, Mathias Dewatripont, and Xavier Gabaix. My deepest thanks go to each one of you. I thank Don Wright who cared enough to encourage me to study overseas. I thank the faculty at Queen's University, in particular Ruqu Wang and Gregor Smith, who prepared me for my studies at MIT. I also thank the high school teachers I encountered who really cared: Brother Dominic, Robert and Maureen Stevenson, Adrian Wells, and Mrs Curran.

I dedicate this thesis to my parents, Ralph and Teresa. They have worked hard their whole lives to provide me with opportunities they never had. I have always had their gentle support throughout my studies: my father's example of reason and respect for learning; and, my mother's passion and unrelenting confidence in all of her sons. Every time I give a talk you are both there with me. You are the best teachers I have ever had.

My brothers have both been a tremendous support. David and I left Australia to study in the U.S. together and Mark has now taken over at MIT as I leave. I hope we can all watch the cricket together again soon. I am so proud of you both.

Throughout my studies I have also benefited from the love and inspiration of my Nan (Nora Roberts) my Aunty Margaret and my late Grand Parents (Gertrude and Berthold Hertzberg).

Thanks are also due to my old friends Eamonn, Chrisostomos, Carlos and Kim. Every phone call, letter, and email has helped keep me sane. I have missed each of you very much. I might never have thought to go to graduate school were it not for Chris. Part of my Ph.D. belongs to him.

At MIT I have been privileged to study with many amazing people. These friendships have

made the last five years so much richer, both academically and personally. While I cannot name everyone I would particularly like to thank James Vickery, Daniel Paravisini, Oscar Landerretche, Alexis Leon, Andy Healy, Bjoern Brugeremann, Andrew Sweeting, Ashley Lester, Nancy Qian, Jon Kearns, Stavros Panageas, Pol Antras, Lucia Breierova, Rita Ramalho, Nada Mora, and Rebecca Zarutskie. I have drawn enormous strength from you all.

Finally I thank Cheryle Steel. I remember standing with Cheryle in the sun room in Culburra and looking out over the Pacific Ocean. North America seemed a long way away, especially for two people who had never left Australia. We were the only people crazy enough to think I could do all this. Walking the streets of Mortdale together, we would talk for hours with unbridled enthusiasm about “the big US schools”. Cheryle reminded me that T. S. Elliot was a banker and helped me to see the passion in economics. I wish I could go back and tell those two crazy kids they were right. I am sorry this has been so hard for us.

Contents

1	Managerial Incentives, Misreporting, and the Timing of Social Learning: A Theory of Slow Booms and Rapid Recessions	11
1.1	Introduction	12
1.2	The Model	15
1.2.1	The Economy	16
1.2.2	The Firm's Problem	17
1.2.3	Contract Design	19
1.3	Learning Dynamics	24
1.3.1	Learning From the Firm	24
1.3.2	Laws of Motion for the Belief	27
1.3.3	The Phases Of Learning	29
1.3.4	The Speed of Learning	32
1.3.5	Direct Learning	33
1.3.6	Delayed Learning	34
1.3.7	Slow Learning	38
1.3.8	Intense Learning	40
1.4	Stock Market and Investment Dynamics	42
1.4.1	Stock Market	42
1.4.2	Investment	50
1.5	Robustness	52
1.5.1	Relaxing the Additivity Assumption	52
1.5.2	Contracting on Aggregate Variables	55

1.5.3	Endogenizing a	57
1.6	Related Literature	60
1.7	Conclusion	62
1.8	Appendix	63
1.8.1	Proof of Proposition 1 and Proposition 10	63
1.8.2	Proof of Proposition 4	68
1.8.3	Proof of Proposition 7	68
1.8.4	Proof of Proposition 12	69
1.8.5	Proof of Proposition 13	70
2	Speculative Trading and Corporate Misreporting	77
2.1	Introduction	77
2.2	Single Firm Problem	80
2.2.1	The Environment	80
2.3	Equilibrium	83
2.3.1	Equilibrium Prices	83
2.3.2	Founder's Problem	85
2.3.3	Founder's Solution	86
2.3.4	Changes in the Heterogeneity of Beliefs	87
2.3.5	Changes in Outside Information τ_s	91
2.4	Endogenizing the External Signal	92
2.4.1	Equilibria	94
2.5	Conclusion	100
2.6	Appendix	100
2.6.1	Proof of Proposition 14.	100
2.6.2	Proof of Proposition 15.	101
2.6.3	Proof of Proposition 16.	101
2.6.4	Proof of Proposition 17.	102
2.6.5	Proof of Proposition 18.	103
2.6.6	Proof of Proposition 19.	103
2.6.7	Proof of Proposition 20.	104

3	White Vans: Optimal Information Release in the Presence of Decentralized Signal Choice	107
3.1	Introduction	107
3.2	The Model	109
3.3	Optimal Signal Choice	111
3.4	Decentralized Signal Choice and Information Policy	113
3.4.1	Discussion	114
3.5	Conclusion	115
3.6	Appendix	115
3.6.1	Proof of Proposition 23	115
3.6.2	Proof of Proposition 24	116
3.6.3	Proof of Proposition 25	117

Chapter 1

Managerial Incentives, Misreporting, and the Timing of Social Learning: A Theory of Slow Booms and Rapid Recessions

This chapter presents a theory of gradual booms and rapid recessions that is motivated by the experience of the US economy over the last decade. The dynamics of the economy are driven by the speed at which agents learn about the unobserved aggregate state. Agents learn from the observed performance of firms. Each firm is controlled by a manager whose compensation is based upon the observed performance of their firm. If a manager is given short-term incentives, she will try to hide weak short-term performance. This makes the short-term performance of a firm less informative for the aggregate state, delaying the release of information until the long-term results of the firm are realized. In equilibrium, when the belief about the aggregate state is high, managers will be given short-term incentives, delaying the release of information. When the belief about the aggregate state is low, long-term incentives will be prevalent and information will be released without delay. This produces asymmetric learning dynamics for the economy, with gradual booms and rapid recessions. In a boom the belief about the aggregate state increases, information is pushed off into the future, and learning is slow. In a recession

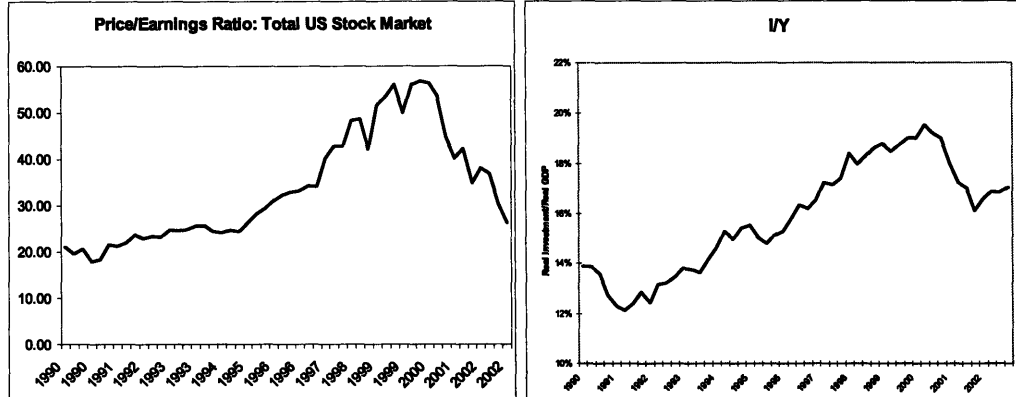


Figure 1-1: The Experience of the 1990s: Stock Market and Investment

the belief is falling, triggering a switch to long-term incentives, that brings forward the release of information and accelerates learning.

1.1 Introduction

Over the 1990s the US economy and stock market experienced a long gradual boom followed by a sharp downturn. The left panel of Figure 1-1, which depicts the path of the price to earnings ratio for the US stock market, shows a clear asymmetry¹. Over the four years from 1996, the P/E ratio rose from 30 to its peak of almost 60. It then fell by the same amount in only half the time. The right panel of Figure 1-1 shows a similar asymmetry for investment². This paper presents a model in which these types of asymmetries occur because learning is slow in booms and intense in recessions.

The mechanism that produces these changes in the speed of learning is motivated by a second feature of the 1990s – corporate fraud. Several examples have become notorious in the popular press: Enron, Kmart, Tyco, Global Crossing, Quest, Worldcom, and Waste Management³.

¹The price/earnings ratio is constructed as follows. The numerator is the real (inflation-adjusted) Datastream Total Market Index; the denominator is a moving average (over preceding ten years) of real earnings corresponding to the index.

²The right panel of Figure 1-1 plots total real investment as a percent of real GDP. Data source: BEA.

³For a more complete list of corporate scandals and dates see Graham, Litan and Suktankar (2002).

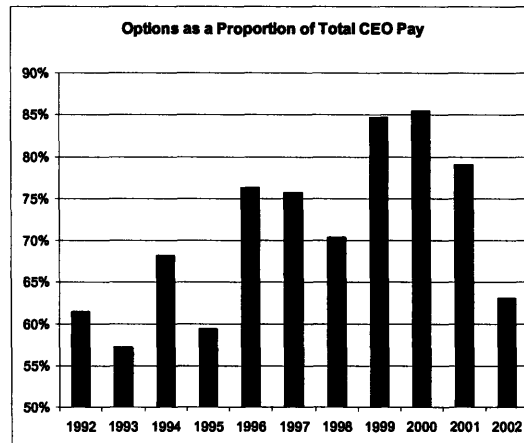


Figure 1-2: Managerial Compensation Over the 1990s

There is also substantial indirect evidence of more widespread fraud. Earnings restatements rose significantly during the last years of the 1990s (Wu (2002)). The gap between aggregate earnings and taxable corporate income⁴ became progressively larger over the 1990s (Desai (2002)). The proportion of firms who beat analyst's forecasts of earnings rose sharply over the 1990s (Matsumoto (2000))⁵. The rise in corporate fraud during the 1990s was closely linked to the way that managerial compensation changed over this time. Figure 1-2 plots the total value of stock options granted to CEOs as a proportion of total CEO pay for firms in the S&P 500⁶. During the boom, the pay of CEOs became increasingly sensitive to the stock price of their firm. This led many managers to undertake actions that distorted the market's perception of their firm's performance. Bergstreser and Phillipon (2003) show that CEOs whose pay was more sensitive to movements in their firm's stock price were more likely to have undertaken earnings management. Kedia (2003) shows that managers with higher stock option grants were more likely to have engaged in aggressive interpretation of accounting standards.

This paper begins by studying how the optimal incentives offered to managers vary with

⁴Taxable income abstracts from most of the estimates and provisions which can be used to manipulate reported earnings.

⁵For a survey of the literature documenting the rise in earnings manipulation over the 1990s see Lev (2003).

⁶Figures reconstructed from data in Hall and Murphy (2003) originally drawn from the ExecuComp Database. I thank Brian Hall and Kevin Murphy for providing me with the underlying data. Note that these are the granted, not realized, value of options, so there is not a mechanical relationship between their level and the stock market.

economic conditions. The trade-off involved in setting managerial pay is of the type studied in the multi-task principal-agent literature (see Holmstrom and Milgrom (1991) and Baker (1992)). In setting a manager's pay, a firm's owners must decide how much weight to assign to the observed short-term performance of the firm compared to its long-term performance. A firm's short-term performance is a relatively informative signal of managerial effort. Incentives sensitive to short-term performance are therefore effective at inducing managerial effort. However short-term performance can be manipulated by the manager, at a cost to the firm. In contrast, long-term incentives deter managers from hiding poor short-term results. The way this trade-off is resolved depends upon the aggregate state. When times are good, only a small fraction of managers find their firm performing poorly. Thus, in good times, short-term incentives are optimal because the probability that they will induce manipulation is low. Conversely, when times are bad, a large fraction of firms will perform poorly. Therefore, if owners offer short-term incentives in bad times, this is likely to result in the manager engaging in costly manipulation. Consequently, owners use incentives that are less sensitive to a firm's short-term performance in bad times. Thus, managers are induced to manipulate their observed short-term performance of their firm only during booms.

Next, I study how the contracting decision of each firm affects the dynamics of the economy. I consider an environment in which the true aggregate state is not directly observable. The dynamics of the economy are generated by changes in the public belief about the aggregate state. The link to the micro problem within each firm is that the short-term performance of the firm plays a dual role. Inside the firm, it is used to provide incentives. Outside the firm, it is used to learn about the state of the economy. A feature of the model is that managerial manipulation, while it does not fool the market, cannot be fully unravelled. Thus in good times, when managers distort the short-term performance of their firms, short-term performance is less informative for the aggregate state. As a result, the manager's private information about the aggregate state is released with delay, when the long-term output of the firm is realized. Thus, when the aggregate state is high, information is pushed off into the future and learning is slow. Conversely, when the belief falls, the process is reversed and learning occurs rapidly. As a result, busts are precipitous relative to the booms which precede them.

I then study the implications of the model for the stock market. As with the belief, move-

ments in the stock market will be asymmetric. This comes from two related channels. First, learning is most intense during a collapse. Second, the sensitivity of the stock market to news about the aggregate state changes over time. If firms have been truthfully revealing their performance, investors have very accurate information about their true value. In this case, additional information about the aggregate state does not have a strong impact upon the value of their firm⁷. If, however, managers have been given incentives that encourage manipulation, investors are more uncertain about the true value of firms in the stock market. Information about the aggregate state helps investors to re-evaluate the true value of any firm suspected of engaging in manipulation. Since manipulation occurs in good times, the stock market is most vulnerable to news about the aggregate state after a long boom.

Finally, I extend the basic model to allow for investment. The optimal level of investment at any point in time is an increasing function of the level of the belief. Thus, the learning dynamics of the economy translate directly into predictions about the speed with which investment changes over the cycle.

The paper proceeds as follows. Section 2 presents the model and studies the optimal contracting decision within each firm. Section 3 studies how the private contracting decisions within each firm affect the way that agents in the economy learn about the aggregate state. Section 4 studies the implications of the model for movements in the stock market and investment. Section 5 relaxes a number of the assumptions of the model in order to demonstrate its robustness. The central result in each case is that the key predictions of the model are not sensitive to the simplifying assumptions that are made to ease its exposition. Section 6 discusses the related literature. Section 7 provides a brief conclusion.

1.2 The Model

This section builds the main model of the paper. I study an environment in which the timing of the release of information varies endogenously, while the actual production of information is exogenously fixed.

⁷In fact, in the model presented here, their value is unaltered by information about the aggregate state.

1.2.1 The Economy

Consider an economy with an OLG production structure. Let time be discrete and indexed by $t \in J^+$. One firm enters in each period. The owners of each firm are risk neutral. The owners must hire a manager to operate the firm. Managers are also risk neutral. The interests of the owners and the manager are not aligned ex ante. Upon entry the owners must write an incentive contract for the manager. This contract will make a payment to the manager conditional upon the observed performance of the firm. The observed performance of the firm has two components; short and long-term. I assume that the true short-term performance of the firm is a sufficient statistic for the effort of the manager⁸. However, the observed short-term performance of the firm can be manipulated by the manager. I will refer to the observed short-term performance of the firm as the manager's signal. Both the manager's signal and the long-term output of the firm are publicly observed. The owners use these outcomes to determine the payment to the manager. Other firms in the economy also use the manager's signal to learn about the true state of the economy.

The true state of the industry at time t is denoted by θ_t and this can be either high ($\theta_t = \theta^H$) or low ($\theta_t = \theta^L$). At $t = 1$ the true state is randomly chosen with equal probability for high and low. Following this, the state can change from one period to the next with a constant transition probability of δ which should be thought of as being close to zero. Let $\Theta_t = \{\theta_1, \theta_2, \dots, \theta_{t-1}, \theta_t\}$ be the history of true realized states at period t . Let H_t be the history of all publicly observable events up to and including period $t - 1$. The objects which comprise H_t will be made clear as the model is developed. It is useful to define

$$\mu_t \equiv \Pr(\theta_t = \theta^H | H_t)$$

the public belief that the state at the beginning of period t is high, given the history H_t . For now we will suppose that firms are only able to write incentive contracts contingent upon information that has been released before period t , that is contingent upon H_t . I will consider

⁸To simplify the model I have assumed that the long term performance of the firm carries no additional information about managerial effort beyond its relation to the firm's short term performance. The assumption is stronger than necessary. All that is required is that the firm's final output not be a sufficient statistic for effort. This ensures that the firm's short-term performance is useful in providing incentives.

the implications of relaxing this assumption later in Section 1.5.

1.2.2 The Firm's Problem

When a firm enters at period t the owners of the firm must hire a manager and offer them an incentive contract. Assume that there is an infinite supply of managers and normalize their outside opportunity wage to zero. The owners, who are assumed to be risk neutral, will select an incentive contract that maximizes the final expected profit of the firm. Profit is simply the output of the firm, net of any payments to the manager. The manager's contribution to the firm is represented by a single choice of effort level e_t , at period t , which can be either high ($e_t = 1$) or low ($e_t = 0$). If a manager chooses low effort they enjoy a personal benefit of $B > 0$. Figure 1-3 illustrates the incentive problem within each firm. The effort choice of the manager and the state of the industry at t determine the true short-term performance of the firm. The true short-term performance of the firm is either high s_h or low s_l . The probability that the firm's true short-term performance is high s_h is $p(e_t, \theta_t)$. For now, assume that effort is additive so that

$$p(e_t = 1, \theta_t) = p(e_t = 0, \theta_t) + \eta \quad (1.1)$$

for $\theta_t \in \{\theta^H, \theta^L\}$ and $\eta > 0$. Assuming that effort increases the probability of high true short-term performance by a constant amount fixes the degree of moral hazard (with respect to effort) to be constant across states (i.e., $p(e_t = 1, \theta_t) - p(e_t = 0, \theta_t)$).⁹ The additivity assumption is relaxed in 1.5. Define

$$p^H \equiv p(e_t = 1, \theta_t = H), \quad p^L \equiv p(e_t = 1, \theta_t = L)$$

and assume that $p^H > p^L$. The probability of high short-term performance is increasing in both effort ($\eta > 0$) and the state of the economy ($p^H > p^L$). For the analysis which follows it will be useful to denote the expected probability that a firm's short-term performance is good

⁹This is a natural assumption to impose at this point since I want to focus on the way that the manager's ability to distort their signal can alter the incentives which they are offered by owners. It would be undesirable to have the choice of incentives differ with the belief simply because the underlying degree of moral hazard varied between the states. The additive assumption is not necessary for the results which follow.

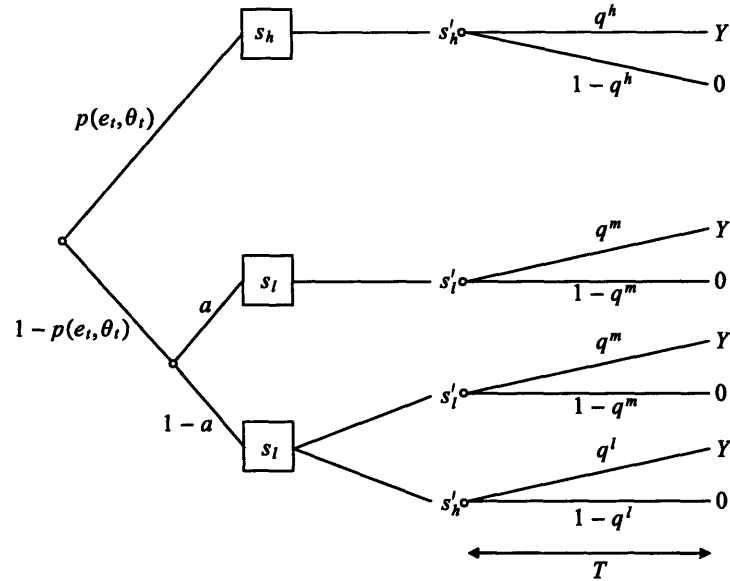


Figure 1-3: The Firm's Problem

after high effort as

$$p_t \equiv \mu_t p^H + (1 - \mu_t) p^L.$$

The true short-term performance of the firm is privately observed by the manager of the firm. If this signal is high, the manager will simply release the high signal s'_h . If, however, the true performance of the firm is low, the manager may be able to distort this signal and report high performance. With probability a the manager is unable to manipulate the signal and hence must report s'_l . The parameter a can be interpreted as the rigor of the audit standards which a manager faces. Alternately, we can take a broader view of what constitutes manipulation and can interpret $1 - a$ as the probability that a manager finds a way to boost the observed short-term performance of the firm¹⁰. In this case, let the manager select $M_t \in \{0, 1\}$. The manager can elect to distort the firm's short-term signal ($M_t = 1$) and thus release s'_h or can elect to release the true signal s'_l ($M_t = 0$). For now, the parameter a will be taken as being exogenously fixed, but later I will allow the owners of the firm to select a as well as the manager's incentives. The qualitative results of the paper are not affected by this. The manager's signal is released

¹⁰See Stein (1989) for further discussion and examples of managerial short-termism.

at period t .

The firm produces a publicly observable level of output at time $t + T$. A firm will either produce $Y > 0$ or zero. If the firm was performing well in the short-term then it produces Y with probability q^h . Conversely, if the firm's short-term performance was low and the manager truthfully signalled¹¹ s'_t , then the firm will produce Y with probability q^m . Finally, if the true short-term performance of the firm was low but the manager elected to manipulate these results, then the firm will produce output with probability q^l . Assume that the probabilities are ordered so that $q^h > q^m > q^l$. Having $q^h > q^m$ characterizes what it means for a firm to truly be performing well in the short-term: their prospects for producing output at $t + T$ are high. In the context of the moral hazard problem we need q^h to be sufficiently larger than q^m so as to make inducing high effort profitable (the exact condition for this will be arrived at below). The requirement that $q^m > q^l$ implies that when a manager distorts the short-term performance of their firm that this destroys long-term value. This can be thought of in a number of ways. If we interpret the manager manipulating their signal as action taken to alter the firm's accounts then this will require time and resources be directed to inflating the appearance of the firm's performance. Ultimately such activities will lower the expected output of the firm relative to the case where these resources had been used productively. More broadly, we can interpret manipulation to mean inflating short-term results. This will require the manager taking projects with lower net present value but whose short-term output is high. Provided that such decisions are not publicly observed then a manager who thinks that their firm is doing poorly may elect to follow short-term strategies to pump up short-term results. Ultimately, however, such activity will lower the expected output of the firm.

1.2.3 Contract Design

For ease of exposition, assume that firm t cannot write a contract conditioned on outcomes of firms who enter after t . It follows that a fully general managerial contract will specify a payment to the manager contingent upon the history H_t , the publicly observed short-term report, and the final output of the firm. For a given history H_t , the contract will be fully characterized by

¹¹This applies both the case where the manager was unable to misreport and to the case where the manager was able to misreport and elected not to do so.

the following set of payments

$$\Gamma \equiv \{w(s', y, H_t) \mid s' \in \{s'_h, s'_l\}, y \in \{0, Y\}\}$$

The manager is risk neutral however, the contract is subject to a limited liability restriction so that $w(s', y, H_t) \geq 0$. In effect this prevents the owners of the firm simply selling the firm to the manager which would costlessly solve the moral hazard problem. The central results of the model also hold if we relax the limited liability restrictions and simply impose that the manager is risk averse and has a fixed reservation utility.

The manager will take the owner's contract as given and will choose $\{e_t, M_t\}$ to maximize their expected payoff. Define the expected payment to a manager who is able to manipulate their signal as

$$G(M_t, \Gamma) \equiv (1 - M_t) [q^m w(s'_l, Y, H_t) + (1 - q^m) w(s'_l, 0, H_t)] + M_t [q^l w(s'_h, Y, H_t) + (1 - q^l) w(s'_h, 0, H_t)]$$

For a given contract Γ the manager's expected payment is

$$\begin{aligned} W(M_t, \Gamma, e_t, H_t) &= (p_t - (1 - e_t)\eta) [q^h w(s'_h, Y, H_t) + (1 - q^h) w(s'_h, 0, H_t)] \\ &\quad + (1 - p_t + (1 - e_t)\eta) a [q^m w(s'_l, Y, H_t) + (1 - q^m) w(s'_l, 0, H_t)] \\ &\quad + (1 - p_t + (1 - e_t)\eta) (1 - a) G(M_t, \Gamma) \end{aligned}$$

The expected output of the firm is

$$R(H_t, M_t, e_t) \equiv (p_t - (1 - e_t)\eta) q^h + (1 - p_t + (1 - e_t)\eta) [a q^m - (1 - a) (q^m - q^l) M_t] Y$$

We are now ready to write the owners' problem¹². They will select Γ so as to maximize the

¹²Implicit in the formulation of this problem is that I have ruled out revelation mechanisms in which the manager can, separate to their signal, report their private information. It is easy to show that the payoff to the owners of the firm is no greater under such a scheme and hence ruling it out is not a strong assumption. It is however important for the learning dynamics which follow that full revelation does not occur.

expected output of the firm net of any payments to the manager,.

$$\begin{aligned} \max_{\Gamma} \Pi &= R(H_t, M_t^*, e_t^*) - W(M_t^*, \Gamma, e_t^*, H_t) \\ &\text{subject to} \\ M_t^* &= \arg \max_{M_t \in \{0,1\}} G(M_t, \Gamma) \\ e_t^* &= \arg \max_{M_t \in \{0,1\}} W(M_t^*, \Gamma, e_t, H_t) \\ w(s', y, H_t) &\geq 0. \end{aligned}$$

Implicit in this program is the assumption that owners do not value any information which the firm may produce about the aggregate state. Their only concern is using the signal and the firms output to provide optimal incentives to the manager. It is assumed that the owners of all firms are unable to write contracts which solve this information externality.

Assume that the parameters of the model are such that it is always optimal for the owner's of the firm to induce high effort. This requires that η be sufficiently large so that the expected gain from inducing high effort outweighs the cost of incentives. The exact condition is given in (1.31) of the appendix.

The optimal contract is solved for in the appendix and the results are collected in Proposition 1.

Proposition 1 *There exists a unique $\bar{\mu}$ such that if $\mu_t \geq \bar{\mu}$ then the unique optimal contract is*

$$\begin{aligned} w(s'_h, Y, H_t) &= \frac{B}{\eta [q^h - (1-a)q^l]}, \\ w(s'_h, 0, H_t) &= w(s'_l, Y, H_t) = w(s'_l, 0, H_t) = 0. \end{aligned} \tag{1.2}$$

This induces the manager to select $e_t = 1$ and $M_t = 1$ and results in an expected profit to the owners of the firm of

$$\Pi^M(\mu_t) = \left[p_t q^h + (1-p_t) (a q^m + (1-a) q^l) \right] Y - \frac{B}{\eta} \left(p_t + \frac{(1-a)q^l}{q^h - (1-a)q^l} \right). \tag{1.3}$$

Alternately, if $\mu_t < \bar{\mu}$ then the optimal contract is

$$\begin{aligned} w(s'_l, Y, H_t) &= w(s'_l, 0, H_t) = \frac{q^l B}{\eta(q^h - q^l)}, \\ w(s'_h, Y, H_t) &= \frac{B}{\eta(q^h - q^l)}, \\ w(s'_h, 0, H_t) &= 0. \end{aligned} \tag{1.4}$$

This induces the manager to select $e_t = 1$ and $M_t = 0$ and results in an expected profit to the owners of the firm of

$$\Pi^T(\mu_t) = \left[p_t q^h + (1 - p_t) q^m \right] Y - \frac{B}{\eta} \left(p_t + \frac{q^l}{q^h - q^l} \right). \tag{1.5}$$

In order to deter manipulation of the signal the incentives given to the manager are made less sensitive to the short-term report. Ideally the owners would pay the manager based on the true short-term performance of the firm since this is the most informative signal of the managers effort choice and hence would minimize the cost of incentives. However, since the observed signal may be manipulated, this is not possible. Despite the possibility that the short-term signal is manipulated, the observation of a high signal and output Y is still the most informative sign of high effort. Manipulation could, for example, be completely avoided if the manager were offered long-term incentives that paid only if the final output of the firm was Y . The cost of such incentives is that observing Y is not as informative for high effort than is observing both $s' = s'_h$ and Y even if the manager is choosing the manipulate their signal. Formally¹³:

$$\frac{\Pr(s' = s'_h, Y | e_t = 1, \mu_t)}{\Pr(s' = s'_h, Y | e_t = 0^l, \mu_t)} > \frac{\Pr(Y | e_t = 1, \mu_t)}{\Pr(Y | e_t = 0, \mu_t)}$$

The optimal contract which avoids manipulation has an unambiguously higher expected cost of incentives because it makes payments to the manager outside of the state where both $s' = s'_h$ and Y are realized. In fact, it makes a fixed payment to the manager whenever the low signal s'_l is observed which is an informative sign of the manager having taken *low* effort. The

¹³Where $\frac{\Pr(s' = s^h, Y | e = e^h, \mu_t)}{\Pr(s' = s^h, Y | e = e^l, \mu_t)} = \frac{p_t q^h + (1 - p_t)(1 - a)q^l}{\sigma_t q^h + (1 - \sigma_t)(1 - a)q^l}$ and $\frac{\Pr(Y | e = e^h, \mu_t)}{\Pr(Y | e = e^l, \mu_t)} = \frac{p_t q^h + (1 - p_t)[(1 - a)q^l + a q^m]}{\sigma_t q^h + (1 - \sigma_t)[(1 - a)q^l + a q^m]}$.

cost of this is that it actually exacerbates the moral hazard problem and requires the payment $w(s'_h, Y, H_t)$ to be all the higher in order to induce high effort. The difference in the cost of incentives is

$$\frac{aBq^l q^h}{\eta(q^h - q^l)(q^h - (1-a)q^l)} > 0$$

However, the cost of allowing manipulation is that it lowers the expected output of the firm. A contract which encourages a manager to distort their signal lowers the expected output of the firm by

$$(1-a)(q^m - q^l)(1-p_t)Y > 0$$

The optimal choice of contract weighs these two costs. Making the manager's contract less sensitive to its short-term signal raises the costs of managerial incentives but discourages value destroying short-term behavior. Optimally, the owners of the firm will elect to discourage manipulation if and only if

$$(1-a)(q^m - q^l)(1-p_t)Y \geq \frac{aBq^l q^h}{\eta(q^h - q^l)(q^h - (1-a)q^l)} \quad (1.6)$$

The central feature of this decision is how this trade-off varies with the belief μ_t . Observe first that it is only a manager who finds their firm performing poorly who may want to engage in costly manipulation. The likelihood that a manager finds herself in this situation depends upon the aggregate state. If times are good then a firm is less likely to end up performing poorly ($p^H > p^L$). In this case, encouraging manipulation is less costly because it only occurs with a low probability. Conversely, when times are bad a manager is likely to find their firm performing poorly. In this instance short-term incentives are less desirable because they lead to manipulation more often. The result is that in equilibrium manipulation is tolerated when the belief is high because owners assess that it is less likely to occur. By this logic owners will only chose to use short-term incentives when the belief is above $\bar{\mu}$. This result matches the stylized fact that it is during times of optimism, such the the boom of the 1990s, that agents enter into contracts which can lead to fraud.

Assume that the parameters of the model are such that the cutoff belief is interior ($\bar{\mu} \in$

$(0, 1)$). For the analysis which follows it will be more convenient to describe the belief using the log likelihood ratio

$$\lambda_t \equiv \ln \left(\frac{\mu_t}{1 - \mu_t} \right).$$

Let $\bar{\lambda}$ be the log likelihood ratio associated with the cutoff belief $\bar{\mu}$. Since λ_t is strictly increasing in μ_t then it follows that the optimal contract will induce the manager to manipulate their signal when $\lambda_t \geq \bar{\lambda}$ and will discourage this when $\lambda_t < \bar{\lambda}$.

We now move on to study the implications for the aggregate release of information given that contracts between owners and managers are signed in this way.

1.3 Learning Dynamics

The previous section fully described the contract choice of each individual firm taking the current belief about the true state as given. We now study how the belief evolves given that the contracting decisions of individual firms is as described in the previous section.

Within each period t the timing of events is as follows: First, the owners of the firm formed at period t select the incentive contract to offer the firm's manager. The choice of contract is publicly observed. The manager then selects effort, observes the manager's private signal of the firm's performance and (if appropriate) decides whether to manipulate the firm's public signal. The public signal of the firm's short-term performance is then released. Finally, the output of the firm created at period $t - T$ is realized and publicly observed. As we shall see below the belief can change with both the realization of firm $t - T$'s results as well as with the signal of the manager during period t . In order to distinguish between these beliefs define: λ_t as the belief at the beginning of the period, λ_t^m as the belief in the middle of the period after the report of firm t is announced, and λ_t^e as the belief at the end of the period after the output of firm $t - T$ is realized. The ordering of events within a period are summarized in Figure 1-4.

1.3.1 Learning From the Firm

Reconsider the problem of the firm as illustrated in Figure 1-3. So far we have considered this problem from the perspective of the firm's owner. Their objective was to use the manager's signal and the firms output to provide optimal incentives to the firm's manager. From an

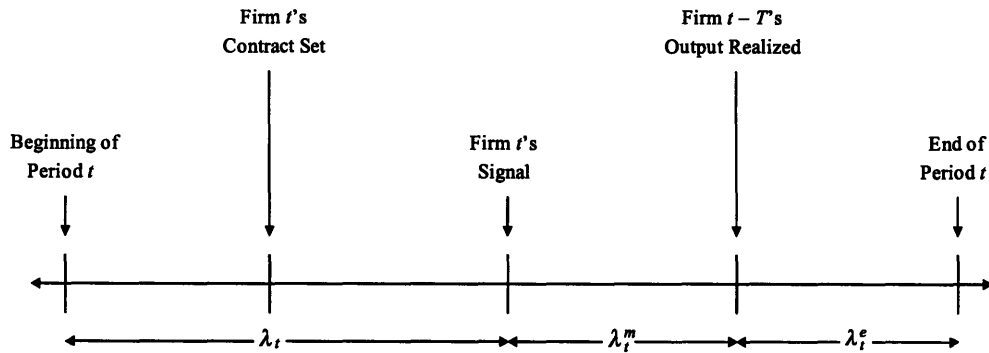


Figure 1-4: Timeline of Events Within a Period

outsider's perspective the publicly observed outcomes of the firm play a very different role. They provide information about the true state of the economy. The model has been constructed so that the only direct role of the aggregate state is to determine the true short-term performance of the firm.¹⁴ In effect then the true short-term performance of the firm contains all the information inside the firm about the aggregate state. Any agent outside the firm will use the manager's signal and (T periods later) the firms output to infer what the true short-term performance of the firm was in order to update their belief about the aggregate state.

If a manager is given incentives which discourage manipulation then it is clear that all the information within the firm is released in the manager's signal. If a manager is given incentives which encourage manipulation and they subsequently release the low signal s'_l then outsiders recognize that the true short-term performance of the firm must have been low s_l . Again the manager has revealed all their information about the aggregate state. Finally, if the manager is given incentives which encourage manipulation and they release the high signal s'_h then it is not clear what the true short-term performance of the firm is. The manager may have observed s_h and truthfully signaled this. Alternately, the manager may have observed s_l and been able to distort their observed performance so that s'_h is released. In this scenario the manager signal is a garbling of their information about the state. The output of the firm now becomes

¹⁴The results of the model would be qualitatively unchanged if this feature was relaxed. The model has been setup in this way for simplicity and to emphasize the role of manipulation in changing the timing of information release.

informative because it helps an outsider infer whether the manager's signal was truthful or not. This follows from the fact that the firm's output is positively correlated with the true short-term performance of the firm. Thus when an outsider observes the firm's output they gain additional information about the state when the manager is suspected of having manipulated short-term results. The important feature of this is that when manipulation occurs it delays the release of information.¹⁵ In contrast when manipulation is not occurring the manager's information is fully released without delay in their signal. This captures the sense in which manipulation changes the timing of social learning by altering when the information inside a firm is released.

At this point it is important to note a crucial feature of the model. When managers manipulate their performance this reduces the informativeness of their signal. There are many models where agents distort their signal but in equilibrium this is fully unravelled. This is a feature of signal jamming models (see for example Fudenberg and Tirole (1996), Holmstrom (1999) and Stein (1989)). Unravelling does not occur in this model for two reasons. The first reason is the fact that performance can only be high or low and so when a manager manipulates their signal they must pool on the high signal. It would be undesirable to have a theory that rested only this. The deeper reason why manipulation is not fully unravelled is that outsiders are unsure exactly what problem the manager is facing when they want to manipulate. This is captured by the fact that a manager privately observes not only her signal but also whether or not she is able to manipulate her signal. My contention is that while unravelling is a useful feature of many models of signal manipulation it does not do a good job of capturing how the flow of information is affected by such activity. Instead the theory here suggests that while we can anticipate manipulation we cannot unravel it if we are not sure what a manager will signal if they are induced to manipulate. This does a better job of capturing the corporate fraud of the 1990s. The final discovery of this activity led to large downward revisions in the prices of those firms involved. The clear implication is that while such activity may reasonably have been suspected it remained unclear if firms were in fact manipulating results or truthfully

¹⁵ Observe that a manager who is suspected of having manipulated their signal will have private information about the true state of the economy for the life of their firm. It is crucial to the story here that the manager does not reveal this information in other ways. I assume that this manager is unable to participate in capital markets where their private information about the state could be revealed through the price impact of their trades. The results of the model continue to work if we suppose that the manager's price impact is partially garbled. This would be true if the manager's participation in the market was limited and noise traders also operated in the market.

revealing their performance.

I now turn to studying the way that the belief is updated.

1.3.2 Laws of Motion for the Belief

Now we begin to study the way in which the belief evolves over time. From one period to the next we have that

$$\lambda_t = \ln \left(\frac{(1 - \delta) \mu_{t-1} + \delta (1 - \mu_{t-1})}{\delta \mu_{t-1} + (1 - \delta) (1 - \mu_{t-1})} \right)$$

I focus upon the case where $\delta \approx 0$, in which case $\lambda_t \approx \lambda_{t-1}^e$.

Next, we ask how the belief will change with the manager's signal. If the manager is given a contract which encourages manipulation and they signal s'_h then the manager's private information is not fully revealed. In this case it is unclear whether the manager did in fact observe the high signal and has released it truthfully or if they observed the low signal and were able to manipulate. In all other scenarios the manager's private information is fully revealed. The laws governing the evolution of the public belief following the manager's signal are collected in Proposition 2.

Proposition 2 *If a manager signals s'_i then $\lambda_t^m = \lambda_t + \lambda^{s'_i}$. If $\lambda_t < \bar{\lambda}$ and the manager signals s'_h then $\lambda_t^m = \lambda_t + \lambda^{s'_h T}$. If $\lambda_t \geq \bar{\lambda}$ and the manager signals s'_h then $\lambda_t^m = \lambda_t + \lambda^{s'_h M}$. Where*

$$\lambda^{s'_i} \equiv \ln \left(\frac{1 - p^H}{1 - p^L} \right) < 0, \lambda^{s'_h T} \equiv \ln \left(\frac{p^H}{p^L} \right) > 0, \lambda^{s'_h M} \equiv \ln \left(\frac{p^H + (1 - p^H)(1 - a)}{p^L + (1 - p^L)(1 - a)} \right) > 0$$

Observing the high signal causes the public belief to increase no matter what the manager's incentives are ($\lambda^{s'_h T}, \lambda^{s'_h M} > 0$). However when the manager is suspected of having manipulated their signal then it is less informative and thus moves the public belief by less ($\lambda^{s'_h T} > \lambda^{s'_h M}$). The severity of the manipulation problem is parameterized by a . It is easy to show that

$$\frac{d\lambda^{s'_h M}}{da} > 0, \lambda^{s'_h M}|_{a=1} = \lambda^{s'_h T}, \lambda^{s'_h M}|_{a=0} = 0$$

The larger is a the more informative is the manager's signal when they are given incentives which encourage manipulation. In the extreme cases ($a = 0, 1$), if a manager who is given

incentives which encourage manipulation then their signal will either completely reveal their private information ($a = 1$) or will be totally uninformative ($a = 0$).¹⁶

Finally, we ask how the realization of firm $t - T$'s output will alter the belief. The output of a firm is only informative if the manager is suspected to have manipulated their. This only occurs when the manager was given incentives which encouraged manipulation and the manager subsequently released the high signal. The output of such a firm is useful in determining if the manager's signal was in fact truthful. In all other scenarios the firms output is not informative for the state. The way in which the public belief evolves following the realization of firm $t - T$'s results is recorded formally in Proposition 3.

Proposition 3 *Following the realization of firm $t - T$'s output at period t for $\delta \approx 0$ the public belief is updated in the following way:*

$$(i) \lambda_t^e = \lambda_t^m \text{ if } \lambda_{t-T} < \bar{\lambda} \text{ or if } \lambda_{t-T} \geq \bar{\lambda} \text{ and } s'_{t-T} = s'_l;$$

$$(ii) \lambda_t^e \approx \lambda_t^m + \lambda^Y \text{ if } \lambda_{t-T} \geq \bar{\lambda}, s'_{t-T} = s'_h \text{ and firm } t - T \text{'s output is } Y$$

$$(iii) \lambda_t^e \approx \lambda_t^m + \lambda^0 \text{ if } \lambda_{t-T} \geq \bar{\lambda}, s'_{t-T} = s'_h \text{ and firm } t - T \text{'s output is } 0$$

$$\text{where } \lambda^Y \equiv \ln \left(\frac{p^h q^h + (1 - p^h) (1 - a) q^l}{p^l q^h + (1 - p^l) (1 - a) q^l} \right) - \lambda^{s'_h M} > 0$$

$$\lambda^0 \equiv \ln \left(\frac{p^h (1 - q^h) + (1 - p^h) (1 - a) (1 - q^l)}{p^l (1 - q^h) + (1 - p^l) (1 - a) (1 - q^l)} \right) - \lambda^{s'_h M} < 0$$

The informativeness of a firm's output depends upon how easy it is for managers to manipulate their signal. If $a = 1$ then managers are unable to distort their signal and as a result their signal fully reveals their information. As a result nothing additional is learned from output ($\lambda^Y = \lambda^0 = 0$). By the same logic, the smaller is a the less information is released in a manager's signal (when $\lambda_{t-T} \geq \bar{\lambda}$). As a result more information is kept inside the firm and

¹⁶Note however that if $a = 0$ then the contract which discourages manipulation will always be optimal. Conversely, if $a = 1$, then a contract based entirely upon the short term signal will be optimal. Technically this contract won't induce manipulation because managers will never be able to manipulate their signal.

delayed until the firms output is released T periods later. This is easily seen by noting that

$$\frac{d\lambda^Y}{da} < 0, \frac{d\lambda^0}{da} > 0.$$

We have now fully described the rules governing the evolution of the belief. The key feature of these dynamics is that manipulation delays the release of information. This is the sense in which manipulation changes the timing of social learning. What is crucial the economy as a whole is that the timing of social learning will vary systematically with the level of the belief itself. When the belief is high so that $\lambda_t \geq \bar{\lambda}$ then the release of information will be delayed, in effect pushing off information into the future. However when the belief is low so that $\lambda_t < \bar{\lambda}$ then information is released quickly. Accordingly the speed of learning will change as the belief about aggregate conditions transitions between these two regions. Our next task is to study these dynamics in detail.

1.3.3 The Phases Of Learning

We can now begin to describe the way that learning varies over the cycle. During any period there are two potential sources of information: firm t 's signal and firm $t - T$'s results. The signal of firm t can either be garbled ($\lambda_t \geq \bar{\lambda}$) or truthful ($\lambda_t < \bar{\lambda}$). Similarly, the results of firm $t - T$ can, in expectation, be informative ($\lambda_{t-T} \geq \bar{\lambda}$) or uninformative ($\lambda_{t-T} < \bar{\lambda}$). Combining these we have four possible phases of learning. These are summarized in Table 1.

Phase	Sources of Information	
	Signal	Results
Direct	Truthful ($\lambda_t < \bar{\lambda}$)	Uninformative ($\lambda_{t-T} < \bar{\lambda}$)
Slow	Garbled ($\lambda_t \geq \bar{\lambda}$)	Uninformative ($\lambda_{t-T} < \bar{\lambda}$)
Delayed	Truthful ($\lambda_t < \bar{\lambda}$)	Informative ($\lambda_{t-T} \geq \bar{\lambda}$)
Intense	Garbled ($\lambda_t \geq \bar{\lambda}$)	Informative ($\lambda_{t-T} \geq \bar{\lambda}$)

Table 1: The Phases of Learning

We can rank the amount of information released in these different phases. Following Blackwell (1951) we have that message service A is more informative than message service B if the

likelihood matrix associated with B can be written as a convex combination of the likelihood matrix of A ¹⁷. In words, A is more informative than B is we can represent B as a garbling of A . This provides an incomplete ranking over message services. Using this concept we can rank the amount of information which is released in the four different phases of learning. This ordering is given in Proposition 4.

Proposition 4 *The amount of information released in the phases of learning is ordered in the following way:*

$$Intense \geq Direct, Delayed \geq Slow$$

Proposition 4 is proved in the appendix.

As the level of the belief changes this will systematically move the economy between these phases of learning. The way that learning varies over the cycle can be seen graphically in Figure 1-5. The top panel of Figure 1-5 displays a realized sample path for the likelihood ratio λ_t . The crucial feature of this path is that it transitions from a low belief ($\lambda_t < \bar{\lambda}$) to a high belief ($\lambda_t \geq \bar{\lambda}$) and then back to low. This path of the belief allows us to capture the four basic phases of learning and the role they play in producing asymmetric dynamics for the economy as a whole. The bottom panel of Figure 1-5 provides a diagrammatic representation of the release of private information. When the belief is low, all of a manager's private information is released immediately. This is represented by a single long vertical arrow. Conversely, when the belief is high then (on average) only a portion of the manager's private information is released through their signal. This is represented by the short vertical arrow. In addition some of the manager's private information is hidden in the firm and this is released when the firm's output is publicly realized T periods later. This delayed release of information is represented by the diagonal arrow. Whenever the belief is above $\bar{\lambda}$ then the private information of managers is released, on average, in both ways. Observe that the timing of information release is determined at each point by whether the belief at that point is above or below $\bar{\lambda}$.

Now follow the example sample path drawn in Figure 1-5. It starts with the belief being below $\bar{\lambda}$ so that the signals of managers immediately release all private information. When this

¹⁷See Laffont (1989) for a formal statement of this.

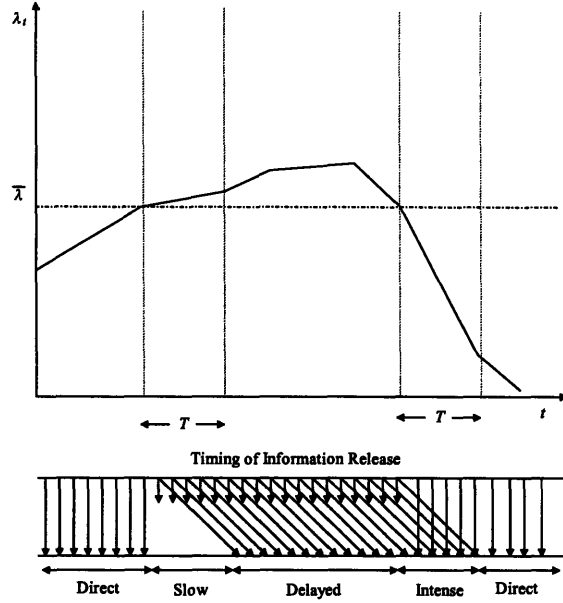


Figure 1-5: The Phases of Learning

is the only source of learning we have “Direct Learning”. The learning dynamics here are as if managers are unable to manipulate their signal.

The sample path then rises above $\bar{\lambda}$. As a result owners now offer managers short-term incentives which encourage them to manipulate their signal. This changes the timing of the release of manager’s private information by pushing information off into the future by T periods. The result is that for the first T periods that the belief is above $\bar{\lambda}$ that social learning only comes from the manipulated signals of managers. Thus as the belief rises the economy enters a phase of “Slow Learning”.

The example sample path remains above $\bar{\lambda}$ for more than T periods. Learning then comes from two sources: firm t ’s garbled signal and the output of firm $t - T$. The central feature of learning in this phase is that relative to direct learning a portion of the information being released is coming from old private information that has been caught up inside firms. It is for this reason that I call this phase “Delayed Learning”.

Next, observe that the sample path returns below $\bar{\lambda}$. This fall in the belief changes the timing of the release of manager’s private information. Now managers are given long-term

incentives and thus will release their private information without delay through their signal. This undoes the change in the timing of information that occurred when the belief rose above $\bar{\lambda}$. For the first T periods that the belief is below $\bar{\lambda}$ learning comes from two sources: the ungarbled signal of firm t as well as the output of firm $t - T$. In effect, this is direct learning *plus* additional information from the past. Thus as the belief falls the economy enters a phase of “Intense Learning”.

After the belief has been below $\bar{\lambda}$ for T periods the information from the past runs out and the system returns to direct learning. Putting these phases together we have that recoveries will be gradual with slow learning and that recessions will be rapid with intense learning. In addition, changes in the state will be detected only with delay when the belief is high so that, in retrospect, booms can continue well beyond the point when the state changes because the release of information is delayed in such times. When the belief is low then information about changes in the state will be released without delay. We now turn to studying the speed of learning in each phase.

1.3.4 The Speed of Learning

The dynamics of this economy are derived from the speed at which the belief changes over time. Booms are gradual in the sense that this is when learning becomes slow and hence the belief is slow to rise. Conversely recessions are rapid in because this is when learning is intense and hence the belief changes quickly.

To capture this formally I measure the speed of learning as the net change in the belief that occurs during a period. Let Δ_t denote this change:

$$\Delta_t \equiv \lambda_t^e - \lambda_t.$$

The magnitude of $|\Delta_t|$ captures the speed at which learning is occurring. The sign of Δ_t indicates the direction in which the belief is moving. It is important to note that I have defined the speed of learning in terms of changes in the log likelihood ratio rather than the belief. This is so as to abstract from the fact that for a given piece of information the belief will move more when it is closer to $\frac{1}{2}$. I abstract from this because my purpose is to study changes in the

speed of learning which come from changes in the amount of information which is released into the economy. It should be noted that although the belief will move more when it is close to $\frac{1}{2}$ that this of itself does not imply any asymmetry for the economy as a whole. The asymmetries which I find here in terms of the log likelihood ratio can themselves be translated directly into asymmetries for movements in the belief.

Our task is now to study systematic changes in the average speed of learning over different phases of the cycle. By rational expectations the expected value of Δ_t at the beginning of period t is zero. However, if we condition on the true state, then our expectation of Δ_t is, in general, non-zero. By studying the expected value of Δ_t , conditional on the state, we can ask: how quickly do we expect the economy to respond to a given realized path of the aggregate state Θ_t .

Before studying the way that the speed of learning changes over the cycle it is useful to make one additional assumption. So far I have only assumed that $p^H > p^L$. This ordering leaves open the possibility that the firm's true performance is more informative in one state than the other. If, for example, $p^H p^L < (1 - p^H)(1 - p^L)$ then the firm's performance is a more informative signal of the true state when the state is low. This of itself would produce asymmetries in the speed of learning: learning, in expectation, would be faster when the state is low. In order to eliminate this second possible source of asymmetry I assume that a firm's performance is a symmetric signal of the state. That is, I assume that

$$p^H = 1 - p^L.$$

I maintain this assumption for the remainder of this section. Under this assumption an economy without manipulation ($a = 1$) will exhibit perfect symmetry. We will see this when we study direct learning.

1.3.5 Direct Learning

Direct learning occurs in period t whenever λ_t and λ_{t-T} are below $\bar{\lambda}$. In this case the only source of social learning comes from the ungarbled signal of managers. Define the following realized path for the belief.

Definition 5 A Low Path is a sequence of realized public beliefs $\{\lambda_i\}_j^k$ for which $\max\{\lambda_i\}_{j-T}^k < \bar{\lambda}$.

Any point on a Low Path will be characterized by Direct Learning. The associated speed of learning is

$$E(\Delta_t | \Theta_t) = \Phi^T(\theta_t)$$

where

$$\Phi^T(\theta_t) \equiv \begin{cases} (2p^H - 1) \lambda^{s_h^T} & \text{if } \theta_t = \theta^H \\ -(2p^H - 1) \lambda^{s_h^T} & \text{if } \theta_t = \theta^L \end{cases}$$

Under direct learning the belief will, in expectation move in the direction of the current state θ_t ¹⁸. This is illustrated in Figure 1-6. The realized path of beliefs from *A* to *C* satisfies the conditions of a low path. As drawn, the true state along this path is low and hence at any point the expected direction of learning is negative. This is illustrated by the negative slope at *B*. At *C* the true state changes from low to high. The expected direction of learning responds without delay and is then positive. In fact, while the stochastic correspondence from *C* remains below $\bar{\lambda}$ then we know that at any realized path will be a low path and hence must exhibit direct learning. Accordingly, we can draw the expected path of the belief as a straight line until the stochastic correspondence rises above $\bar{\lambda}$. This is shown by the dotted line from *C* to *D*. The slope of this line indicates the expected speed and direction of learning under direct learning when the state is high. Observe also that under direct learning the speed of learning is perfectly symmetric. Thus if we removed the possibility of manipulation ($a = 1$) then the economy would exhibit perfect symmetry in booms and recessions.

1.3.6 Delayed Learning

Delayed learning occurs in period t whenever λ_t and λ_{t-T} are above $\bar{\lambda}$ and as such is the natural counterpart to direct learning. In this case social learning comes from two sources: the garbled signal of the manager of firm t and the realized output of the firm created T periods ago. Define the following realized path for the belief.

¹⁸Note that $(2p^H - 1) \lambda^{s_h^T} > 0$. To see this simply observe that $p^H > \frac{1}{2}$ which follows from symmetry the assumption of $(p^H = 1 - p^L)$ and that $p^H > p^L$.

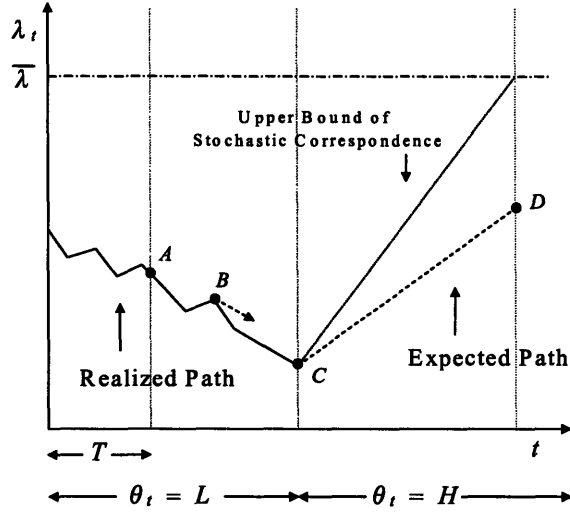


Figure 1-6: A Change in the State under Direct Learning.

Definition 6 A High Path is a sequence of realized public beliefs $\{\lambda_i\}_j^k$ for which $\min \{\lambda_i\}_{j-T}^k \geq \bar{\lambda}$.

Any point on a High Path will be characterized by delayed learning. During delayed learning the expected speed of learning has two components:

$$E(\Delta_t | \Theta_t) \approx \Phi^M(\theta_t) + \Omega(\theta_{t-T})$$

where

$$\Phi^M(\theta_t) \equiv [1 - a(1 - p^{\theta_t})] \lambda^{s_h^M} + a(1 - p^{\theta_t}) \lambda^{s_l^M}$$

$$\Omega(\theta_{t-T}) \equiv [p^{\theta_t} q^h + (1 - a)(1 - p^{\theta_t}) q^l] \lambda^Y$$

$$+ [p^{\theta_t}(1 - q^h) + (1 - a)(1 - p^{\theta_t})(1 - q^l)] \lambda^0$$

The first term $\Phi^M(\theta_t)$ represents the expected update coming from the garbled report of the manager. Note that

$$\Phi^M(\theta_t = H) > 0, \Phi^M(\theta_t = L) < 0$$

so that this component moves the belief, in expectation in the direction of the true state at t . The second term $\Omega(\theta_{t-T})$ represents the expected update to the belief coming from the results

of the firm created at $t - T$. The sign of this component is determined by the value of θ_{t-T} so that

$$\Omega(\theta_{t-T} = H) > 0, \quad \Omega(\theta_{t-T} = L) < 0.$$

This second term captures the sense in which learning is delayed. This second component of learning provides information about the value of the state T periods ago. This is of course informative for the value of the state at t because $\delta \approx 0$.

It is clear that if the state is the same at t and $t - T$ ($\theta_t = \theta_{t-T}$) then in expectation the belief will move in the direction of the current state. In this scenario the speed at which learning occurs is strictly slower than in the case of direct learning. Formally stated:

$$|E(\Delta_t | \theta_t = \theta_{t-T}, \lambda_{t-T} \geq \bar{\lambda}, \lambda_t \geq \bar{\lambda})| \geq |E(\Delta_t | \theta_t = \theta_{t-T}, \lambda_{t-T} < \bar{\lambda}, \lambda_t < \bar{\lambda})|$$

There are two reasons why learning is slower in the delayed learning phase relative to direct learning when $\theta_t = \theta_{t-T}$. The first is that while the results of the firm are informative they do not fully reveal the manager's private information. This is because a manager who has manipulated their signal can, with probability q^l produce output Y . Conversely a manager who has truly observed high performance may produce 0 with probability $1 - q^h$. As such, the output of the firm is a garbling of the managers information about the aggregate state. The second reason that delayed learning is slower than direct learning is (when $\theta_t = \theta_{t-T}$) is that information about θ_{t-T} is not as informative about the current state θ_t because of the possibility that the state has changed in the interim. This effect is however small when $\delta \approx 0$.

Next consider the scenario in which the state has changed during a phase of delayed learning so that $\theta_t \neq \theta_{t-T}$. The effect of this is to move the expected speed of learning in the direction of the old value of the state. As a result, in a phase of delayed learning the belief is, in expectation, slow to respond to a change in the true state. The strength of this effect depends upon the relative informativeness of the manager's garbled signal and the output of the firm. The informativeness of the manager's signal is determined by a . If managers have few opportunities to manipulate their signal (a is high) then the signal of manager's with high powered incentives are still highly informative. As a result the expected amount of learning from the results of firms will be small since, in expectation, most information will come from the

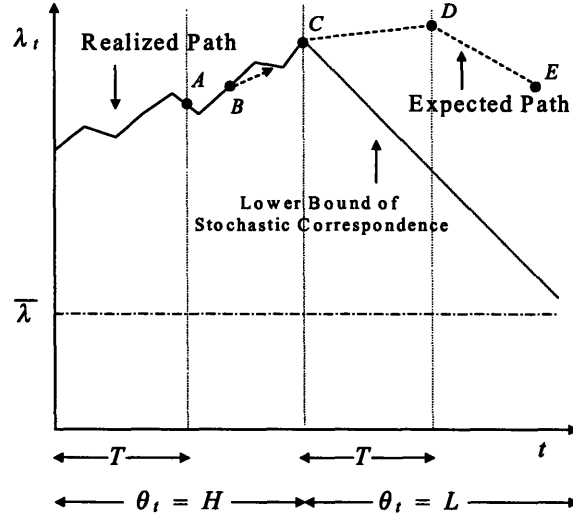


Figure 1-7: A Change in the State under Delayed Learning when $a \leq \hat{a}$.

manager's signal. In the extreme $a = 1$ so that manager's signals are fully reliable and output contains no additional information about the state. By this logic, when a is high, $\Phi^M(\theta_t)$ will dominate $\Omega(\theta_{t-T})$ and hence under delayed learning the belief will in expectation move in the direction of the true current state. If however a is low then the manager's report will, when $\lambda_t \geq \bar{\lambda}$, be relatively uninformative and as a result the expected direction of learning will be determined by θ_{t-T} rather than θ_t . This is formalized in the following proposition.

Proposition 7 *Suppose that $\delta \approx 0$, then there exists an $\hat{a} \in [0, 1]$ such that for $a \leq \hat{a}$*

$$E(\Delta_t | \theta_t = L, \theta_{t-T} = H, \lambda_{t-T} > \bar{\lambda}, \lambda_t \geq \bar{\lambda}) \geq 0$$

$$E(\Delta_t | \theta_t = H, \theta_{t-T} = L, \lambda_{t-T} > \bar{\lambda}, \lambda_t \geq \bar{\lambda}) \leq 0$$

Proposition 7 is proved in the appendix. If $a \leq \hat{a}$ then the belief will, in expectation continue to move in the direction of the old state. This scenario is illustrated in Figure 1-7. The diagram shows a realized high path from A to C . The expected speed of learning along this path is indicated by the slope of the arrow at B . At C the aggregate state changes from high to low. The resulting expected speed of learning is indicated by the slope of the dashed line between

C and D . Expected speed of learning is less than before but remains positive for the next T periods because learning is, in expectation, dominated by the realization of results which are coming from the past. At D the second component of learning finally reacts to the change in the state and the belief, in expectation, begins to fall. We are able to draw this expected path as two straight lines because, as drawn, the stochastic correspondence remains above $\bar{\lambda}$. This contrasts with the same scenario under direct learning where, in expectation the change in the belief responded immediately to a change in the state. Note that when a is large enough so that $\Phi^M(\theta_t)$ dominates $\Omega(\theta_{t-T})$ then the speed at which the belief moves in the direction of the new state is, in expectation, still slowed by the force of information coming from the past.

Since manipulation is a phenomenon which the economy experiences only when the belief is high then it is only during such episodes that the economy is slow to detect changes in the state. We therefore can have scenarios where, at the end of a boom the belief continues to rise despite a fall in the state. There is no counterpart in bad times, the belief will not, in expectation, continue to fall when the state improves.

Having studied the phases of learning when the belief is low (direct) and high (delayed) we now consider the properties of learning when the belief is moving between high and low. We will see that learning is slow when the belief rises and intense when it falls.

1.3.7 Slow Learning

Suppose now that the true aggregate state is high so that the belief is rising over time. We will see that the speed of learning will slow as the belief rises above $\bar{\lambda}$. To begin with, define the following realized path of beliefs.

Definition 8 *A Rising Path is a sequence of realized public beliefs $\{\lambda_i\}_j^k$ for which \exists an n such that $\max\{\lambda_i\}_{j-T}^n < \bar{\lambda}$ and $\min\{\lambda_i\}_{n+1}^k \geq \bar{\lambda}$.*

Observe that on a rising path $\{\lambda_i\}_j^n$ satisfies the conditions for a low path and hence will display the properties of direct learning. For $\{\lambda_i\}_j^n$ the only source of learning will be the ungarbled signals of managers. When the belief rises above $\bar{\lambda}$ then the source of learning will change. For the next T periods $\{\lambda_i\}_{n+1}^{n+1+T}$ the only source of public information will be the garbled signals of managers. Formally, on a Rising Path the expected speed of learning at any

point in $\{\lambda_i\}_{n+1}^{n+1+T}$ will be

$$E(\Delta_i | \theta_t, \theta_{t-T}) = \Phi^M(\theta_t)$$

where $\Phi^M(\theta_t) \equiv [1 - a(1 - p^{\theta_t})] \lambda^{s'_h M} + a(1 - p^{\theta_t}) \lambda^{s'_i}$

This is simply the first component of learning that we studied in the delayed learning phase. The expected speed of learning here is less (in magnitude) than under either direct or delayed learning. In slow learning the belief always moves in expectation in the direction of the true current state. However observe that if the manipulation problem is pronounced (when a is small) then the speed of learning in this phase goes to zero. For $\{\lambda_i\}_{n+1}^k$ that the realized path is now a high path and exhibits the properties of delayed learning.

Figure 1-8 illustrates the slow down in learning which occurs on a Rising Path. Observe first that the realized path from A to G satisfies the conditions for a Rising Path. The path from A to C satisfies the conditions for a low path and hence the economy experiences direct learning here: the source of learning is simply the ungarbled report of each manager. The speed of learning is shown at a point such as B . At C the aggregate belief rises above $\bar{\lambda}$ and the result is to change the type of incentives which owners offer managers. Now managers are given incentives which encourage them to distort their perceived short-term performance. For the next T periods from C to E social learning is only based upon these garbled signals the speed of learning slows to the slope shown at D . After T periods at E the economy enters a phase of delayed learning. Now the information from manager's garbled signals are augmented from the realization of results from the firm created T periods earlier. It is only after having been above $\bar{\lambda}$ for T periods that results become informative. From E to G the economy experiences delayed learning and the expected speed of learning at any point on this path is indicated at F . On a rising path learning slows because as the belief crosses $\bar{\lambda}$ the timing of the release of manager's private information about the state is pushed off into the future. When the belief falls this process is unwounded and the economy experiences a period of intense learning.

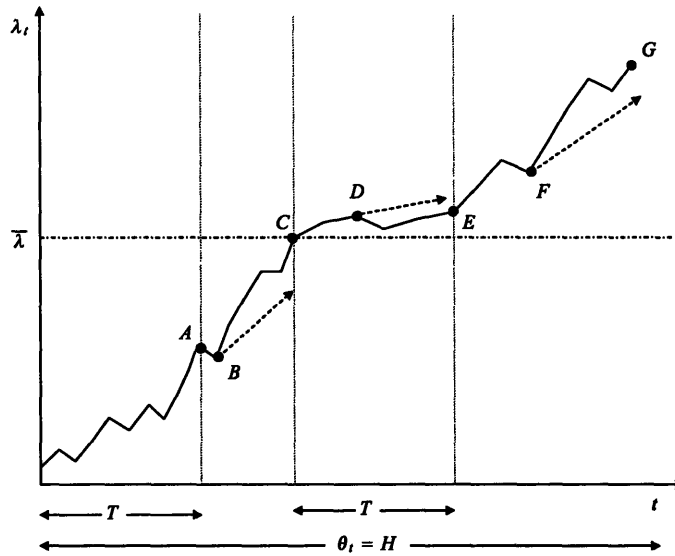


Figure 1-8: Learning Slows as the Belief Rises

1.3.8 Intense Learning

The phase of intense learning is the counterpart to slow learning. Suppose that the state is low so that the belief is falling over time. As λ_t falls below $\bar{\lambda}$ managers will be given contracts which do not induce them to manipulate their signal. As a result, the timing of the release of their private information will be brought forward and will be coupled with the delayed information coming from the results of firms created when the belief was above $\bar{\lambda}$. To study this phase define the following path.

Definition 9 A *Falling Path* is a sequence of realized public beliefs $\{\lambda_i\}_j^k$ for which \exists an n such that $\min \{\lambda_i\}_{j-T}^n \geq \bar{\lambda}$ and $\max \{\lambda_i\}_{n+1}^k < \bar{\lambda}$.

First observe that $\{\lambda_i\}_j^n$ is a high path and hence will display the properties of delayed learning. Over this region learning will come from manager's garbled signals as well as from results. When the belief falls below $\bar{\lambda}$ the contracts offered to managers will change, and managers will begin to truthfully reveal their private information. For $\{\lambda_i\}_{n+1}^{n+1+T}$ public information will then come from the truthful signal of firm t as well as the result of firm $t - T$. In effect this is

direct learning *plus* additional information from the past. Over this region the expected speed of learning is

$$\begin{aligned}
E(\Delta_t | \theta_t, \theta_{t-T}) &= \Phi^T(\theta_t) + \Omega(\theta_{t-T}) \\
\text{where } \Phi^T(\theta_t) &\equiv \begin{cases} (2p^H - 1) \lambda^{s_h^T} & \text{if } \theta_t = \theta^H \\ -(2p^H - 1) \lambda^{s_h^T} & \text{if } \theta_t = \theta^L \end{cases} \\
\Omega(\theta_{t-T}) &\equiv \left[p^{\theta_t} q^h + (1-a)(1-p^{\theta_t})q^l \right] \lambda^Y \\
&+ \left[p^{\theta_t} (1-q^h) + (1-a)(1-p^{\theta_t})(1-q^l) \right] \lambda^0
\end{aligned}$$

As before we have that

$$\begin{aligned}
\Phi^T(\theta_t = H) &> 0, \quad \Phi^T(\theta_t = L) < 0 \\
\Omega(\theta_{t-T} = H) &> 0, \quad \Omega(\theta_{t-T} = L) < 0
\end{aligned}$$

The first term $\Phi^T(\theta_t)$ represents the expected change in the belief due to the truthful signal of the current manager. The second term $\Omega(\theta_{t-T})$ represents the expected change in the belief coming from the results of the firm created T periods ago. It is easy to show that the first term will always dominate and hence during a phase of intense learning the belief will, in expectation, always move in the direction of the current state θ_t . Suppose that the state is unchanged over the last T periods then it is clear that the speed of learning is faster than under direct learning. The difference in speed is simply $\Omega(\theta_{t-T})$. This difference is stronger the more informative are results. When a is small, for example, results become more informative and the speed of learning in a phase of intense learning is, as a result, all the stronger.

Figure 1-9 illustrates the learning dynamics on a Falling Path A to G . From A to C the economy experiences delayed learning. The associated speed is show by the slope at B . From C to E the economy experiences intense learning. Now information comes from truth signals as well as results from the past. The speed of learning accelerates and is shown by the slope at D . After the belief has been below $\bar{\lambda}$ for T periods the results from the past are exhausted. From E to G the economy returns to direct learning. Results are no longer informative because the managers of firms created T periods fully revealed their private information with their signal.

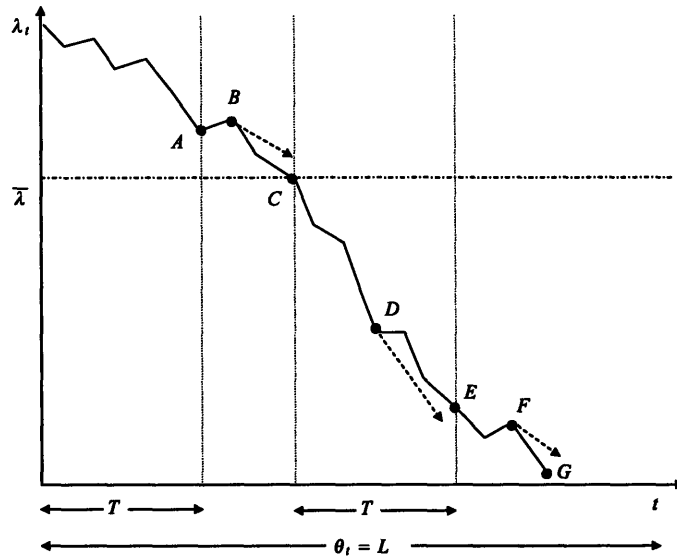


Figure 1-9: Learning Accelerates as the Belief Falls

The episode of intense learning explains why busts are precipitous. Its a time when we discover the deception of the boom and begin to sign contracts which no longer encourage such actions. This shifts the timing of information forward and creates an period when the information from the boom overlaps with the information from the bust.

1.4 Stock Market and Investment Dynamics

1.4.1 Stock Market

The previous section showed how the amount of information released into the economy changes systematically over the cycle. This was shown to produce a path for the belief characterized by slow booms and precipitous declines. We now want to show how these movements in the belief affect the value of the stock market. Our original observation was that the stock market rose gradually over the course of the 1990s and then fell much faster in the years after. So far the model has suggested an explanation to this observed fact: more information came out during the crash due to a shift in the timing of the release of manager's private information. Indeed

this section will confirm that logic. In addition, there is a second force at work. The impact that a change in the belief has upon the value of the market changes over time. In particular this depends upon how many firms are in existence who are suspected to have released manipulated signals. It is these firms whose value is most sensitive to changes in the belief. For a firm suspected of manipulation, additional information about the state help investors learn about the true performance of these firms. It is thus at the end of a boom, when the economy has many firms whose value is made ambiguous by manipulation that the stock market is most vulnerable to movements in the belief. So at the start of a collapse not only is learning strong but the number of firms whose value it affects is higher.

To study the stock market I want to talk about the perceived value of the stock of firms that is in existence. I abstract from, risk aversion and discounting. The value of a firm at a point in time t is defined as the expected value of its output less any payments it may have to make to the manager. This expectation will always be taken at the beginning of period t and is based all the information contained in H_t . Thus it does not include any information which has not been publicly released and it is not based upon the true value of the state since both of these are from the perspective of an investor unobservable.

The Value of an Existing Firm

To begin studying the value of the stock market we need to characterize the value of each firm that is in existence. A firm is said to be in existence at (the start of) period t if it has already entered the economy but has yet to realize its output. For a firm j to be in existence at t it must be that $t - T \leq j \leq t - 1$. Observe that the OLG structure of the firms ensures that there are always T firms in existence at the beginning of any period.

In order to keep track of the characteristics of the firms in existence at any period define the following indicator variables:

$$\vartheta_j = \begin{cases} 1 & \text{if } \lambda_j \geq \bar{\lambda} \\ 0 & \text{if } \lambda_j < \bar{\lambda} \end{cases}, \quad \varsigma_j = \begin{cases} 1 & \text{if } s'_j = s'_h \\ 0 & \text{if } s'_j = s'_l \end{cases}$$

The first indicator variable ϑ_j records the type of incentives offered to the manager of firm j . If $\vartheta_j = 1$ ($\vartheta_j = 0$) this indicates that the manager of firm j was given incentives which

induce (discourage) manipulation. The second indicator variable ς_j records the signal that the manager of firm j released. If $\varsigma_j = 1$ ($\varsigma_j = 0$) this indicates that the manager of firm j released the high (low) signal. The value of any firm in existence will depend upon both these variables.

Consider a firm j for which $\vartheta_j = \varsigma_j = 0$. The manager of this firm was given incentives which discouraged manipulation. Subsequently the manager revealed that the firm's short-term performance was low. Its expected profit is

$$v_{00} \equiv q^m Y - \frac{q^l B}{\eta(q^h - q^l)}.$$

Similarly, consider a firm j for which $\vartheta_j = 0$ and $\varsigma_j = 1$. The manager of this firm was given incentives which discouraged manipulation. Subsequently the manager revealed that the firm's short-term performance was high. It has a value

$$v_{01} \equiv q^h \left(Y - \frac{B}{\eta(q^h - q^l)} \right).$$

Next consider a firm j for which $\vartheta_j = 1$ and $\varsigma_j = 0$. The manager of this firm was given incentives which encouraged manipulation. However the manager must have observed low performance but was unable to hide it. When the manager releases the low signal this then fully reveals the true short-term performance of the firm. The value of this firm is simply

$$v_{10} \equiv q^m Y$$

Notice that each of these three values $\{v_{00}, v_{01}, v_{10}\}$ is a constant. All three are unaffected by movements in the belief. The reason is that in each case the true short-term performance of the firm is public knowledge after the manager releases their signal. Knowledge about the aggregate state is only useful in so far as it provides information about the likely short-term performance of the firm. This contrasts with the final case.

Consider a firm j for which $\vartheta_j = \varsigma_j = 1$. The manager of this firm was given incentives which encouraged manipulation. Subsequently the manager reported the high signal. In this scenario it is unclear whether the firm is in fact performing well or if the firm is performing

poorly and the manager has been able to hide this. At time t the value of such a firm is

$$\begin{aligned} \widehat{v}_j(H_t) &= \Phi_j(H_t) v_{11} \\ \text{where} \quad \Phi_j(H_t) &\equiv \Pr(y_j = Y | H_t) \\ v_{11} &\equiv \left(Y - \frac{B}{\eta[q^h - (1-a)q^l]} \right) \end{aligned}$$

$\Phi_j(H_t)$ is the assessed probability, at the start of period t , that firm j will produce positive output conditional upon the observed history H_t . If the firm does achieve positive output then the firm's profit will be v_{11} . Investors continue to use the information that is released after the firm's entry to evaluate the probability that the high signal was truthful. It is only firms who are suspected of manipulation whose value continues to change over the course of their life. For $\delta \approx 0$ we can approximate¹⁹ this assessed probability by

$$\Phi_j(H_t) \approx \widetilde{\Phi}(\mu_t) \equiv \frac{q^h p_t + q^l (1-p_t)(1-a)}{1-a(1-p_t)}.$$

Observe that this is strictly increasing in the belief

$$\frac{\partial \widetilde{\Phi}(\mu_t)}{\partial \mu_t} = \frac{(1-a)(p^H - p^L)(q^h - q^l)}{[1-a(1-p_t)]^2} > 0.$$

Thus an increase (decrease) in the belief will raise (lower) the value of such a firm. When the belief increases, it is more likely (from the perspective of an outsider) that the manager did in fact observe high performance and truthfully reported it. Conversely, when the belief falls investors become more convinced that the manager of the firm manipulated their signal and that the firm was in fact performing poorly. As a result the assessed value of the firm falls.

Putting these four scenarios together we can write that at the start of period t the value of any firm that is in existence is

$$\begin{aligned} v_j(\vartheta_j, \varsigma_j, H_t) &= \vartheta_j [\varsigma_j \Phi_j(H_t) v_{11} + (1-\varsigma_j) v_{10}] \\ &\quad + (1-\vartheta_j) [\varsigma_j v_{01} + (1-\varsigma_j) v_{00}] \end{aligned}$$

¹⁹To see this observe that $\Phi_j(H_t) = \frac{q^h p_j(H_t) + q^l (1-p_j(H_t))(1-a)}{p_j(H_t) + (1-p_j(H_t))(1-a)}$ where $p_j(H_t) \equiv p^H \Pr(\theta_j = \theta^H | H_t) + p^L (1 - \Pr(\theta_j = \theta^H | H_t))$. If $\delta \approx 0$ then it follows that $p_j(H_t) \approx p_t$ and hence the approximation follows.

Stock Market Dynamics

Having defined the value of any firm in existence at t we now turn to studying the way that the value of the stock market changes over time. Let $V_t(H_t)$ denote the average value of the stock of firms in existence at the start of period t :

$$V_t(H_t) = \frac{1}{T} \sum_{j=t-T}^{t-1} v_j(\vartheta_j, \varsigma_j, H_t)$$

We now ask how this will change over time. The change in the stock market from one period to the next is

$$\begin{aligned} V_{t+1}(H_{t+1}) - V_t(H_t) &= \Delta V_{t+1}^{Churn} + \Delta V_{t+1}^{Exist} \\ \text{where } \Delta V_{t+1}^{Churn} &\equiv \frac{1}{T} [v_t(\vartheta_t, \varsigma_t, H_{t+1}) - v_{t-T}(\vartheta_{t-T}, \varsigma_{t-T}, H_t)] \\ \Delta V_{t+1}^{Exist} &\equiv \frac{1}{T} \sum_{j=t-T+1}^{t-1} [v_j(\vartheta_j, \varsigma_j, H_{t+1}) - v_j(\vartheta_j, \varsigma_j, H_t)] \end{aligned}$$

The change in the value of the stock market has two components. The first term ΔV_{t+1}^{Churn} indicates that the value of the market will change because of the OLG nature of the production structure. During period t firm t will be created and firm $t - T$ will exit. Thus the value of the market will be changed by the difference in the value of the entering and exiting firm. This term is of the order of magnitude of T^{-1} and thus should be thought to be small.

The second term ΔV_{t+1}^{Exist} represents how the value of the $T - 1$ firms who are in existence both at t and $t-1$ will change over the period. This change comes from the fact that information is released during the period and this may cause investors to reevaluate the value of these firms. Recall from above that the value of a firm only changes if investors are uncertain about its true short-term performance. That is, if $\vartheta_j = \varsigma_j = 1$. These firms are suspected of having manipulated their signal and so investors will use the information which is released during period t to re-evaluate the value of such a firm. To see this, observe that

$$\Delta V_{t+1}^{Exist} \equiv \frac{1}{T} v_{11} \sum_{j=t-T+1}^{t-1} \vartheta_j \varsigma_j [\Phi_j(H_{t+1}) - \Phi_j(H_t)]$$

Any firm that is in existence that for which $\vartheta_j \varsigma_j = 0$ then the true short-term performance of the firm was fully revealed with the manager's signal. Accordingly, their value is unaltered by any further information about the aggregate state and thus does not change from period t to period $t - 1$. It follows that the strength of this effect will depend upon the number of firms who are in existence who are suspected of having manipulated their signal. The number of such firms in existence at the start of period t and period $t + 1$ is.

$$z_t \equiv \sum_{j=t-T+1}^{t-1} \vartheta_j \varsigma_j$$

At any point on a low path we must have that $z_t = 0$ and so it follows that $\Delta V_{t+1}^{Exist} = 0$. Thus when the belief is low the stock of firms have an unambiguous value. The manager of each firm has been given incentives which discourage manipulation and so the value of each firm is fully revealed when the manager releases their signal. Thus, the release of information on a low path does not move the stock market over and above the way it relates to changes in ΔV_{t+1}^{Churn} . When $z_t \geq 1$ then the information which is released during the period will be used to update the assessed valuation of these z_t existing firms. A useful way to approximate ΔV_{t+1}^{Exist} is to use the fact that $\Phi_j(H_t) \approx \tilde{\Phi}(\mu_t)$. This gives us that

$$\Delta V_{t+1}^{Exist} \approx \frac{1}{T} v_{11} z_t \left[\tilde{\Phi}(\mu_{t+1}) - \tilde{\Phi}(\mu_t) \right]$$

If the information released during period t increases the belief then the value of these z_t will rise also. In this case it is more likely that the manager of any of these firms has in fact truthfully reported high performance. Conversely if bad news is released during the period then it is more likely that these managers have been deceiving the market. Their value is thus decreased as a result.

There are two sources of variation in ΔV_{t+1}^{Exist} . The first of these is that the amount of information which is released in a period changes over time. The previous section already showed that the magnitude of changes in the belief will vary with the cycle. Learning is slow during a boom and thus $\tilde{\Phi}(\mu_{t+1}) - \tilde{\Phi}(\mu_t)$ will be small. Conversely, when the market crashes learning is intense. Thus more information is released at these times and $\tilde{\Phi}(\mu_{t+1}) - \tilde{\Phi}(\mu_t)$ is

large. We saw that on a Falling Path the economy goes through a period of intense learning. The effect of this will be to generate large downward revisions in the estimated value of the current stock of existing firms.

The second source of variation in ΔV_{t+1}^{Exist} comes from z_t . The number of firms whose value is re-evaluated with changes over the cycle. As already mentioned, z_t is zero at any point on a Low Path. At any point that does not satisfy the conditions of a Low Path then there exists at least one firm in existence who was given incentives which encouraged manipulation. The fraction of these firms who have an uncertain value depends upon the number who were able to report the high signal (either truthfully or through manipulation). Conditioning on the true history of the state we can write the expected value of z_t as

$$E(z_t|\Theta_t) = \sum_{j=t-T+1}^{t-1} \vartheta_j \left[1 - a \left(1 - p^{\theta_j} \right) \right]$$

At any point on a High Path the manager of every firm in existence has been given incentive which induce manipulation. Of these a fraction will have released the high signal: either truthfully or through manipulation. If the value of the state has been unchanged for the last T periods then the expected value of z_t is

$$\begin{aligned} z^H &\equiv (T-1) [1 - a (1 - p^H)] \text{ if } \theta_t = \theta_{t-1} = \dots \theta_{t-T} = \theta^H \\ z^L &\equiv (T-1) [1 - a (1 - p^L)] \text{ if } \theta_t = \theta_{t-1} = \dots \theta_{t-T} = \theta^L \end{aligned}$$

Figure 1-10 shows how $E(z_t|\Theta_t)$ evolves over time. The top panel of Figure 1-10 draws an example realized path for the belief. Below this panel the true value of the state is shown. The dynamics of the belief correspond to the phases of learning studied in the previous section. Here we can see how the number of firms whose value is altered by the belief changes during these phases. The Figure is drawn supposing that the economy begins on a low path. Hence z_t is zero until the belief rises above $\bar{\lambda}$. This occurs at A . At this point owners begin to offer managers incentives which encourage manipulation. Over the next T periods the stock of firms changes so that by B the manager of each firm has been given the same incentives. Thus at B the expected number of firms whose value is changing with the belief is z^H . While the state

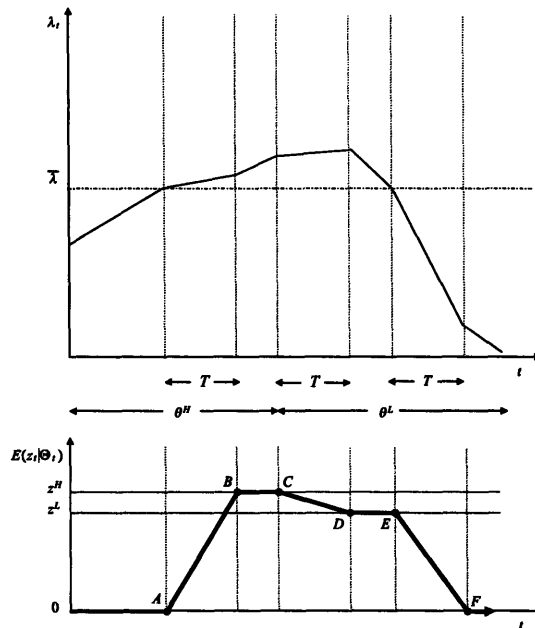


Figure 1-10: $E(z_t|\Theta_t)$ Over the Cycle

remains high the expected value of z_t remains at this level. At C the true value of the state changes from high to low and thus the proportion of firms who release the high signal will fall in expectation. Over the next T periods $E(z_t|\Theta_t)$ will fall to D . The expected number of firms whose value is uncertain will then remain at this level until the belief crosses $\bar{\lambda}$. This occurs at E . At this point managers are now no longer given incentives which induce manipulation. Accordingly the number of these firms will, with certainty, return to zero at F .

The central insight here is that the way that changes in the belief affect the stock market changes over time. In bad times when managers are discouraged from manipulation investors are well informed about the value of every firm. It follows that when information about the state is released it has a small impact upon the market. This model shows an extreme case where when the belief is low the belief does not move the market at all. The opposite is true when the belief is high. Now the stock market is full of firms who have been given short-term incentives which encourage manipulation. As a result investors are more uncertain about their value. When information about the true value of the state is released it causes investors to

update their belief about the truth of any given manager's signal. At the end of a boom there are many of these firms. Thus when bad news starts to come it causes large movements in the stock market. This is consistent with the findings of Graham, Litan and Sukhtankar (2002) who estimate the impact of the Enron and Worldcom scandals upon the US economy. They argue that these events had wide reaching effects because they sparked a crisis in confidence for all firms. Their estimates suggest that the total loss in stock market wealth attributable to these events was at least 10 percent. At the end of the boom of the 1990s, many managers were suspected of having manipulated their signals. For this reason the stock market was so sensitive to news about the aggregate state.

Our study of the endogenous release of information showed that a lot of information is released into the economy between E and F (on Figure 1-10). We see now that at the start of this episode the number of firms whose value is moved by the news is (in expectation) z^L . It is thus during a boom that the stock market is most vulnerable to bad news. This mechanism thus complements the dynamics of learning studied earlier and shows how the stock market can under go major downward revisions at the end of boom.

1.4.2 Investment

We now want to show that movements in the belief can easily be translated into changes in aggregate investment. The experience of investment in the US economy over the 1990s was shown above in Figure 1-1. The growth of investment over the boom of the 1990s was slow when compared to the speed with which it fell subsequently.

The previous section of this paper demonstrated the way that the incentives provided to managers induced slow learning when the belief was rising and intense learning when the belief was falling. The purpose of this section is to demonstrate in a simple extension to the basic model that these dynamics for the belief imply similar behavior for investment.

The basic model does not have any implications for investment: it was simply imposed that one firm of fixed size enters in each period. In order to think about investment then I now suppose that at the birth of each firm the owners also select the size of the firm. In particular, let I_t be the amount invested in firm t at the time of its entry. Suppose that there are only two differences in the firm's problem. First, let the final output of the firm now be scaled by I_t so

that it is $I_t Y$ when positive and 0 otherwise. Second suppose now that the private benefit of the manager for exerting low effort is $I_t B$. The fixed scale case analyzed so far is thus just a special case where we restrict $I_t = 1$. It is easy to see that the solution to the firm's problem is the same as in Proposition 1 except now all quantities are scaled by I_t . It follows then that the value of $\bar{\mu}$ is unaffected by the choice of I_t . The learning dynamics of the model are then identical to the fixed scale case.

We then need to ask: how will the choice of scale vary with the belief? For ease of exposition assume that the cost of investment in the firm $C(I_t)$ is quadratic²⁰ so that $C(I_t) = \frac{c}{2} I_t^2$. The optimal choice of I_t will solve

$$\max_{I_t \geq 0} \Pi^*(\mu_t) I_t - C(I_t)$$

where $\Pi^*(\mu_t) = \max \{ \Pi^T(\mu_t), \Pi^M(\mu_t) \}$

The optimal level of investment is

$$I_t^*(\mu_t) = \begin{cases} \omega^H \mu_t + I_1 & \text{if } \mu_t \geq \bar{\mu} \\ \omega^L \mu_t + I_0 & \text{if } \mu_t < \bar{\mu} \end{cases} \quad (1.7)$$

where $I_1 \equiv \frac{1}{c} \Pi^M(\mu_t = 0)$, $I_0 \equiv \frac{1}{c} \Pi^T(\mu_t = 0)$

$$\omega^H \equiv \frac{1}{c} (p^H - p^L) \left[(q^h - (aq^m + (1-a)q^l)) Y - \frac{B}{\eta} \right] > 0$$

$$\omega^L \equiv \frac{1}{c} (p^H - p^L) \left[(q^h - q^m) Y - \frac{B}{\eta} \right] > 0$$

Observe that $I_t^*(\mu_t)$ is strictly increasing in the belief²¹. There is thus a direct mapping from

²⁰The qualitative results of this section simply require that $C(I)$ be strictly convex.

²¹Note that

$$\omega^H - \omega^L = \frac{1}{c} (p^H - p^L) (q^m - q^l) (1-a) > 0$$

so that investment is more responsive to changes in the belief when the belief is high ($\mu_t \geq \bar{\mu}$). To understand this, first consider the case when the belief is low, so that managers are offered contracts which discourage manipulation ($\mu_t < \bar{\mu}$). Here the marginal return to investment is increasing with the belief simply because this increases the probability that the firm will have high short term performance. This, in turn, implies a higher probability of final output. This force is also in effect when the belief is high ($\mu_t \geq \bar{\mu}$) and is augmented by a second force. When manager's are offered contracts which encourage manipulation then an increase in the belief lowers the expected probability that the manager will take value destroying actions to boost the short term performance of the firm. As a result, investment responds more to a change in the belief at times when managers are offered contracts which encourage manipulation.

the dynamics of the belief set out earlier to the dynamics of investment.

The previous section studied the dynamics of the belief over time. The investment function (1.7) shows that changes in the belief can be translated directly into changes in the level of investment. When the economy undergoes a period of Slow Learning when the belief is rising this will translate into slow investment growth. The delayed reaction of the belief to a change in the state when the belief is high translates into an equally delayed response for investment. In an episode like the one illustrated in Figure 1-7 the belief and investment will, in expectation, continue to rise for T periods despite the change in the state from high to low. The rapid decline in the belief which occurs during an episode of intense learning will produce a precipitous collapse in investment.

1.5 Robustness

This section demonstrates that the central results of the model are robust to a number of the assumptions which were made to ease its exposition.

1.5.1 Relaxing the Additivity Assumption

I now demonstrate that the central results obtained in Section 2 were not driven by the fact that effort was assumed to be additive. In particular, I now relax (1.1) and simply impose the following ordering;

$$\begin{aligned}
 p(e_t = 1, \theta_t) &> p(e_t = 0, \theta_t) \text{ for } \theta_t \in \{\theta^H, \theta^L\} \\
 p(e_t, \theta_t = \theta^H) &> p(e_t, \theta_t = \theta^L) \text{ for } e_t \in \{0, 1\}
 \end{aligned}$$

These assumption simply impose that the performance of the firm is increasing in both managerial effort and the aggregate state. It is useful to define the following:

$$\begin{aligned}
 p_t &\equiv \mu_t p(e_t = 1, \theta_t = \theta^H) + (1 - \mu_t) p(e_t = 1, \theta_t = \theta^L) \\
 \sigma_t &\equiv \mu_t p(e_t = 0, \theta_t = \theta^H) + (1 - \mu_t) p(e_t = 0, \theta_t = \theta^L)
 \end{aligned}$$

These are the expected probability of success given that the manager provides high (p_t) or low (σ_t) effort. The owner's problem is completely analogous to the additive case and so I do not write it out again. To best understand how relaxing the additive assumption alters the trade-off involved in contract choice it is useful to start by solving for the optimal contracts which induce the manager to choose $M_t = 1$ and $M_t = 0$. These are solved for in the appendix and are presented in the following Proposition.

Proposition 10 *The optimal contract which induces the manager to select $e_t = 1$ and $M_t = 1$ is*

$$\begin{aligned} w(s'_h, Y, H_t) &= \frac{B}{(p_t - \sigma_t)[q^h - (1-a)q^l]} \\ w(s'_h, 0, H_t) &= w(s'_l, Y, H_t) = w(s'_l, 0, H_t) = 0. \end{aligned}$$

This results in an expected profit to the owners of the firm of

$$\Pi^M(\mu_t) = \left[p_t q^h + (1-p_t)(aq^m + (1-a)q^l) \right] Y - \frac{B}{(p_t - \sigma_t)} \left(p_t + \frac{(1-a)q^l}{q^h - (1-a)q^l} \right)$$

The optimal contract which induces the manager to select $e_t = 1$ and $M_t = 0$ is

$$\begin{aligned} w(s'_l, Y, H_t) &= w(s'_l, 0, H_t) = \frac{q^l B}{(p_t - \sigma_t)(q^h - q^l)} \\ w(s'_h, Y, H_t) &= \frac{B}{(p_t - \sigma_t)(q^h - q^l)} \\ w(s'_h, 0, H_t) &= 0 \end{aligned}$$

This results in an expected profit to the owners of the firm of

$$\Pi^T(\mu_t) = \left[p_t q^h + (1-p_t)q^m \right] Y - \frac{B}{(p_t - \sigma_t)} \left(p_t + \frac{q^l}{q^h - q^l} \right)$$

The optimal contracts in each case take the same form as in the additive case. The key difference is that the trade-off between whether or not to discourage manipulation has changed slightly. In the additive case the difference between $\Pi^M(\mu_t) - \Pi^T(\mu_t)$ is linear and strictly downward sloping. This guarantees the existence of a unique $\bar{\mu} \in R$. Now however, $\Pi^M(\mu_t) - \Pi^T(\mu_t)$ is neither downward sloping or continuous over the entire range of $\mu_t \in R$. If however we

focus our attention on the region of $\mu_t \in [0, 1]$ then the trade-off between these contract choices operates in the same direction under a simple condition. This is established in Proposition 11 and is proven in the appendix.

Proposition 11 $\frac{\partial(\Pi^M(\mu_t) - \Pi^T(\mu_t))}{\partial\mu_t} > 0 \forall \mu_t \in [0, 1]$ if and only if the following condition holds:

$$(p^L - \sigma^L)^2 > \frac{aBq^l q^h [(p^H - \sigma^H) - (p^L - \sigma^L)]}{(q^h - q^l)(q^h - (1-a)q^l)Y(1-a)(q^m - q^l)(p^H - p^L)}.$$

To interpret this condition begin by observing that it holds whenever $(p^L - \sigma^L) \geq (p^H - \sigma^H)$. The difference $(p^i - \sigma^i)$ measures the degree of moral hazard in state i . If for example $p^i - \sigma^i = 1$ then (absent the ability to manipulation) there is in effect no moral hazard in that state since the short-term performance of the firm fully reveals the action of the manager. The condition then simply requires that the degree of moral hazard not be much more severe in the low state than in the high state. That is $(p^H - \sigma^H)$ can be “too large” relative to $(p^L - \sigma^L)$. If this were the case, it may be optimal to allow manipulation when the belief is low because the degree of moral hazard was so much more severe in that state and the short-term contract would be a less expensive way of inducing effort.

If we assume that the condition in Proposition 11 is satisfied then we simply need to assume that

$$\begin{aligned} \Pi^T(0) &> \Pi^M(0) \\ \Pi^T(1) &< \Pi^M(1) \end{aligned}$$

to ensure that there exists a unique $\bar{\mu} \in (0, 1)$ such that it is optimal to induce the manager to select $e_t = 1$ and $M_t = 0$ for $\mu_t < \bar{\mu}$ and to choose $e_t = 1$ and $M_t = 1$ when $\mu_t \geq \bar{\mu}$. The most important qualitative feature of the additive case is preserved: owners offer managers high powered incentives when the belief is high. As before high powered incentives are (in expectation) less costly when μ_t is large because the manager is less likely to privately observe their firm performing poorly. Now that effort is not additive we simply need to ensure that the degree of moral hazard is not much stronger in the low state or else high powered incentives may become optimal when μ_t is low.

1.5.2 Contracting on Aggregate Variables

Until now I have assumed that the incentive contracts offered to managers cannot be made contingent on events which occur after t . However in an optimal contracting environment owners may wish to make the payment to the manager conditioned upon events which are informative for the state at t . The purpose of this section is to demonstrate that allowing contracts to be written contingent on aggregate information will not change the fundamental features of the equilibrium learning dynamics studied above.

In equilibrium the amount of information which is potentially available in the future will be determined by the actions of other firms. To begin with however, suppose that firm t 's owners have available to them a public signal x which can take on N_t possible values. Let α_i^j be the probability that x takes on the value j if the state at t is i . Order these signals so that

$$\frac{\alpha_H^1}{\alpha_L^1} < \frac{\alpha_H^2}{\alpha_L^2} < \dots < \frac{\alpha_H^{N_t}}{\alpha_L^{N_t}}.$$

By this ordering x^1 is the most informative signal that the state is low and x^{N_t} is the strongest signal that the state is high. We now allow the contracts which managers write be of the form $w(s', y, H_t, x^j)$. That is, for a given history they make a payment to the manager conditional upon the manager's signal, the realized level of output and the outcome of the public signal. Otherwise the firm's problem is the same as before. First consider the optimal contract which allows manipulation.

Proposition 12 *The optimal contract that induces the manager to select $e_t = 1$ and $M_t = 1$, and is written contingent on the aggregate signal, sets*

$$w(s'_h, Y, H_t, x^1) = \frac{B}{\eta(q^h - (1-a)q^l)(\mu_t\alpha_H^1 + (1-\mu_t)\alpha_L^1)}$$

and all other payments to zero. This results in an expected profit to the owners of

$$\begin{aligned} \Pi^{MC}(\mu_t) &= \left[p_t q^h + (1-p_t)(aq^m + (1-a)q^l) \right] Y \\ &\quad - \frac{B}{\eta} \left(\frac{(1-a)q^l}{q^h - (1-a)q^l} + \frac{p^H \mu_t \alpha_H^1 + p^L (1-\mu_t) \alpha_L^1}{\mu_t \alpha_H^1 + (1-\mu_t) \alpha_L^1} \right) \end{aligned}$$

This is proved in the appendix. It is easy to check that the expected profits here are strictly higher than in the case where contracting on the signal was not permitted. This contract has a familiar form. It pays the manager only when the firm performs well and the signal indicates that the state is most likely low. This is the standard relative performance evaluation motive. Assuming that effort makes an additive contribution to the probability of the firm's success ensures that high performance in the low state is the most informative signal of managerial effort.

Next, consider the optimal contract which discourages manipulation. Now the ability to write contingent contracts serves two roles. The first standard function is to direct payments to the state in which high performance is the most informative for effort. The second function is to exploit the private information of the manager. When a manager observes that their private signal is low then their private belief is updated and they are less optimistic about the aggregate state. It follows that a manager who is deciding whether or not to manipulate is now doing so with a belief below μ_t . We saw above that in order to discourage manipulation, managers were paid a fee for announcing poor performance. Owners can reduce the ex ante expected cost of this fee by making it positively correlated with the manager's private information when they observe the low signal. This force encourages the owners to pay for low performance contingent on the state being low and to pay for high performance contingent on the state being high. There are thus two countervailing forces in choosing between $w(s'_h, Y, H_t, x^1)$ and $w(s'_h, Y, H_t, x^N)$. Paying $w(s'_h, Y, H_t, x^1)$ has the standard benefit of relative performance evaluation by making the payment to the manager in the most informative state of the world. However paying $w(s'_h, Y, H_t, x^N)$ has the benefit that this is a payment less desirable to a manager once they have observed a low signal privately. In turn, this will make it less expensive to deter the manager from manipulation. The optimal contract must weigh these forces. This is summarized in the following proposition.

Proposition 13 *If (1.39) and (1.40) holds, then the optimal contract that induces the manager to select $e_t = 1$ and $M_t = 0$, and is written contingent on the aggregate signal, sets*

$$w(s'_h, Y, H_t, x^1) = \frac{B}{\eta(q^h - q^l)(\mu_t \alpha_H^1 + (1 - \mu_t) \alpha_L^1)}$$

$$w(s'_l, Y, H_t, x^1) = w(s'_l, 0, H_t, x^1) = \frac{q^l B}{\eta(q^h - q^l)(\mu_t \alpha_H^1 + (1 - \mu_t) \alpha_L^1)}$$

and all other payments to zero. This results in an expected profit to the owners of

$$\Pi^{TC}(\mu_t) = [p_t q^h + (1 - p_t) q^m] Y - \frac{B}{\eta} \left(\frac{q^l}{q^h - q^l} + \frac{p^H \mu_t \alpha_H^1 + p^L (1 - \mu_t) \alpha_L^1}{\mu_t \alpha_H^1 + (1 - \mu_t) \alpha_L^1} \right)$$

We can now study whether the firm's owners will elect to discourage manipulation or not. Inspecting the expressions for $\Pi^{MC}(\mu_t)$ and $\Pi^{TC}(\mu_t)$ we see immediately that the term value of contracting on aggregate information has exactly the same benefit to each contract. As a result we are left with exactly the same trade-off that was studied above. The cutoff $\bar{\mu}$ is exactly as before and hence the model will exhibit exactly the same learning dynamics.

1.5.3 Endogenizing a

Until now the manager's ability to manipulate their signal has been taken as given. However in reality many of the variables which affect this probability can be controlled by the owners of a firm. Most obviously they can choose how closely the manager is monitored by the board and the firm's auditors. In this subsection I show the central features of the model continue to hold where a is selected optimally by the owners of the firm at the same time that they set the incentive contract of the manager. I assume that the owners are able to credibly commit to an observed level of a .

Assume that increasing a is costly. In particular suppose that owners must pay $F(a)$ where

$$F(a) = 0 \text{ for } a \in [0, \hat{a}]$$

$$F'(\hat{a}) = 0$$

$$F''(a) > 0 \text{ for } a \geq \hat{a}$$

and that $F(1)$ is prohibitively expensive. Here $\hat{a} \in (0, 1)$ is a base level of regulation of the manager and owners must pay to increase a above this base. The assumptions on the shape of $F(a)$ guarantee a unique interior solution.

We now ask: what is the owner's optimal selection of incentive scheme and level of auditing. Suppose first that the optimal incentive scheme discourages manipulation. The expected payoff

to the owners in this case is $\Pi^T(\mu_t)$ as given in Proposition 1. These expected profits are unchanged by the level of a and hence if these incentives are ever optimal then the optimal choice of auditing is $a = \hat{a}$. The intuition is that in this case the weakened incentives rule out the possibility of manipulation. Notice that increasing a does not help to discourage manipulation. For this reason increasing a will not alter the cost of incentives for the contract. In addition, since the manager is discouraged from distorting their signal then increasing a can have no effect upon the expected output of the firm. Allowing owners to invest in setting $a \geq \hat{a}$ therefore does not alter this contract. It will however alter the region of beliefs over which this contract is optimally altered (that is change $\bar{\mu}$).

Suppose now that the optimal incentives to the manager do encourage manipulation. The expected profit to the firm in this case $\Pi^M(\mu_t)$ is strictly increasing in a . Let $a^*(\mu_t)$ denote the optimal choice of a for a given level of the belief. This is characterized by the following first order condition.

$$F'(a^*(\mu_t)) = (1 - p_t)(q^m - q^l)Y + \left(\frac{B}{\eta}\right) \left[\frac{q^h q^l}{(q^h - (1 - a^*(\mu_t))q^l)^2} \right] \quad (1.8)$$

The right hand side of (1.8) is simply the marginal benefit of raising a (that is $\frac{\partial \Pi^M}{\partial a}$). The first term on the right hand side of (1.8) is the increase in expected output by increasing a . Since the manager manipulates whenever they are able, then raising a limits the probability of the manager taking these value destroying actions. This term is unambiguously positive. The second term represents the fact that raising a lowers the expected cost of incentives. By raising a the owners make the signal of the manager a more informative for the manager's effort and thus reduce the severity of the moral hazard problem. This term is also unambiguously positive. Finally, notice that the right hand side of (1.8) is strictly decreasing in a^* and the left hand side is strictly increasing in a^* and hence the first order condition characterizes a unique optimal choice of a^* for any given level of the belief μ_t provided that a contract which allows manipulation is in fact optimal.

If we continue to suppose that the optimal contract allows manipulation we can now ask

how the choice of a^* will vary with the belief. Totally differentiating (1.8) gives us that

$$\frac{da^*(\mu_t)}{d\mu_t} = \frac{-[(p^H - p^L)(q^m - q^l)Y]}{F''(a^*) + \left(\frac{2Bq^h}{\eta}\right) \left[\frac{(q^l)^2}{(q^h - (1 - a^*(\mu_t))q^l)^3}\right]} < 0 \quad (1.9)$$

The optimal degree of auditing is strictly decreasing in the belief. The reason is that the probability that a manager finds their firm performing poorly and hence wants to manipulate their signal is strictly decreasing in μ_t . Thus the benefit to deterring costly manipulation falls as the belief increases.

Finally we need to show how the choice of a affects the choice of incentives. The key aim here is to show that the choice is summarized by a belief $\bar{\mu}$ for which incentives (fully) discourage manipulation only when the belief is below this cutoff. Observe that

$$\frac{d\Pi^T(\mu_t)}{d\mu_t} = (p^H - p^L)(q^h - q^m)Y - \left(\frac{B(p^H - p^L)}{p_t - \sigma_t}\right)$$

and

$$\frac{d\Pi^M(\mu_t, a^*(\mu_t))}{d\mu_t} = (p^H - p^L)(q^h - (a^*(\mu_t)q^m + (1 - a^*(\mu_t))q^l))Y - \left(\frac{B(p^H - p^L)}{p_t - \sigma_t}\right)$$

where the second equality follows from the envelope theorem. It follows then that

$$\frac{d\Pi^M(\mu_t, a^*(\mu_t))}{d\mu_t} - \frac{d\Pi^T(\mu_t)}{d\mu_t} = (p^H - p^L)(q^m - q^l)(1 - a^*(\mu_t))Y > 0$$

and thus to ensure that we have a cutoff we simply need to assume that

$$\begin{aligned} \Pi^T(0) &> \Pi^M(0, a^*(0)) \\ \Pi^T(1) &< \Pi^M(1, a^*(1)) \end{aligned}$$

From this it follows that, as before, there exists a unique $\bar{\mu}^a$ for which contracts which fully discourage manipulation are offered when the belief is below this cutoff. In as much, the learning dynamics of the model with endogenous auditing will be qualitatively similar to the model where a was constant. When the belief is below $\bar{\mu}^a$ owners will write contracts which fully discourage

manipulation and hence all of the manager's private information will be released without delay. When the belief rises above $\bar{\mu}^a$ then the manager will be given high powered incentives which encourage them to distort their signal if they are able. However unlike the case where a was fixed the informativeness of manager's signals will now continue to decline with the belief. This is because the equilibrium choice of a^* is strictly decreasing in the belief. As the owners of the firm become more confident that the state is high they select lower levels of auditing because they are more confident that the manager will not find their firm performing poorly and hence attempt to undertake costly manipulation. When the belief is above $\bar{\mu}^a$ the update that is made to the belief following the manager releasing the high signal is

$$\lambda_t^e = \lambda_t + \lambda^{s^H M}(\mu_t)$$

where
$$\lambda^{s^H M}(\mu_t) \equiv \ln \left(\frac{p^H + (1-p^H)(1-a^*(\mu_t))}{p^L + (1-p^L)(1-a^*(\mu_t))} \right)$$

and observe that the magnitude of this updated will fall as the belief rises, that is

$$\frac{d\lambda^{s^H M}(\mu_t)}{d\mu_t} = \frac{\frac{da^*(\mu_t)}{d\mu_t} (p^H - p^L)}{(p^H + (1-p^H)(1-a^*(\mu_t)))(p^L + (1-p^L)(1-a^*(\mu_t)))} < 0$$

because $\frac{da^*(\mu_t)}{d\mu_t} < 0$. When the belief rises, the release of information will become increasingly delayed, strengthening the learning dynamics studied for the case where a was constant.

1.6 Related Literature

This paper is related to existing literatures on both learning and managerial manipulation. Turning first to the learning literature, there are several papers that explain why learning can occur at variable speeds.²² Bikhchandani, Hirshleifer and Welch (1992) and Caplin and Leahy (1993,1994) generate informational cascades in an environment where agents reveal their private information through discrete actions. Discrete action models exogenously impose the noise with which agents release their private information. In my paper, the accuracy with which managers release their private information is endogenously determined by the type of incentives they are

²²For a comprehensive survey see Chamley (2004).

offered in equilibrium. A second important difference is that discrete action models do not predict asymmetries; both booms and crashes can be dramatic.

This paper is more closely related to number of papers that generate economic asymmetries from changes in the speed of learning. Rajan (1994) presents a model in which agents hide bad performance in good times. The result is that learning is slow when times are good and information is released dramatically when times are bad. The asymmetry in Rajan's model comes from the assumption that it is less damaging to an agent's reputation to admit bad performance when times are bad. In contrast, in my model, the manager's incentives to hide bad performance change endogenously over the cycle.

Papers by Van Nieuwerburgh and Veldkamp (2003) and Veldkamp (2002) also use variations in the speed of learning to generate business cycle asymmetry. In these papers, the amount of information released about the aggregate state is determined by the level of economic activity. When the state changes during high activity, the economy learns quickly, producing a sudden crash. Conversely, booms are gradual because they start with low activity which causes information to be realized slowly. In my paper, the amount of information produced in any period is fixed, and only the timing of its release varies. Further work could combine these effects, allowing both the amount and the timing of information to vary endogenously over the cycle.

Several other mechanisms have also been shown to generate learning asymmetries. In Chalkley and Lee (1998) risk aversion leads agents to act slowly on good news but to act promptly on bad news. The result is that good news is revealed more slowly than bad news. Hong and Stein (2002) generate a similar result in an environment where agents have heterogeneous beliefs and short sale constraints.

The literature on managerial manipulation is smaller. In parallel research, Povel, Singh, and Winton (2003) provide an alternative theory as to why fraud occurs in good times. They consider a credit market where the degree of adverse selection varies with the state. Investors can undertake costly monitoring to determine which firms are of high type. In good times, most firms are of the high type and hence investors choose not to monitor. Conversely, monitoring is used in bad times because this is when adverse selection is the most severe. The result is that low types can pose as high types and fraudulently raise funds for their firm, only in good times.

There are a number of papers in which managers take value destroying actions in order

to improve their short-term results. Stein (1989) shows that this happens in an environment where the manager is given incentives sensitive to their firm's current stock price. In that paper the manager's incentives are set exogenously. Bolton, Scheinkman, and Xiong (2003) consider a model in which owners use short-term incentives for the express purpose of inducing the manager to undertake short-term "glamor projects". This is shown to be optimal in an environment where glamor projects can generate speculative rents.

1.7 Conclusion

The primary motivation for this paper is to explain some stylized facts about the US economy over the past decade. In particular, I presented a model which generates gradual booms and rapid recessions of the type observed during this period. In the model, in good times owners offer managers incentives which are highly sensitive to the short-term performance of the firm. This induces managers who are performing poorly to inflate the perceived short-term performance of their firm. While this manipulation does not fool the market, it cannot be unravelled. Thus, when managers manipulate, it slows down the release of their private information about the aggregate state of the economy. Accordingly, booms are gradual because learning is pushed off into the future. Recessions are rapid because this is when the process is reversed. I also showed that the stock market is most vulnerable to news about the state of the economy after a prolonged boom.

The central idea in this paper can, with some qualifications, be applied beyond the 1990s. Several other booms have also given rise to episodes of corporate fraud. Examples include the 1920s (Galbraith (1972) and Baskin and Miranti (1997)), the "go-go" years of the 1960s (Brooks(1973)), and, the leveraged buy-out boom of the 1980s (Kaplan and Stein (1993)). In each instance, the fraud that was undertaken during the boom was discovered in the following downturn. The asymmetry of the economy in the 1990s is not unique. A long held view of the business cycle is that recessions are steeper than booms (see Mitchell (1927), Keynes (1936), and, Burns and Mitchell (1946)). Empirical evidence of this asymmetry has been provided for output (Potter (1995)) and unemployment (Neftci (1984) and Sichel(1993))²³. Others have

²³These studies focus on the US economy. For international evidence see Razzak (2001).

shown that the stock market shows a similar asymmetry; downward movements tend to be steeper than upward movements (see for example Chen, Hong, and Stein (2001)).

This paper has focused upon the role of managerial compensation in determining the timing of information release. I do not suggest that this particular mechanism was in operation during these other episodes. However, the fundamental contracting trade-off I considered here can be applied well beyond the question of managerial incentives. Any contract between a principal and an agent can deter or encourage fraud. If the contract pays the agent based on the ultimate outcome of their relationship, fraud will be deterred. While long-term contracts can always deter fraud, they can come at a cost. In my model, the cost is that long-term incentives are less effective at inducing effort from the agent. In other contexts long-term contracts may, for example, come at the cost that they require a suboptimal allocation of risk. Taken in this broader context, the model in this paper can explain why booms have tended to be times of fraud. Moreover, the paper shows that the learning dynamics arising from the principal-agent problem produce gradual booms and rapid recessions.

1.8 Appendix

1.8.1 Proof of Proposition 1 and Proposition 10

I solve the problem of the owner without assuming additivity. Define

$$\begin{aligned} p_t &\equiv \mu_t p(e_t = 1, \theta_t = \theta^H) + (1 - \mu_t) p(e_t = 1, \theta_t = \theta^L) \\ \sigma_t &\equiv \mu_t p(e_t = 0, \theta_t = \theta^H) + (1 - \mu_t) p(e_t = 0, \theta_t = \theta^L) \end{aligned}$$

The optimal contract which induces $M_t = 1, e_t = 1$ solves the following problem:

$$\begin{aligned} \min & [p_t q^h + (1 - p_t)(1 - a)q^l] W^{s^{h'Y}} + [p_t(1 - q^h) + (1 - p_t)(1 - a)(1 - q^l)] W^{s^{h'0}} \\ & + (1 - p_t)a [q^m W^{s^{l'Y}} + (1 - q^m) W^{s^{l'0}}] \end{aligned} \quad (1.10)$$

subject to

$$[q^h - (1 - a)q^l] W^{s^{h'Y}} + [(1 - q^h) - (1 - a)(1 - q^l)] W^{s^{h'0}} - a [q^m W^{s^{l'Y}} + (1 - q^m) W^{s^{l'0}}] \geq \frac{B}{p_t - \sigma_t} \quad (1.11)$$

$$q^l W^{s^{h'Y}} + (1 - q^l) W^{s^{h'0}} \geq q^m W^{s^{l'Y}} + (1 - q^m) W^{s^{l'0}}, \quad (1.12)$$

$$W^{s^{h'Y}}, W^{s^{h'0}}, W^{s^{l'Y}}, W^{s^{l'0}} \geq 0. \quad (1.13)$$

In words, the contract minimizes the expected payment to the manager (1.10) subject to the manager's incentive compatibility condition (1.11), the condition which ensures that the manager does in fact choose to manipulate (1.12), and the limited liability restrictions (1.13). To demonstrate that (1.2) is in fact optimal we can use the method of lagrange. In particular we can verify this solution if we can find a set of positive multipliers which support it. For now let's ignore (1.12) and we can verify later that it will be satisfied strictly by the solution. Letting $\psi, \psi^{s^h Y}, \psi^{s^h 0}, \psi^{s^l Y}, \psi^{s^l 0} \geq 0$ be the appropriately defined lagrange multipliers associated with each constraint we can write the Lagrangian as:

$$\begin{aligned} \min L = & [p_t q^h + (1 - p_t)(1 - a)q^l] W^{s^h Y} + [p_t(1 - q^h) + (1 - p_t)(1 - a)(1 - q^l)] W^{s^h 0} \\ & + (1 - p_t)a [q^m W^{s^l Y} + (1 - q^m) W^{s^l 0}] + \psi \left\{ \frac{B}{p_t - \sigma_t} - [q^h - (1 - a)q^l] W^{s^h Y} \right. \\ & - [(1 - q^h) - (1 - a)(1 - q^l)] W^{s^h 0} + a [q^m W^{s^l Y} + (1 - q^m) W^{s^l 0}] \left. \right\} \\ & - \psi^{s^h Y} W^{s^h Y} - \psi^{s^h 0} W^{s^h 0} - \psi^{s^l Y} W^{s^l Y} - \psi^{s^l 0} W^{s^l 0} \end{aligned}$$

The first order conditions for this problem are:

$$W^{s^h Y} : [p_t q^h + (1 - p_t)(1 - a)q^l] - \psi [q^h - (1 - a)q^l] - \psi^{s^h Y} = 0 \quad (1.14)$$

$$W^{s^h 0} : [p_t(1 - q^h) + (1 - p_t)(1 - a)(1 - q^l)] - \psi [(1 - q^h) - (1 - a)(1 - q^l)] - \psi^{s^h 0} = 0 \quad (1.15)$$

$$W^{s^l Y} : (1 - p_t)a q^m + \psi a q^m - \psi^{s^l Y} = 0 \quad (1.16)$$

$$W^{s^l 0} : (1 - p_t)a(1 - q^m) + \psi a(1 - q^m) - \psi^{s^l 0} = 0 \quad (1.17)$$

The associated complimentary slackness conditions are:

$$\begin{aligned} \psi \left\{ \frac{B}{p_t - \sigma_t} - [q^h - (1 - a)q^l] W^{s^h Y} - [(1 - q^h) - (1 - a)(1 - q^l)] W^{s^h 0} \right. \\ \left. + a [q^m W^{s^l Y} + (1 - q^m) W^{s^l 0}] \right\} = 0 \end{aligned} \quad (1.18)$$

$$\psi^{s^h Y} W^{s^h Y} = \psi^{s^h 0} W^{s^h 0} = \psi^{s^l Y} W^{s^l Y} = \psi^{s^l 0} W^{s^l 0} = 0 \quad (1.19)$$

Observe that (1.16) and (1.17) ensure that $\psi^{s^l Y}, \psi^{s^l 0} > 0$ and so when combined with (1.18) it must be that $W^{s^l Y} = W^{s^l 0} = 0$. Next suppose that as conjectured $W^{s^h Y} > 0$. This implies that $\psi^{s^h Y} = 0$ and hence that

$$\psi = \frac{p_t q^h + (1 - p_t)(1 - a)q^l}{q^h - (1 - a)q^l} > 0$$

Finally observe that this implies that

$$\psi^{s^{h'0}} = \left(\frac{(1-a)(q^h - q^l)}{q^h - (1-a)q^l} \right) > 0$$

which verifies that the solution does indeed have $W^{s^{h'Y}} > W^{s^{h'0}} = 0$. With this we need only find the minimum value of $W^{s^{h'Y}}$ which satisfies (1.11) when $W^{s^{l'Y}} = W^{s^{l'0}} = W^{s^{h'0}} = 0$. This is given in (1.2). To find the expected cost of these incentives this is simply substituted into (1.10). Subtracting this from the expected output of a firm in which a manager is going to manipulate gives (1.3).

Next to solve for the optimal contract which discourages misreporting. This problem is:

$$\min p_t \left[q^h W^{s^{h'Y}} + (1 - q^h) W^{s^{h'0}} \right] + (1 - p_t) \left[q^m W^{s^{l'Y}} + (1 - q^m) W^{s^{l'0}} \right] \quad (1.20)$$

subject to

$$\left[q^h W^{s^{h'Y}} + (1 - q^h) W^{s^{h'0}} \right] - \left[q^m W^{s^{l'Y}} + (1 - q^m) W^{s^{l'0}} \right] \geq \frac{B}{p_t - \sigma_t}; \quad (1.21)$$

$$q^m W^{s^{l'Y}} + (1 - q^m) W^{s^{l'0}} \geq q^l W^{s^{h'Y}} + (1 - q^l) W^{s^{h'0}}; \quad (1.22)$$

$$W^{s^{h'Y}}, W^{s^{h'0}}, W^{s^{l'Y}}, W^{s^{l'0}} \geq 0. \quad (1.23)$$

This contract minimizes the expected payment to the manager (1.20) subject to the incentive compatibility constraint (1.21), the condition which ensures that the manager does not elect to misreport when they are able to do so (1.22), and the limited liability conditions (1.23). Let $\gamma, \varphi, \psi^{s^{h'Y}}, \psi^{s^{h'0}}, \psi^{s^{l'Y}}, \psi^{s^{l'0}} \geq 0$ be the lagrange multiplier associated with each of these constraints. The Lagrangian is then:

$$\begin{aligned} \min L = & p_t \left[q^h W^{s^{h'Y}} + (1 - q^h) W^{s^{h'0}} \right] + (1 - p_t) \left[q^m W^{s^{l'Y}} + (1 - q^m) W^{s^{l'0}} \right] \\ & + \gamma \left[\frac{B}{p_t - \sigma_t} - \left[q^h W^{s^{h'Y}} + (1 - q^h) W^{s^{h'0}} \right] + \left[q^m W^{s^{l'Y}} + (1 - q^m) W^{s^{l'0}} \right] \right] \\ & + \varphi \left[q^l W^{s^{h'Y}} + (1 - q^l) W^{s^{h'0}} - q^m W^{s^{l'Y}} - (1 - q^m) W^{s^{l'0}} \right] \\ & - \psi^{s^{h'Y}} W^{s^{h'Y}} - \psi^{s^{h'0}} W^{s^{h'0}} - \psi^{s^{l'Y}} W^{s^{l'Y}} - \psi^{s^{l'0}} W^{s^{l'0}} \end{aligned}$$

The first order conditions for this problem are:

$$W^{s^{h'Y}} : p_t q^h - \gamma q^h + \varphi q^l - \psi^{s^{h'Y}} = 0 \quad (1.24)$$

$$W^{s^{h'0}} : p_t (1 - q^h) - \gamma (1 - q^h) + \varphi (1 - q^l) - \psi^{s^{h'0}} = 0 \quad (1.25)$$

$$W^{s^{l'Y}} : (1 - p_t) q^m + \gamma q^m - \varphi q^m - \psi^{s^{l'Y}} = 0 \quad (1.26)$$

$$W^{s^{l'0}} : (1 - p_t) (1 - q^m) + \gamma (1 - q^m) - \varphi (1 - q^m) - \psi^{s^{l'0}} = 0 \quad (1.27)$$

The associated complimentary slackness conditions are:

$$\gamma \left[\frac{B}{p_t - \sigma_t} - \left[q^h W^{s^h Y} + (1 - q^h) W^{s^h 0} \right] + \left[q^m W^{s^m Y} + (1 - q^m) W^{s^m 0} \right] \right] = 0 \quad (1.28)$$

$$\varphi \left[q^l W^{s^l Y} + (1 - q^l) W^{s^l 0} - q^m W^{s^m Y} - (1 - q^m) W^{s^m 0} \right] = 0 \quad (1.29)$$

$$\psi^{s^h Y} W^{s^h Y} = \psi^{s^h 0} W^{s^h 0} = \psi^{s^m Y} W^{s^m Y} = \psi^{s^m 0} W^{s^m 0} = 0 \quad (1.30)$$

To show that having $W^{s^h Y}, W^{s^m Y}, W^{s^m 0} > 0$ is indeed optimal note first that by (1.30) this implies that

$$\psi^{s^h Y} = \psi^{s^m Y} = \psi^{s^m 0} = 0.$$

Substituting this into the first order conditions yields

$$\begin{aligned} \gamma &= p_t + \frac{q^l}{q^h - q^l} > 0, \\ \varphi &= 1 + \frac{q^l}{q^h - q^l} > 0, \\ \psi^{s^h 0} &= 1. \end{aligned}$$

Having found positive values for the other multipliers we can conclude that $W^{s^h Y}, W^{s^m Y}, W^{s^m 0} > 0$ and $W^{s^h 0} = 0$ is indeed optimal. Moreover (1.28) and (1.29) combined with $\gamma, \varphi > 0$ imply that both constraints must bind. There is an ambiguity however in solving for $W^{s^h Y}, W^{s^m Y}$, and $W^{s^m 0}$ because there are only two constraints to satisfy. In particular for (1.21) and (1.22) to bind we must have

$$\begin{aligned} \left[q^m W^{s^m Y} + (1 - q^m) W^{s^m 0} \right] &= q^h W^{s^h Y} - \frac{B}{p_t - \sigma_t} \\ \left[q^m W^{s^m Y} + (1 - q^m) W^{s^m 0} \right] &= q^l W^{s^h Y} \end{aligned}$$

From here it is clear that $W^{s^h Y}$ is uniquely determined as

$$W^{s^h Y} = \frac{B}{(p_t - \sigma_t)(q^h - q^l)}$$

and there is a locus of points continuum of $W^{s^m Y}, W^{s^m 0}$ values such that

$$\left[q^m W^{s^m Y} + (1 - q^m) W^{s^m 0} \right] = \frac{q^l B}{(p_t - \sigma_t)(q^h - q^l)}$$

which are in fact optimal. To select one of these solutions let $W^{s''Y} = W^{s''0} = W^{s''}$ which gives

$$W^{s''} = \frac{q^l B}{(p_t - \sigma_t)(q^h - q^l)}$$

Substituting this solution into the objective function gives the expected profit to the owners from a contract which discourages misreporting.

Next we need to show that the optimal contract does in fact induce high effort. It is easy to see that the optimal contract which does not induce high effort will simply set

$$w(s'_h, Y) w(s'_h, 0) = w(s'_l, Y) = w(s'_l, 0) = 0.$$

Given such a contract the manager will not misreport. The resulting profits to the owners will be

$$\Pi^l(\mu_t) = [(p_t - \eta)q^h + (1 - p_t + \eta)q^m] Y$$

Such a contract is never optimal if and only if

$$\Pi^l(\mu_t) \leq \max\{\Pi^M(\mu_t), \Pi^T(\mu_t)\}.$$

Writing these expressions out and rearranging terms shows that this holds whenever η satisfies

$$\eta \geq \max \left\{ \sqrt{\frac{B}{Y(q^h - q^m)} \left(p_t + \frac{q^l}{q^h - q^l} \right)}, \frac{(1 - p_t)(1 - a)(q^m - q^l)Y}{2Y(q^h - q^m)} + \left(2Y(q^h - q^m) \right)^{-1} \cdot \left[\sqrt{[(1 - p_t)(1 - a)(q^m - q^l)Y]^2 + 4(q^h - q^m)B \left[p_t + \frac{(1 - a)q^l}{q^h - (1 - a)q^l} \right]} \right] \right\}$$

Next we must determine when the optimal contract encourages misreporting. First, observe that

$$\frac{\partial(\Pi^M(\mu_t) - \Pi^T(\mu_t))}{\partial \mu_t} = Y(1 - a)(q^m - q^l)(p^H - p^L) - \frac{aBq^l q^h [(p^H - \sigma^H) - (p^L - \sigma^L)]}{(q^h - q^l)(q^h - (1 - a)q^l)(p_t - \sigma_t)^2}$$

Note that the first term is unambiguously positive. If $(p^L - \sigma^L) \geq (p^H - \sigma^H)$ then the second term is also positive and it must be that $\frac{\partial(\Pi^M(\mu_t) - \Pi^T(\mu_t))}{\partial \mu_t} > 0$. If however $(p^H - \sigma^H) > (p^L - \sigma^L)$ then to

have $\frac{\partial(\Pi^M(\mu_t) - \Pi^T(\mu_t))}{\partial\mu_t} > 0$ requires that

$$(p_t - \sigma_t)^2 > \frac{aBq^l q^h [(p^H - \sigma^H) - (p^L - \sigma^L)]}{(q^h - q^l)(q^h - (1-a)q^l)Y(1-a)(q^m - q^l)(p^H - p^L)}$$

Since under this scenario we have that $(p^H - \sigma^H) > (p^L - \sigma^L)$ then it follows that $(p_t - \sigma_t)^2 \geq (p^L - \sigma^L)^2$ where the equality is strict only when $\mu_t = 0$. Therefore to ensure this condition holds $\forall \mu_t \in [0, 1]$ we must assume that it holds when $(p_t - \sigma_t) = (p^L - \sigma^L)$. This requires that

$$(p^L - \sigma^L)^2 > \frac{aBq^l q^h [(p^H - \sigma^H) - (p^L - \sigma^L)]}{(q^h - q^l)(q^h - (1-a)q^l)Y(1-a)(q^m - q^l)(p^H - p^L)}$$

but then observe that this is also satisfied whenever $(p^L - \sigma^L) \geq (p^H - \sigma^H)$ and hence is by itself a necessary and sufficient condition. This proves Proposition 11. In the additive case where $(p^L - \sigma^L) = (p^H - \sigma^H) = \eta$ that

$$\frac{\partial(\Pi^M(\mu_t) - \Pi^T(\mu_t))}{\partial\mu_t} = Y(1-a)(q^m - q^l)(p^H - p^L) > 0$$

In this case $\Pi^M(\mu_t) - \Pi^T(\mu_t)$ is linear and hence there must exist a unique $\bar{\mu}$ at which $\Pi^M(\bar{\mu}) = \Pi^T(\bar{\mu})$. With $\Pi^M(\mu_t) \geq \Pi^T(\mu_t)$ if and only if $\mu_t \geq \bar{\mu}$. This establishes Proposition 1 and Proposition 10. ■

1.8.2 Proof of Proposition 4

We can rank these phases by showing that the information can be cast as a garbling of the information in another. To see that *Intense* \geq *Direct* simply observe that intense learning is direct learning plus an additional signal (i.e., results). Similarly *Delayed* \geq *Slow* by exactly the same logic. *Intense* \geq *Delayed* because delayed learning involves managers garbling their signal whereas this is ungarbled in intense learning. The same logic establishes that *Direct* \geq *Slow*. This establishes Proposition 4. ■

1.8.3 Proof of Proposition 7

To prove $E(\Delta_t | \theta_t = \theta^L, \theta_{t-T} = \theta^H, \lambda_{t-T} > \bar{\lambda}, \lambda_t \geq \bar{\lambda}) \geq 0$. First observe that $\Phi^M(\theta_t)|_{a=0} = 0$ and $\Phi^M(\theta_t)|_{a=1} = (2p^{\theta_t} - 1)$. Next, we can show that $\frac{\partial\Phi^M(\theta_t)}{\partial a} > 0$ if $\theta_t = \theta^H$ and $\frac{\partial\Phi^M(\theta_t)}{\partial a} < 0$ if

$\theta_t = \theta^L$. This simply comes from the fact that decreasing a garbles the managers signal further. Similarly observe that $\Omega(\theta_{t-T})|_{a=1} = 0$. Next we can show that $\frac{\partial \Omega(\theta_{t-T})}{\partial a} < 0$ if $\theta_{t-T} = \theta^H$ and $\frac{\partial \Omega(\theta_{t-T})}{\partial a} > 0$ if $\theta_{t-T} = \theta^L$. This comes from the fact that output is in expectation less informative when managers have less chance to manipulate. Thus if $\theta_t = \theta^L, \theta_{t-T} = \theta^H$ then $[\Phi^M(\theta^L) + \Omega(\theta^H)]|_{a=0} > 0$ and $[\Phi^M(\theta^L) + \Omega(\theta^H)]|_{a=1} < 0$. It also follows that $\frac{\partial}{\partial a} [\Phi^M(\theta^L) + \Omega(\theta^H)] < 0$ and so there exists a unique \hat{a} for which $[\Phi^M(\theta^L) + \Omega(\theta^H)]|_{a=\hat{a}} = 0$. If $a < \hat{a}$ then then $\Phi^M(\theta^L) + \Omega(\theta^H) > 0$ and if $a > \hat{a}$ then $\Phi^M(\theta^L) + \Omega(\theta^H) < 0$. A symmetric argument applies for the case where $\theta_t = \theta^L, \theta_{t-T} = \theta^H$. This establishes Proposition 7. ■

1.8.4 Proof of Proposition 12

Define $\mu_t^{\theta^j} = \Pr(\theta_t = \theta^j | H_t)$. The optimal contract that induces $e_t = M_t = 1$ solves

$$\begin{aligned} \min_{\theta} \sum_{\theta} \mu_t^{\theta} \sum_{i=1}^{N_t} \alpha_{\theta}^i & \left[[p^{\theta} q^h + (1-p^{\theta})(1-a)q^l] w(s'_h, Y, x_i) \right. \\ & + [p^{\theta}(1-q^h) + (1-p^{\theta})(1-a)(1-q^l)] w(s'_h, 0, x_i) \\ & \left. + (1-p^{\theta}) a q^m w(s'_l, Y, x_i) + (1-p^{\theta}) a (1-q^m) w(s'_l, 0, x_i) \right] \end{aligned}$$

subject to:

$$\begin{aligned} \sum_{\theta} \mu_t^{\theta} \sum_{i=1}^{N_t} \alpha_{\theta}^i & \left[[q^h - (1-a)q^l] w(s'_h, Y, x_i) \right. \\ & + [(1-q^h) - (1-a)(1-q^l)] w(s'_h, 0, x_i) \\ & \left. - a q^m w(s'_l, Y, x_i) - a(1-q^m) w(s'_l, 0, x_i) \right] \geq \frac{B}{\eta} \end{aligned} \tag{1.32}$$

and limited liability $w(s'_h, Y, x_i), w(s'_h, 0, x_i), w(s'_l, Y, x_i), w(s'_l, 0, x_i) \geq 0$. Note that (1.32) ensures that the manager prefers $e_t = M_t = 1$ to $e_t = 0, M_t = 1$ ²⁴. For brevity I sketch the solution of this problem. It must be optimal to set $w(s'_l, Y, x_i) = w(s'_l, 0, x_i) = 0 \forall i$ since these unambiguously exacerbate (1.32). Next write the problem as a Lagrangian. Let $\phi_j^y \geq 0$ be the

²⁴I have ignored two additional constraints which ensure the manager does not select. The first ensures that the manager does not select $e_t = 1, M_t = 0$ or $e_t = M_t = 0$. It can be verified that the solution deters these choices.

multiplier on $w(s'_h, y, x_j)$ and $\gamma \geq 0$ be the multiplier on (1.32). The first order conditions for each $w(s'_h, y, x_j)$ give

$$\sum_{\theta} \mu_t^{\theta} \alpha_{\theta}^i \left[p^{\theta} q^h + (1 - p^{\theta}) (1 - a) q^l - \gamma \left[q^h - (1 - a) q^l \right] \right] = \phi_j^Y \quad (1.33)$$

$$\begin{aligned} \sum_{\theta} \mu_t^{\theta} \alpha_{\theta}^i \left[p^{\theta} (1 - q^h) + (1 - p^{\theta}) (1 - a) (1 - q^l) \right. \\ \left. - \gamma \left[(1 - q^h) - (1 - a) (1 - q^l) \right] \right] = \phi_j^Y \end{aligned} \quad (1.34)$$

Setting $\phi_j^Y = 0$ in 1.34 and substituting into 1.33 implies that $\phi_j^Y < 0$. Hence by complimentary slackness $w(s'_h, 0, x_j) = 0 \forall j$. We must have at least one $w(s'_h, Y, x_j) > 0$ to satisfy (1.32). If $w(s'_h, Y, x_j) > 0$ then this must be consistent with having $\phi_k^Y \geq 0$ for all $k \neq j, k \in \{1, 2, \dots, N_t\}$. Rearranging the first order conditions this holds if and only if $\frac{\alpha_{\theta}^k}{\alpha_{\theta}^l} \geq \frac{\alpha_{\theta}^j}{\alpha_{\theta}^l}$. Since this must hold for all $k \neq j, k \in \{1, 2, \dots, N_t\}$ then it follows that $w(s'_h, Y, x_1) > 0$ and $w(s'_h, Y, x_i) = 0 \forall i \geq 2$. From here the solution is obtained by finding the lowest $w(s'_h, Y, x_1)$ which satisfies (1.32). This establishes Proposition 12. ■

1.8.5 Proof of Proposition 13

The optimal contract which induces $e_t = 1, M_t = 0$ solves

$$\begin{aligned} \min \sum_{\theta} \mu_t^{\theta} \sum_{i=1}^{N_t} \alpha_{\theta}^i \left[p^{\theta} q^h w(s'_h, Y, x_i) + p^{\theta} (1 - q^h) w(s'_h, 0, x_i) \right. \\ \left. + (1 - p^{\theta}) q^m w(s'_h, Y, x_i) + (1 - p^{\theta}) (1 - q^m) w(s'_h, 0, x_i) \right] \end{aligned} \quad (1.35)$$

subject to

$$\begin{aligned} \sum_{\theta} \mu_t^{\theta} \sum_{i=1}^{N_t} \alpha_{\theta}^i \left[(1 - p^{\theta}) q^m w(s'_h, Y, x_i) + (1 - p^{\theta}) (1 - q^m) w(s'_h, 0, x_i) \right] \geq \quad (1.36) \\ \sum_{\theta} \mu_t^{\theta} \sum_{i=1}^{N_t} \alpha_{\theta}^i \left[(1 - p^{\theta}) q^l w(s'_h, Y, x_i) + (1 - p^{\theta}) (1 - q^l) w(s'_h, 0, x_i) \right] \end{aligned}$$

$$\begin{aligned} & \sum_{\theta} \mu_{\theta}^{\theta} \sum_{i=1}^{N_t} \alpha_{\theta}^i \left[q^h w(s'_h, Y, x_i) + (1 - q^h) w(s'_h, 0, x_i) \right. \\ & \left. - q^m w(s'_i, Y, x_i) - (1 - q^m) w(s'_i, 0, x_i) \right] \geq \frac{B}{\eta} \end{aligned} \quad (1.37)$$

$$\begin{aligned} & \sum_{\theta} \mu_{\theta}^{\theta} \sum_{i=1}^{N_t} \alpha_{\theta}^i \left[p^{\theta} q^h w(s'_h, Y, x_i) + p^{\theta} (1 - q^h) w(s'_h, 0, x_i) \right. \\ & \left. + (1 - p^{\theta}) q^m w(s'_i, Y, x_i) + (1 - p^{\theta}) (1 - q^m) w(s'_i, 0, x_i) \right] \geq \\ & B + \sum_{\theta} \mu_{\theta}^{\theta} \sum_{i=1}^{N_t} \alpha_{\theta}^i \left[\left[(p^{\theta} - \eta) q^h + (1 - a) (1 - p^{\theta} + \eta) q^l \right] w(s'_h, Y, x_i) \right. \\ & \left. + \left[(p^{\theta} - \eta) (1 - q^h) + (1 - a) (1 - p^{\theta} + \eta) (1 - q^l) \right] w(s'_h, 0, x_i) \right. \\ & \left. + (1 - p^{\theta} + \eta) q^m a w(s'_i, Y, x_i) + (1 - p^{\theta} + \eta) (1 - q^m) a w(s'_i, 0, x_i) \right] \end{aligned} \quad (1.38)$$

and limited liability. Note that (1.36) deters $M_t = 1, e_t = 1$, (1.37) deters $M_t = 0, e_t = 0$, and (1.38) deters $M_t = 1, e_t = 0$. I sketch the solution to this problem. As per the analysis without contracting on aggregate information, we can show $w(s'_h, 0, x_i) = 0 \forall i$. Also, we can set $w(s'_i, Y, x_i) = w(s'_i, 0, x_i) = w(s'_i, x_i)$. In order to satisfy (1.37) we must have $w(s'_h, Y, x_i) > 0$ for some i and thus in order to satisfy (1.36) we must have $w(s'_i, x_j) > 0$ for some j . Writing the Lagrangian we can show that $w(s'_i, x_1) > 0$ and $w(s'_i, x_j) = 0 \forall j \geq 2$. Finally we must determine the i for which $w(s'_h, Y, x_i) > 0$. Positing $w(s'_h, Y, x_1) > 0$ as the solution, we can find positive values for each Lagrange multiplier if

$$\sum_{\theta} \mu_{\theta}^{\theta} \alpha_{\theta}^1 \left[\eta a - (1 - p^{\theta}) (1 - a) \right] \geq 0 \quad (1.39)$$

$$\frac{C_1}{C_2} \left[q^h \sum_{\theta} \mu_{\theta}^{\theta} \alpha_{\theta}^{N_t} p^{\theta} + q^l \sum_{\theta} \mu_{\theta}^{\theta} \alpha_{\theta}^{N_t} (1 - p^{\theta}) \right] - 1 \geq 0 \quad (1.40)$$

$$\begin{aligned}
C_1 &\equiv \frac{1}{\eta} \left(\frac{q^l}{q^h - q^l} + \frac{p^H \mu_t \alpha_H^1 + p^L (1 - \mu_t) \alpha_L^1}{\mu_t \alpha_H^1 + (1 - \mu_t) \alpha_L^1} \right) \\
C_2 &\equiv \sum_{\theta} \mu_t^{\theta} \alpha_{\theta}^{N_t} \left[\eta q^h - (1 - a) (1 - p^{\theta} + \eta) q^l \right] \\
&\quad + \frac{q^l \sum_{\theta} \mu_t^{\theta} \alpha_{\theta}^{N_t} (1 - p^{\theta})}{q^l \sum_{\theta} \mu_t^{\theta} \alpha_{\theta}^1 (1 - p^{\theta})} \sum_{\theta} \mu_t^{\theta} \alpha_{\theta}^1 \left[\eta a - (1 - p^{\theta}) (1 - a) \right]
\end{aligned}$$

It can be shown that these conditions are satisfied when q^l is small or when a is sufficiently large.

Bibliography

- [1] Baker, George P. (1992), "Incentive Contracts and Performance Measurement", *Journal of Political Economy*, 100(3), 598-614.
- [2] Baskin, Jonathan and Paul Miranti (1997), *A History of Corporate Finance*, Cambridge University Press.
- [3] Bergstresser, Daniel and Thomas Phillipon (2003), "CEO Incentives and Earnings Management: Evidence From the 1990s", mimeo.
- [4] Bikhchandani, Sushil, David Hirshleifer and Ivo Welch (1992), "A Theory of Fads, Fashion, Custom, and Cultural Change as Information Cascades", *Journal of Political Economy*, 100(5), 992-1026.
- [5] Blackwell, David (1951), "Comparison of Experiments", In J. Neymann (ed.) *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press.
- [6] Bolton, Patrick and Jose Scheinkman, and Wei Xiong (2003) "Executive Compensation and Short-termist Behavior in Speculative Markets", NBER Working Paper No. w9722.
- [7] Brooks, John (1973) *The Go-Go Years*, New York: Weybright and Talley.
- [8] Burns, Arthur and Wesley Mitchell (1946), *Measuring Business Cycles*, New York: Columbia University Press.
- [9] Caplin, Andrew and John Leahy (1993), "Sectoral Shocks, Learning, and Aggregate Fluctuations", *Review of Economic Studies*, 60, 777-794.

- [10] Caplin, Andrew and John Leahy (1994), "Business as Usual, market Crashes, and Wisdom After the Fact", *American Economic Review*, 84(3), 548-565.
- [11] Chalkley, Martin and In Ho Lee (1998) "Learning and Asymmetric Business Cycles", *Review of Economic Dynamics*, 623-645.
- [12] Chamley, Christophe (2004), *Rational Herds*, Cambridge University Press.
- [13] Chen, Joseph, Harrison Hong and Jeremy Stein (2001), "Forecasting Crashes: Trading Volume, past Returns, and Conditional Skewness in Stock Prices", *Journal of Financial Economics*, 61, 345-381.
- [14] Desai, Mihir (2002), "The Corporate Profit Base, Tax Sheltering Activity, and the Changing Nature of Employee Compensation", NBER Working Paper No. 8866.
- [15] Fudenberg, Drew and Jean Tirole (1986), "A 'Signal-Jamming' Theory of Predation", *RAND Journal of Economics*, 39, 366-376.
- [16] Galbraith, John K. (1972), *The Great Crash: 1929*, Boston: Houghton Mifflin Company.
- [17] Graham, Carol, Robert Litan and Sandip Sukhtankar (2002), "The Bigger They Are, The Harder They Fall: An Estimate of the Costs of the Crisis in Corporate Governance", Brookings Working Paper.
- [18] Hall, Brian J. and Kevin J. Murphy (2003), "The Trouble with Stock Options", *Journal of Economic Perspectives*, 17(3), 49-70.
- [19] Holmstrom, Bengt (1999) "Managerial Incentive Problems: A Dynamic Perspective", *Review of Economic Studies*, 66(1), 169-82.
- [20] Hong, Harrison and Jeremy Stein (2002) "Differences of Opinion, Short-Sales Constraints and Market Crashes", Mimeo.
- [21] Holmstrom, Bengt and Paul Milgrom (1991), "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design", *Journal of Law, Economics, and Organization*, 7, 24-52.

- [22] Kaplan, Steven and Jeremy Stein (1993), "The Evolution of Buyout Pricing and Financial Structure in the 1980s", *Quarterly Journal of Economics*, 108(2), 313-357.
- [23] Kedia, Simi (2003), "Do Executive Stock Options Generate Incentives for Earnings Management? Evidence from Accounting Restatements", mimeo.
- [24] Keynes, John M. (1936) *The General Theory of Employment, Interest and Money*, London: Macmillan.
- [25] Laffont, Jean-Jacques, (1989), *The Economics of Uncertainty and Information*, MIT Press.
- [26] Lev, Baruch (2003) "Corporate Earnings: Facts and Fiction", *Journal of Economic Perspectives*, 17(2), 27-50.
- [27] Matsumoto, Dawn (2000) "Management's Incentives to Guide Analysts' Forecasts", Working Paper, Harvard University.
- [28] Mitchell, Wesley (1927), *Business Cycles: The Problem and Its Setting*, New York: National Bureau of Economic Research.
- [29] Neftci, Salih N. (1984) "Are Economic Time Series Asymmetric Over the Business Cycle?", *Journal of Political Economy*, 92, 307-328.
- [30] Potter, Simon (1995), "A Nonlinear Approach to US GNP", *Journal of Applied Econometrics*, 10, 109-125.
- [31] Povel, Paul, Rajdeep Singh and Andrew Winton (2003) "Booms, Busts. and Fraud", mimeo.
- [32] Rajan, Raghuram G. (1994), "Why bank Credit Policies Fluctuate: A Theory and Some Evidence", *Quarterly Journal of Economics*, 109(2), 399-441.
- [33] Razzak, W. (2001), "Business Cycle Asymmetries: International Evidence", *Review of Economic Dynamics*, 4, 230-243.
- [34] Sichel, Daniel (1993), "Business Cycle Asymmetry: A Deeper Look", *Economic Inquiry*, 31(2), 224-236.

- [35] Stein, Jeremy (1989), "Efficient Capital Markets, Inefficient Firms: A Model of Myopic Corporate Behavior", *Quarterly Journal of Economics*, 104(4), 655-669.
- [36] Van Nieuwerburgh, Stijn and Laura Veldkamp (2003), "Learning Asymmetries in Real Business Cycles", mimeo.
- [37] Veldkamp, Laura (2002), "Slow Boom, Sudden Crash", mimeo.
- [38] Wu, Min (2002), "Earnings Restatements: A Capital Market Perspective", Ph.D. Thesis, New York University.

Chapter 2

Speculative Trading and Corporate Misreporting

This chapter presents a model of corporate misreporting in an environment where investors have heterogeneous beliefs and short sale constraints. The disagreement between investors provides a motive for agents who start a firm to limit the amount of information which it releases to the public so as to sponsor speculation over its value. This incentive to limit information is stronger when the heterogeneity of beliefs among investors is stronger. Investors also learn about a firm's expected profitability from the information released by other firms in the industry. I show that this creates a strategic complementarity in the precision of information released by each firm. This can give rise to multiple equilibria: one in which all firms release precise reports and one in which their reports are inaccurate.

2.1 Introduction

The aftermath of the 1990s boom has brought questions regarding corporate governance and accounting standards to prominence. The discovery of accounting fraud at companies such as Enron, Kmart, Tyco, Global Crossing, Quest, Worldcom, and, Waste Management has made headline news in the popular press¹. Wu (2002) documents the rise in earning restatements

¹For a more complete list of corporate scandals and dates see Graham, Litan and Suktankar (2002).

during the final years of the 1990s. Practices to distort accounting information were particularly widespread among “New Economy” firms (see for example Kahn (2000), Macdonald (2000) and Davis (2002)). The first question asked in this paper is: why would the owners of a firm allow a firm to issue imprecise information? The information published by a firm is used to evaluate the performance of its manager. The more accurate is this information the better the owners of the firm are able to monitor the manager. A simple model would suggest that the owners of a firm will implement accounting accuracy so that the marginal cost of extra scrutiny is equal to its marginal contribution to improved governance. In this paper I suppose that the owners of the firm can costlessly set any level of precision for the firm's accounts. In this simple model they would always opt for perfectly informative accounts.

This paper considers an environment in which the market value of the firm can potentially have a speculative component. This comes from the fact that investors have heterogeneous prior beliefs about the value of the firm and hence will potentially disagree about its value. This disagreement, combined with short sales constraints, produces a speculative component to the firm's price. This presents owners with an incentive to release imprecise information about the firm. Releasing perfectly informative accounts will fully reveal the value of the firm and thus end any disagreement over its value. Owners may choose to commit to imprecise accounting standards so as to sponsor the speculative component of the firm's price. This comes at the cost of less effective monitoring of the manager. I show that such a policy will only be optimal when the heterogeneity of beliefs among investors is sufficiently large. I also show that the owner of a firm will only choose to pursue this strategy if any external information about the firm's value is sufficiently imprecise. The logic here is similar. If external information about the firm's value is sufficiently precise then there is little for investors to disagree over and so the speculative motive is weak.

Historically, episodes of strong stock market speculation and imprecise accounting have gone hand in hand. Besides the 1990s examples include the 1920s (Galbraith(1972) and Baskin and Mirati (1997)) the “go-go” years of the 1960s (Brooks(1973)), and, the leveraged buy-out boom of the 1980s (Kaplan and Stein(1993)). The model in this paper predicts that speculation and imprecise information are closely linked and that these episodes will occur in waves. To demonstrate this I consider an environment in which there are many identical firms in the same

industry. The information from any one firm is used by investors to evaluate the profitability of each firm in the industry. I show that this generates a strategic complementarity in accounting standards. When each firm in the industry releases perfectly precise information there is no uncertainty about industry profitability and therefore no speculation. Thus each firm has no individual incentive to limit the precision of their accounts. Conversely when every other firm is releasing imperfect information this opens the possibility for speculation and thus encourages each individual firm to follow such a policy. I demonstrate that there is always an equilibrium in which all firms release perfectly informative accounts. In addition, when the heterogeneity of beliefs among investors is strong, there is also an equilibrium in which each firm releases imperfectly informative accounts and enjoys a speculative premium in the firm's price. Thus the theory suggests a reason why episodes of accounting fraud happen in waves - due to the strategic complementarity information. It also shows that this phenomenon is closely linked to episodes of elevated stock prices and speculation.

This paper is most closely related to recent work by Bolton, Scheinkman and Xiong (2003). They also consider an environment in which investors have heterogeneous beliefs and short sale constraints. They argue that owners may offer short term incentives to a manager to induce them to invest in high variance zero expected value projects. The uncertainty of these projects permits disagreement between investors and thus creates a speculative component to the firm's price. The decision to invest in variance projects in their model is analogous to creating uncertainty by releasing imperfect information as per the framework in this paper. The key difference between their paper and mine is that I demonstrate a strategic complementarity in the decision to release information between firms.

The value of performance information in solving principal agent relationships has been well understood since Holmstrom (1979). Few papers have explored countervailing costs of uncovering such information. One exception is Arya, Glover and Sunder (1998). They consider an environment in which the owners of a firm are unable to commit not to fire the manager if the firm's short-term performance is bad. To circumvent this, the firm's owners commit to imperfect performance monitoring so as to prevent the manager from being overly concerned with the short term performance of the firm. In their paper owners commit to imperfect monitoring to tie their own hands whereas in the model presented in this paper owners limit

this information because of its affect upon the degree of speculation among investors over the firm's price. Also, their model does not suggest any link between the actions of individual firms in an industry.

This paper proceeds as follows. Section 2 describes the basic single firm model. Section 3 analyses the optimal accounting choice of the owner. Section 4 extends the environment to one in which there are multiple firms in the industry. Here I show the strategic complementarity in information choice. Section 5 provides a brief conclusion.

2.2 Single Firm Problem

2.2.1 The Environment

To begin the analysis I consider the problem of a single firm in isolation. The true underlying profitability of the firm is denoted by θ . The true value of θ is unknown and investors disagree about its distribution. This is the source of speculation in the model. There are 3 periods denoted by $t = 0, 1, 2$. For simplicity I abstract from discounting. The timing of events is as follows.

At the start of period 0 the firm is fully owned by an agent who shall be referred to as the founder of the firm. The operation of the firm requires a manager. The founder sets the incentive contract for the manager and chooses the governance structure of the firm. The governance structure is represented by a signal which is released at date 1. The founder must choose the precision (τ_a) of this signal. This decision should be thought as a reduced form way for representing things such as the type of auditor the owner selects. I assume that whatever information the owner receives about the performance of the firm cannot be hidden from investors and hence it becomes public information. Subsequent to the selection of τ_a the founder sells the firm to an investor (there is no loss of generality here since the founder could trivially sell the firm to herself). Finally, the manager selects their effort level e . This choice is unobservable and hence the firm is subject to a classic moral hazard problem.

In period 1 two signals of the firm's profitability are publicly released. The first of these, a , shall be referred to as the accounts of the firm. The second signal, s , comes from outside information about θ . From the perspective of the firm, the informativeness of the outside

signal is taken as given. Subsequent to these signals being released investors are able to trade the firm among themselves.

In period 3 the output of the firm y is realized and is publicly observed. The output of the firm is paid to whoever owns the firm at date 3. The owner must also pay whatever is due to the manager.

The realized output of the firm is determined by the profitability of the firm and the effort of the manager. Assume that these two components enter additively so that

$$y = \theta + e.$$

The accounts of the firm provide a noisy signal of the final output of the firm. In particular

$$a = \theta + e + \varepsilon_a$$

where the additive error term ε_a is distributed $N\left(0, \frac{1}{\tau_a}\right)$. The founder is able to costlessly select any non-negative value for τ_a . The outside signal which is released at the start of period 1 is an imperfect signal of the firm's profitability θ . Specifically,

$$s = \theta + \varepsilon_s$$

where the additive error term ε_s is distributed $N\left(0, \frac{1}{\tau_s}\right)$. The precision of this signal, τ_s , is determined exogenously and is public knowledge. This signal will be endogenized in Section 4 where it will come from the accounts released by other firms in the industry.

The profitability of the firm, θ , is distributed normal with a zero mean. Investors disagree only in their assessment of the precision of this distribution. Investors are indexed by i so that investor i believes that $\theta \sim N\left(0, \frac{1}{\tau_\theta^i}\right)$. Let τ_θ^+ (τ_θ^-) denote the precision used by the agent with the strongest (weakest) prior. We will see below that it is only the value of these extreme beliefs which will matter in equilibrium. I assume that agents have heterogeneous beliefs in the sense that they agree to disagree. Agent i knows that other agents assess the variance of θ differently but this does not cause them to alter their τ_θ^i . This is the only departure from full rationality in the model. The founder believes that $\theta \sim N\left(0, \frac{1}{\tau_\theta^F}\right)$ with $\tau_\theta^F \in [\tau_\theta^-, \tau_\theta^+]$. The

exact value of τ_θ^F will not affect the equilibrium. All investors and the founder are assumed to be risk neutral.

The manager believes that $\theta \sim N\left(0, \frac{1}{\tau_\theta^M}\right)$. The manager's private cost of exerting effort is

$$c(e) = \frac{e^2}{2\beta}.$$

The manager's utility function is of the form

$$u(W, e) = -e^{-r(W-c(e))}$$

where W is the total monetary payment to the manager at date 2. The manager has an outside opportunity of $-e^{-rw}$.

I assume that investors are unable to short sell the firm's stock. This prevents infinite trade volumes which would arise when agents disagree over the value of the firm.

The contract which the owner can offer the manager is limited to being linear in form. In addition, I assume for simplicity that the manager's contract is only contingent upon observables at date 1, a and s . The payment to the manager is thus

$$W = w_0 + w_a a + w_s s.$$

It is important to understand the role of these two assumptions. I have assumed that the manager is not paid based upon the final output of the firm to simplify the model. Since y is informative for the manager's choice of effort then an (unrestricted) optimal contract *would* be sensitive to y . However, allowing for this means that date 1 investors not only disagree about the expected output of the firm but also the expected payment to the manager. This complicates the exposition without adding new ideas. This assumption does not qualitatively alter the analysis. The assumption can be justified by supposing that the manager has a shorter horizon than the firm so that they must be paid at the end of date 1. This is reasonable if we interpret the final output of the firm as its entire future net worth. The assumption that the manager's incentive scheme must be linear is primarily made to avoid the Mirrlees problem (Mirrlees (1974)). Mirrlees shows that the tails of the normal distribution are perfectly informative for

the actions of the manager and hence moral hazard problem can be completely avoided through the use of extreme punishments and rewards. Hence if the manager's contract is unrestricted there is no governance benefit to improved monitoring. If we suppose that there is some finite limit to the magnitude of the transfer which can be enforced between the manager and the owners of the firm (such as a standard limited liability assumption) then such a scheme is no longer viable. With a limited liability assumption the firm's accounts would be valuable in providing incentives to the manager. This is all I need to generate the qualitative results I find below. I have made the stronger assumption of linearity so as that I can find an analytic solution for the manager's contract.

2.3 Equilibrium

We now solve the model. Before considering the decision of the founders we solve for the equilibrium prices in each period as a function of the contract and accounts chosen by the founder.

2.3.1 Equilibrium Prices

Let p_t denote the market clearing price of the firm in period t . The price of the firm in period 2 is equal to its realized value so that

$$p_2 = y - W.$$

In period 1 investors have observed both a and s and use these to update their belief about the expected value of the firm in date 2. Investor i 's updated expectation of p_2 is

$$E_1(p_2|a, s; \tau_\theta^i) = e - w_0 - w_a a - w_s s + \left(\frac{\tau_a \tilde{a} + \tau_s s}{\tau_\theta^i + \tau_a + \tau_s} \right)$$

where I have defined $\tilde{a} = a - e$. The short sale constraint implies that at any point in time the price of the firm is determined by the investor with the highest expected value of the firm.

Thus

$$p_1 = e - w_0 - w_a a - w_s s + \max_i \left\{ \frac{\tau_a \tilde{a} + \tau_s s}{\tau_\theta^i + \tau_a + \tau_s} \right\}.$$

From this expression we can see that only the investor with the strongest or weakest prior will determine the date 1 value of the firm. If good news is released at date 1 ($\tau_a \tilde{a} + \tau_s s > 0$) then the agent with the weakest prior (τ_θ^-) will be the most optimistic about the firm and thus will set its value. Conversely if the news is below expectations ($\tau_a \tilde{a} + \tau_s s < 0$) then the investor with the strongest prior will be least swayed by the news and thus will hold the highest valuation of the firm. The equilibrium value for the price of the firm at date 1 can thus be written as

$$p_1 = e - w_0 - w_a a - w_s s + \max \left\{ \left(\frac{\tau_a \tilde{a} + \tau_s s}{\tau_\theta^+ + \tau_a + \tau_s} \right), \left(\frac{\tau_a \tilde{a} + \tau_s s}{\tau_\theta^- + \tau_a + \tau_s} \right) \right\}.$$

Returning to date 0, investor i 's expectation of p_1 is

$$E_0(p_1; \tau_\theta^i) = (1 - w_a) e - w_0 + \frac{C}{\sqrt{2\pi}} \sqrt{(\tau_a + \tau_s) \left[\frac{(\tau_a + \tau_s)}{\tau_\theta^i} + 1 \right]}$$

where

$$C \equiv \frac{(\tau_\theta^+ - \tau_\theta^-)}{(\tau_\theta^- + \tau_a + \tau_s)(\tau_\theta^+ + \tau_a + \tau_s)} > 0.$$

It is clear that $E_0(p_1; \tau_\theta^i)$ is strictly decreasing in τ_θ^i and hence the date zero value of the firm will be determined by the belief of the investor with the weakest prior. Therefore the equilibrium price of the firm at date zero is

$$\begin{aligned} p_0 &= [(1 - w_a) e - w_0] + V^{\text{Spec}}(\tau_a) \\ &\text{where} \\ V^{\text{Spec}}(\tau_a) &\equiv \frac{(\tau_\theta^+ - \tau_\theta^-)}{\sqrt{2\pi\tau_\theta^-}} \eta^{\frac{1}{2}} (\tau_\theta^- + \eta)^{-\frac{1}{2}} (\tau_\theta^+ + \eta)^{-1} \\ \eta &\equiv \tau_a + \tau_s \end{aligned}$$

where η is the precision of all the information released at date 1. Thus at the end of date 0 the founder will sell the firm to the investor who has the weakest prior belief about the firm. This investor buys the firm at date 0 because they believe that θ is the most volatile and hence expect the biggest divergence in beliefs at date 1. Put differently, this investor expects the largest speculative resale premium at date 1.

There are two components to the firm's price at date 0. The first of these represents the “real” value of the firm, its expected output less any payments to the manager. Absent the possibility of speculation this would be the date 0 price of the firm. The second term, V^{Spec} , represents the speculative component of the firm's price. This comes purely from the fact that investors will disagree in their assessment of the firm's value at date 1. The date 0 investor values this because it presents the possibility that they will be able to sell to more optimistic investor at date 1. Observe that this speculative component is zero if all investors have the same prior ($\tau_{\theta}^+ - \tau_{\theta}^-$). The speculative component disappears if the information released at date 1 fully reveals the value of firm ($\eta = \infty$) or is not informative at all ($\eta = 0$). This is because the speculative component of the firm's price requires that investors disagree over the firm's value at date 1. This requires that they have some information to disagree over ($\eta > 0$) but that this information does not reveal the firm's value with certainty ($\eta < \infty$).

2.3.2 Founder's Problem

The founder selects the precision of the firm's accounts (τ_a) and the incentive contract of the manager (w_0, w_a, w_s, e) so as to maximize the price at which they can sell the firm in date 0. Formally, her problem is

$$\max_{w_0, w_a, w_s, \tau_a, e} (1 - w_a)e - w_0 + V^{\text{Spec}}(\tau_a) \quad (2.1)$$

subject to

$$c'(e) = w_a \quad (2.2)$$

$$\begin{aligned} & w_0 + w_a e - c(e) \\ -\frac{r}{2} \left[\frac{(w_a + w_s)^2}{\tau_{\theta}^M} + \frac{w_a^2}{\tau_a} + \frac{w_s^2}{\tau_s} \right] & \geq \underline{w} \end{aligned} \quad (2.3)$$

$$\tau_a \geq 0 \quad (2.4)$$

The first constraint (2.2) is the incentive compatibility constraint which ensures that effort level e is chosen by the manager. The second constraint (2.3) is the standard managerial

participation constraint. For the remainder of the analysis I normalize $\underline{w} = 0$. This ensures that the founder will always wish to employ the manager.

2.3.3 Founder's Solution

The optimal managerial contract is given in the appendix. The founder's optimal choice of τ_a is characterized in the following proposition.

Proposition 14 *Let τ_a^* denote the founder's optimal choice of the precision of the firm's accounts. It must be that $\tau_a^* + \tau_s > \hat{\eta}$ where*

$$\hat{\eta} \equiv \frac{1}{4} \left(-\tau_\theta^- + \sqrt{(\tau_\theta^-)^2 + 8\tau_\theta^- \tau_\theta^+} \right)$$

If $\tau_a^* \in (0, \infty)$ then τ_a^* must satisfy

$$(2r\beta^2) \left(\frac{V^{Real}(\tau_a^*)}{\tau_a^*} \right)^2 = V^{Spec}(\tau_a^*) \left[(\tau_\theta^+ + \eta^*)^{-1} + \frac{1}{2} (\tau_\theta^- + \eta^*)^{-\frac{1}{2}} - \frac{1}{2} \eta^{*-1} \right]$$

where

$$\begin{aligned} V^{Real}(\tau_a) &\equiv \left[2\beta \left(\frac{r\beta}{\tau_\theta^M + \tau_s} + 1 + \frac{r\beta}{\tau_a} \right) \right]^{-1} \\ \eta^* &\equiv \tau_a^* + \tau_s. \end{aligned}$$

Proof. See Appendix. ■

To understand Proposition 14 it is useful to write the founder's objective given optimal choices of $\{w_0, w_a, w_s, e\}$ for a given value of τ_a . Written in this way the founder's objective is:

$$V(\tau_a) = V^{Real}(\tau_a) + V^{Spec}(\tau_a) \tag{2.5}$$

The first term $V^{Real}(\tau_a)$ is simply the value of the firm absent any speculation ($\tau_\theta^+ - \tau_\theta^- = 0$). It represents the expected value created by the effort of the manager net of any payments required to induce that effort. $V^{Real}(\tau_a)$ is unambiguously increasing in τ_a . As the precision of the firm's accounts improves the manager's effort choice becomes increasingly observable. This allows the founder to offer higher powered incentives without forcing the manager to bear

more risk. In an environment without speculation the founder would set $\tau_a^* = \infty$.

The second term in the founder's objective $V^{\text{Spec}}(\tau_a)$ represents the speculative component of the firm's date 0 price. Observe that $V^{\text{Spec}}(\tau_a) > 0$ only if there is heterogeneity among investor's beliefs ($\tau_\theta^+ > \tau_\theta^-$). There is a non-monotonic relationship between the accuracy of the firm's accounts and the expected speculative component of the firm's price:

$$\begin{aligned} \frac{\partial V^{\text{Spec}}(\tau_a)}{\partial \tau_a} &\geq 0 \text{ for } \eta \leq \hat{\eta} \\ \frac{\partial V^{\text{Spec}}(\tau_a)}{\partial \tau_a} &\leq 0 \text{ for } \eta \geq \hat{\eta} \end{aligned}$$

The intuition for this is as follows. Speculation over the value of the firm comes from the fact that investor disagree about how to use the signals that are released at date 1. If no information is released at date 1 ($\eta = 0$) then there is nothing for investors to disagree over and hence there is no speculative premium at date 1. Starting from this point, an increase in τ_a gives investors some imperfect information which sponsors disagreement over the value of the firm. This occurs up until the total precision of all information released at date exceeds $\hat{\eta}$. When τ_a increases beyond this point it serves to reduce the divergence in beliefs. As the firm's accounts become increasingly informative then the accuracy of the firm's accounts outweigh any initial heterogeneity in beliefs. In the limit when $\tau_a \rightarrow \infty$ each investors expected value of the firm converges to a and thus there is no speculation.

From this discussion it is clear that we must have $\tau_a^* + \tau_s > \hat{\eta}$. Increasing τ_a when η is below $\hat{\eta}$ serves to improve the monitoring of the manager as well as increase the expected speculation at date 1. When η is greater than $\hat{\eta}$ there is a trade-off in the founders choice. Improving the accounts of the firm helps solve the managerial incentive problem ($\frac{\partial}{\partial \tau_a} V^{\text{Real}}(\tau_a) > 0$) but by releasing more information the accounts serve to reduce the difference of opinion and hence the degree speculation at date 1 ($\frac{\partial}{\partial \tau_a} V^{\text{Spec}}(\tau_a) < 0$).

2.3.4 Changes in the Heterogeneity of Beliefs

We now ask how changes in the degree of heterogeneity of investor beliefs affects the founder's choice of accounting accuracy τ_a . As already noted above, if there is no heterogeneity in investors beliefs ($\tau_\theta^+ - \tau_\theta^- = 0$) the founder will optimally select perfect accounting accuracy.

In this instance the firm's price has no speculative component and thus the founder chooses to monitor the performance of the manager without noise. If the disagreement among investors is sufficiently large then the optimal selection of τ_a will be finite in order to induce speculation at date 1. This is formalized in the following Proposition².

Proposition 15 *If $\hat{\eta} \leq \tau_s$ then τ_a^* must be finite if*

$$\frac{(\tau_\theta^+ - \tau_\theta^-)}{\sqrt{\tau_\theta^-}} > \frac{\sqrt{2\pi} \left[2\beta \left(\frac{r\beta}{\tau_\theta^M + \tau_s} + 1 \right) \right]^{-1}}{\tau_s^{\frac{1}{2}} (\tau_\theta^- + \tau_s)^{-\frac{1}{2}} (\tau_\theta^+ + \tau_s)^{-1}} \quad (2.6)$$

Conversely, if $\hat{\eta} > \tau_s$ then τ_a^ must be finite if*

$$\frac{(\tau_\theta^+ - \tau_\theta^-)}{\sqrt{\tau_\theta^-}} \geq \frac{\sqrt{2\pi} \left[2\beta \left(\frac{r\beta}{\tau_\theta^M + \tau_s} + 1 \right) \right]^{-1}}{(\hat{\eta})^{\frac{1}{2}} (\tau_\theta^- + \hat{\eta})^{-\frac{1}{2}} (\tau_\theta^+ + \hat{\eta})^{-1}}. \quad (2.7)$$

Proof. See Appendix. ■

Next, we consider how an increase in the heterogeneity of beliefs affects the founder's choice of accounting accuracy. In particular, consider what happens when the pool of investors expands to include investors who have weaker priors. This corresponds to a decrease in τ_θ^- while holding τ_θ^+ constant.

Proposition 16 *τ_a^* is weakly increasing in τ_θ^- .*

Proof. See Appendix. ■

To understand this result observe that the real component of the firm's value $V^{\text{Real}}(\tau_a)$ is unaffected by the heterogeneity in investor's beliefs (τ_θ^- and τ_θ^+). A change in τ_θ^- alters the speculative component of the firm's price through two channels. First, a fall in τ_θ^- increases the heterogeneity of beliefs at date 1. For any investor at date 0 this increases the speculative premium at date 0. The second effect comes from the fact that it is the investor with the weakest prior who purchases the firm in the first period. It is thus their evaluation of the degree of speculation at date 1 that determines the market price of this speculation at date 0.

²Note that the conditions given in Proposition 15 are not necessary conditions.

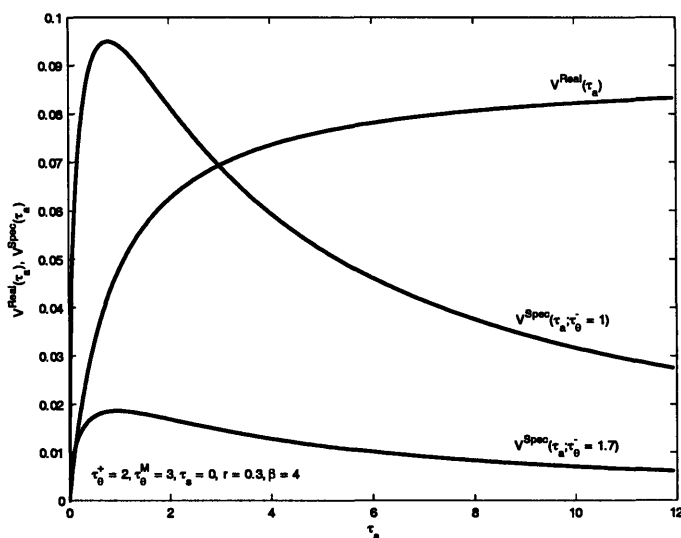


Figure 2-1: A Decrease in τ_{θ}^{-}

As the prior of this investor becomes weaker they anticipate increased uncertainty at date 1 and hence place a greater value upon the speculative value of the firm. Taken together these two forces increase the relative importance for the founder of sponsoring greater speculation by lowering τ_a . Figure 2-1 provides a numerical example that illustrates the effect of an increase in τ_{θ}^{-} . For any level of accounting accuracy τ_a the speculative component of the firm's price is decreased³. In the limit, as τ_{θ}^{-} approaches τ_{θ}^{+} , the speculative component of the firm's price disappears altogether. The result is that a given increase in τ_a (above $\widehat{\tau_a}$) reduces the speculative component of the firm's price by less (in absolute terms) while giving the same gain to the real value of the firm.

Extending this logic, we have that an increase in τ_{θ}^{-} must decrease the speculative component of the firm's price as well as the firm's price overall. This is formalized in the following Proposition.

Proposition 17 $V^{Spec}(\tau_a^*)$ and $V(\tau_a^*)$ are weakly decreasing in τ_{θ}^{-} .

Proof. See Appendix. ■

³It is easily verified that that $\frac{\partial}{\partial \tau_{\theta}^{-}} V^{Spec}(\tau_a) < 0$.

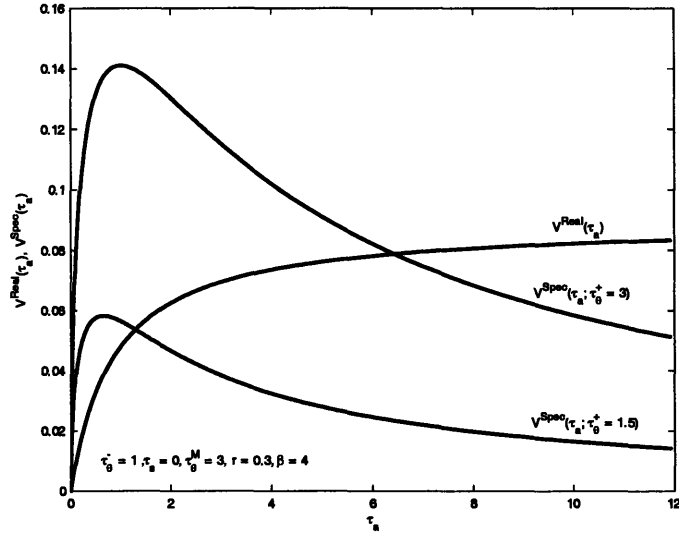


Figure 2-2: An Increase in τ_θ^+

Next, consider a strengthening of the belief of the investor with the strongest belief (i.e. an increase in τ_θ^+). The optimized value of the firm must rise as a result.

Proposition 18 $V(\tau_a^*)$ is weakly increasing in τ_θ^+ .

Proof. See Appendix. ■

Proposition 18 follows directly from the fact that $V^{\text{Spec}}(\tau_a)$ is increasing in τ_θ^+ for every value of τ_a . This increase in the firm's speculative premium implies that the founder can at least achieve the same date 0 price that would have been achieved prior to the increase. Figure 2-2 provides a numerical example that illustrates the effect of an increase in τ_θ^+ .

Although a rise in τ_θ^+ increases the date 0 value of the firm, its effect upon the optimal choice of accounting accuracy τ_a is ambiguous. This is because there are two conflicting channels through which τ_θ^+ affects the founder's choice. First, a rise in τ_θ^+ increases the disagreement among investors at date 0 and thus raises the value of the speculative component of the firm's price. If we began with a case where there was no disagreement among investors ($\tau_\theta^+ = \tau_\theta^-$) then the founder would optimally choose fully accurate accounts (because $V^{\text{Spec}}(\tau_a) = 0$). A rise in τ_θ^+ would then enable the founder to achieve speculative rents if they offered imperfect accounts (finite τ_a). A rise in τ_θ^+ has a second, countervailing influence upon the founder's decision.

As the belief of the investor with the strongest belief is further strengthened (τ_θ^+ rises) they become less influenced by the firm's accounts at date 1. As a result when the founder increases τ_a this will have a smaller impact upon the disagreement between investors. Through this channel the founder pays a smaller cost of marginal improvements in governance. The net effect of these two forces is, in general, ambiguous. Since a rise in τ_θ^+ has an ambiguous effect upon the choice of τ_a^* it is also true that the speculative component of the firm's price may also rise or fall as a result of the change.

Combining Propositions 17 and 18 we have that an increase in the heterogeneity of beliefs, through some combination of an increase in τ_θ^+ and a decrease in τ_θ^- must lead to an increase in the date zero value of the firm.

2.3.5 Changes in Outside Information τ_s

Not all the information which is released about the firm at date 1 can be controlled by the founder. External sources of information about the expected profits of the firm can come from many sources. Examples include analyst's research or information in the business press. In the next section I will endogenize this external information by supposing that it comes from the accounts of other firms. But for now we will treat it as being exogenous. We now ask what is the effect of a change in the precision of the external signal (τ_s).

The value of speculation at date 1 is determined by the total amount of information η which is released in that period. An increase in the precision of the outside signal can be fully undone by a decrease in the precision of the firm's accounts. However the concavity of $V^{\text{Real}}(\tau_a^*)$ implies that the founder faces increasing marginal costs of such an offset. As a result, it cannot be optimal for the founder to fully offset (or more) an increase in external information. It follows that an increase in the precision of the external signal must lead to an increase in the total amount of information released at date 1. As a result the speculative component of the firm's price must decrease⁴. This is formalized in the following Proposition.

Proposition 19 *The combined precision of the firm's accounts and the external signal $\tau_a^* + \tau_s$ is weakly increasing in τ_s . The speculative component of the firm's date 0 price $V^{\text{Spec}}(\tau_a^*)$ is*

⁴This follows from the fact that the founder always sets τ_a above $\widehat{\tau}_a$ so that $V^{\text{Spec}}(\tau_a)$ is unambiguously decreasing in τ_a .

weakly decreasing in τ_s .

Proof. See Appendix. ■

Faced with an increase in the precision of outside information the founder may choose to raise or lower the precision of their own firm's accounts (depending upon parameter values). Figure 2-3 provides a numerical example which illustrates how an increase in τ_s alters the founder's objective. The speculative component of the firm's price $V^{\text{Spec}}(\tau_a)$ is shifted horizontally by the size of the change in τ_s . It is clear that as τ_s grows large (above $\hat{\eta}$) that this diminishes the speculative rents which the firm is able to command. The intuition is simply that as more accurate external information is made available at date 1 then the difference of opinion between investors at date 1 is diminished. Thus lowering the speculative component of the firm's price. In this limit, as the external signal becomes perfectly informative ($\tau_s \rightarrow \infty$) for θ there will be no disagreement among investors at all. In this case $V^{\text{Spec}}(\tau_a) = 0$ and the founder will simply choose their own accounts to be fully informative ($\tau_a = \infty$). This is formalized in the following Proposition.

Proposition 20 *If the outside signal contains no noise (so that $\tau_s = \infty$) then the founder will choose perfectly informative accounts ($\tau_a^* = \infty$).*

Proof. See Appendix. ■

As Figure 2-3 indicates, the real component of the firm's price is increasing in the precision of the outside signal ($\frac{\partial}{\partial \tau_s} V^{\text{Real}}(\tau_a) > 0$). This information is used by the founder to benchmark the performance of the manager. As information on θ becomes more accurate the founder is better able to disentangle the firm's accounts as to provide a more accurate measure of the manager's effort. As a result, an increase in the precision of the outside signal has an ambiguous effect upon the firm's value. It will lower the speculative component of the firm's price but may lead to an increase in its value overall by allowing the founder to more effectively monitor the manager.

2.4 Endogenizing the External Signal

Until now the precision of the external signal which is released at date 1 has been taken as given. However if we think of θ as the profitability of an industry then it is natural to suppose

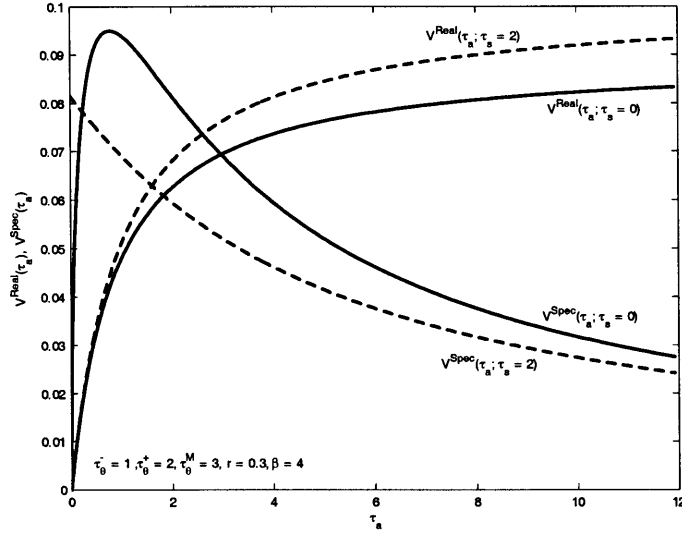


Figure 2-3: An Increase in τ_s

that an important source of information which will be released at date 1 is the accounts of other firms in the same industry. To formalize this suppose that there are N firms in the industry indexed by $j \in \{1, 2, \dots, N\}$. Each firm shares the common underlying profitability θ . Suppose also that the only information released at date 1 is the accounts of each firm⁵. Assume that the noise in each firm's accounts $\varepsilon_{a,j}$ is independent across firms. From the perspective of firm j the external signal is now the accounts of all other firms. By Baye's rule, this signal is a weighted sum of the accounts of every other firm:

$$s_i = \sum_{k \neq j} \tau_{a,k} (a_k - e_k)$$

so that

$$s_j \sim N\left(0, \frac{1}{\tau_{s,j}}\right) \text{ with } \tau_{s,j} = \sum_{k \neq j} \varepsilon_{a,k}.$$

Assume that at date 0 each founder selects $\tau_{a,j}$ and the incentive contract of their manager simultaneously. Founders act independently and do not take into account the effect of their

⁵The analysis could easily be extended to also allow for an additional exogenous source of information. This would not qualitatively alter the results obtained below.

choices upon the value of other firms. Trading takes place subsequent to each founder setting $\tau_{a,j}$ and their manager's incentive contract. At date 1 the accounts of each firm are released simultaneously. Investors observe each set of accounts and then trade based upon their updated information. Date 2 is unaffected. Note that the incentive scheme which the founder offers the manager is now, in effect, a function of the accounts of every firm.

The equilibrium market prices and the pattern of trading remains as described in the analysis of the single firm. As a result, the founder's objective remains as in (2.5) since $\tau_{s,j}$ is taken as given. An equilibrium is now defined as Nash vector of choices $\{\tau_{s,1}^{**}, \tau_{s,2}^{**}, \dots, \tau_{s,N}^{**}\}$ where each founder's choice of $\tau_{s,j}^{**}$ is optimal given the choices of all other founder's $\tau_{s,-j}^{**}$.

2.4.1 Equilibria

There is always an equilibrium in which each founder selects perfect accounting precision.

Proposition 21 $\tau_{a,j}^* = \infty \forall j \in \{1, \dots, N\}$ is a Nash equilibrium. In this equilibrium $V_j^{Spec}(\tau_{a,j}^*; \tau_{a,-j}^*) = 0 \forall j \in \{1, \dots, N\}$.

Proof. See Appendix. ■

Intuitively, Proposition 21 follows directly from Proposition 20. If every other firm chooses perfectly accurate accounts then the true value of θ will be known with certainty at date 1. The speculative component of the firm's price will therefore be zero. The founder's only incentive then is to set $\tau_{a,j} = \infty$ so as to maximize the real value of the firm. Observe that this equilibrium always exists no matter how diverse the beliefs of investors are. Extending the logic of Proposition 21 produces the following Corollary.

Corollary 22 *If there exists a Nash equilibrium in which any firm sets $\tau_{a,j}^*$ finite then it must be that $\tau_{a,k}^*$ is finite $\forall k \neq j$.*

This stresses the strategic complementarity between each founder's decision. In order for a founder to sponsor the speculative component of their firm's price it is required that other firms will release noisy information so as not to destroy the speculation.

While perfect information revelation is always a Nash equilibrium it is not necessarily optimal from the perspective of all N founders collectively. For example, this is true when the

heterogeneity of beliefs is sufficiently strong so that (2.7) holds. In this case, if each founder set $\tau_{a,j} = \frac{\hat{\eta}}{N}$ then the date zero price of each firm would be strictly higher than under the equilibrium in Proposition 21. In this example the speculative component of each firm's price would be maximized while the real component of their price would be strictly less than under the perfect accounting accuracy. However, $\tau_{a,j}^* = \frac{\hat{\eta}}{N} \forall j \in \{1, \dots, N\}$ is not a Nash equilibrium. Taking this strategy as given, the best response of any individual founder would be to raise⁶ $\tau_{a,j}$ above $\frac{\hat{\eta}}{N}$. The central message here is that the strategic complementarity involved in each founder trying to sponsor speculation can give rise to a prisoner's dilemma. Whether or not a social planner would like to facilitate coordination among founders so as to avoid the perfect information equilibrium would depend upon how the planner values the speculative component of each firm's price. If this is not valued by the planner then they would in fact wish to prohibit any attempt by founder's to collude in order to escape this equilibrium.

Equilibria With Imperfect Information

Under some parameter configurations there also exist equilibria in which each founder chooses to release imperfect accounts and thereby take advantage of speculation on the value of their firm. Providing an analytical characterization of these equilibria is made difficult by the fact that the best response correspondence for $\tau_{a,i}^*$ is not well behaved. To avoid these problems I investigate these finite Nash equilibria numerically.

Figure 2-4 draws the best response function for founder j for different values of τ_{θ}^- . It follows from Proposition 16 that the best response mapping is increasing in τ_{θ}^- . In the example, a symmetric Nash equilibria in which all firms choose accounts of limited precision exist for $\tau_{\theta}^- = 0.5, 1$. These are found where the best response functions intersects the 45 degree line. However when τ_{θ}^- increases to 1.5 the speculative motive becomes sufficiently weak that no Nash equilibrium exists in which any founder chooses finite accounting precision. This example indicates that all founders will choose to sponsor speculation if the heterogeneity of beliefs among investors is sufficiently strong. This is confirmed by Figure 2-5 which shows how the value of the finite symmetric Nash equilibria τ_a^{**} changes with τ_{θ}^- . As τ_{θ}^- falls, and the speculative motive becomes stronger, an equilibrium in which each founder commits to lower

⁶The best reponse would not necessarily be to raise $\tau_{a,j}$ to ∞ .

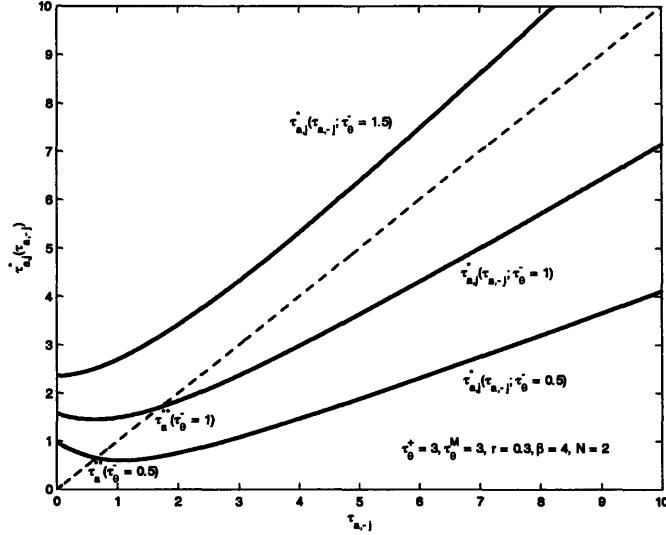


Figure 2-4: Founder's Best Response Function

levels of accounting accuracy is sustainable. These equilibria only exist for $\tau_\theta^- \leq 1.33$. The dotted line in Figure 2-5 shows the date 0 price of each firm associated with the finite nash equilibria. As τ_θ^- increases, and the available speculative rents diminish the price of each firm falls. Note that the real component of each firm's value must be increasing with τ_θ^- since each founder is choosing increasingly accurate accounts⁷. The numerical example indicates that the loss to the speculative value of the firm, both through the reduction in the heterogeneity of beliefs as well as the increased precision of information released at date 1 outweighs this.

For the parameter values in Figure 2-5 where $\tau_\theta^- \leq 1.33$ there exists two equilibria: the finite equilibria indicated and one in which each founder selects perfect accounting precision. The date 0 price which each founder receives in the equilibrium with perfect accounting accuracy is $V_j(\tau_{a,j}^* = \infty; \tau_{a,-j}^* = \infty) = 0.125$. In this example there exists a $\widetilde{\tau_\theta^-} < 1.33$ such that if $\tau_\theta^- < \widetilde{\tau_\theta^-}$ then each founder receives a strictly higher date 0 price in the equilibrium with limited accounting accuracy. Conversely when $\tau_\theta^- \geq \widetilde{\tau_\theta^-}$ each founder would receive a higher price in

⁷Note that this increase in $V^{\text{Real}}(\tau_a)$ will come through two channels. First an increase in the precision of the firm's own accounts which improves the monitoring of the manager. In addition the accuracy of all other firm's accounts are increasing so that a founder is better able to benchmark the performance of their own manager. Both forces will serve to raise $V^{\text{Real}}(\tau_a)$.

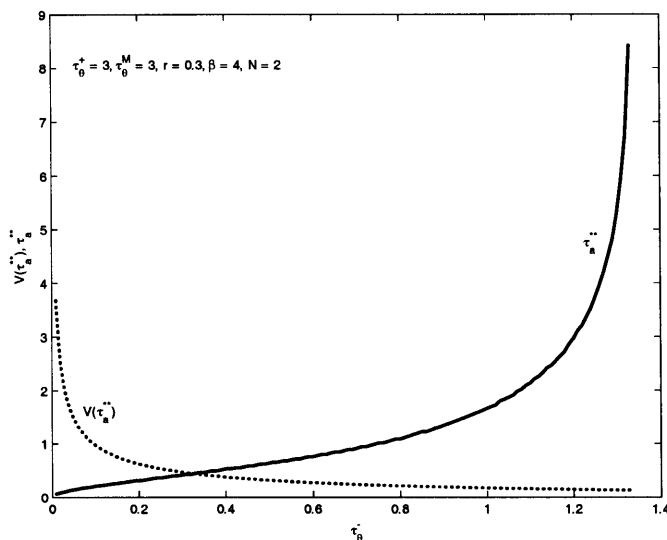


Figure 2-5: Nash Equilibria as τ_{θ}^- Changes

the equilibrium with perfect accounting accuracy. In this example $\widetilde{\tau_{\theta}^-} \approx 1.32$.

When $\tau_{\theta}^- < \widetilde{\tau_{\theta}^-}$ then the speculative component of each firm's price outweighs the loss to the real value of the firm. If some form of communication among each founder were possible then they would choose to coordinate on the equilibrium with limited accounting accuracy. In fact this equilibrium Pareto dominates the equilibrium with full accounting accuracy⁸. The suggestion is that if a social planner could influence which equilibrium was selected they would encourage the equilibrium in which each firm issued imperfect accounts. If the planner only valued the real component of each firm's price then it is clear that they would always encourage the equilibrium with perfect accounting accuracy.

The converse is true when $\tau_{\theta}^- \geq \widetilde{\tau_{\theta}^-}$ but sufficiently low so that a finite equilibrium still exists. In this case each founder would prefer to be in an equilibrium where each firm elected to use perfect accounting accuracy⁹. In this scenario a social planner would wish to foster

⁸This is easily seen by observing that the price in the stock market each period leaves each investor with zero expected utility. Similarly the manager's contract always leaves them with expected utility equal to their outside option. As a result the equilibrium in which the founder's utility is highest must Pareto dominate the other.

⁹This possibility of Pareto inferior equilibria with finite precision comes from the fact that $\frac{\partial^2}{\partial \tau_a \partial a} V^{\text{Real}}(\tau_a) > 0$.

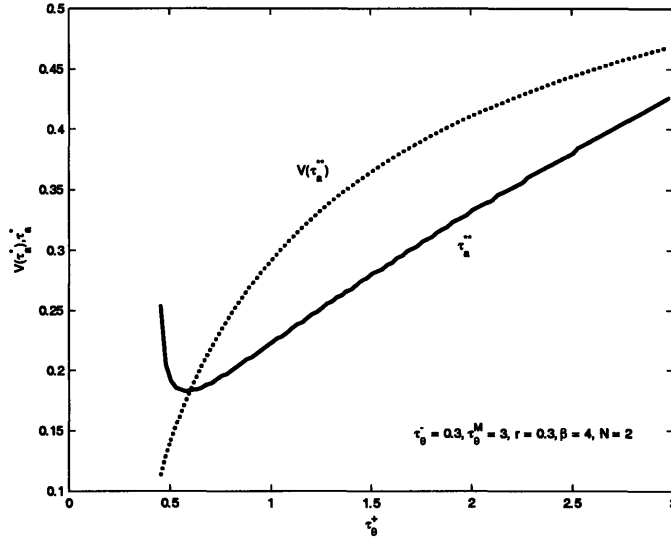


Figure 2-6: Nash Equilibria as τ_θ^+ Changes

coordination on the equilibrium with complete information release. Such a policy, would serve to raise the value of each firm. If in practice it is difficult to evaluate which parameter regime exists the role for policy (if any) maybe hard to determine. Simulating this model over a wide range of parameters suggests that the region over which the equilibrium with finite accounting accuracy depresses the price of each firm relative to perfect accounts is very narrow (as with this example).

Figure 2-6 provides another numerical example of parameters under which there exists an equilibrium where each founder sets finite accounting accuracy. This example demonstrates how that equilibrium changes with τ_θ^+ . In this example finite equilibria exist whenever $\tau_\theta^+ \geq 0.452$. As above, the heterogeneity of beliefs among investors must be sufficiently strong to sustain such an equilibrium. Similar to the case above there exists a $\widetilde{\tau_\theta^+} \approx 0.472$ such that when $\tau_\theta^+ \in [0.452, 0.472]$ then the price which founder's achieve under the finite equilibrium is less than they would obtain in the equilibrium with perfect information. Conversely when heterogeneity among investors beliefs is sufficiently strong so that $\tau_\theta^+ > \widetilde{\tau_\theta^+}$ then the price of each firm is strictly higher under the equilibrium where each firm produces accounts of limited accuracy.

The precision of accounts which founders choose in these finite equilibria has a non-monotonic

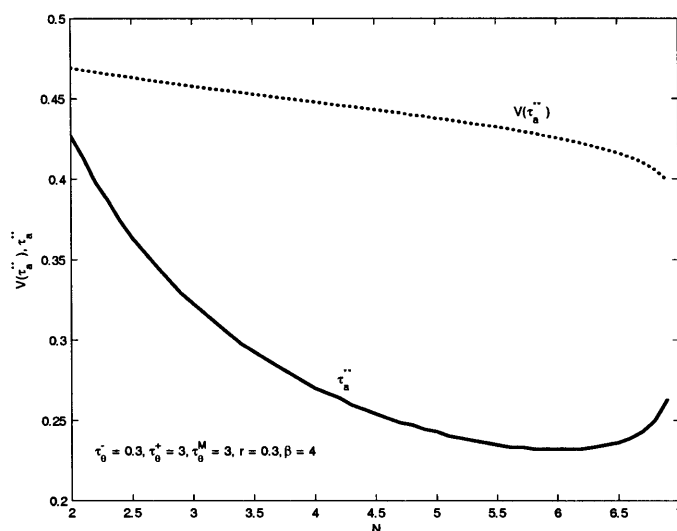


Figure 2-7: Nash Equilibria as N Changes

relationship with τ_θ^+ . At first, as τ_θ^+ grows above 0.452 the predominate force is that investor heterogeneity is growing and founder's opt for less informative reports so as to exploit this increased heterogeneity. However as τ_θ^+ continues to rise a second countervailing force dominates. Now founder's make use of the fact that the most confident investors have increasingly strong priors and as such are less influenced by the accounts of the firm. As a result increasing the accuracy of the firm's accounts carries a lower cost to the speculative value of the firm. The result is that as τ_θ^+ becomes progressively large the finite equilibrium involves more accurate accounts. This raises the real value of the firm, both through the improve signal from the firm as well as the improved ability to benchmark the manager's performance. Of course the real value of the firm remains strictly below that achieved under the equilibrium with full accounting accuracy.

Figure 2-6 provides an example of finite symmetric Nash equilibria and illustrates how these change with the number of firms in the industry N . In this example Nash equilibria in which each founder selects finite precision for their accounts exist whenever $N \leq 7.03$. Equilibria with finite accounting precision are harder to obtain when there are more firms in industry. This is because accounting imprecision is (among founders) a public good. An the number

of firms in the industry grows the “tragedy of the commons” between them becomes stronger. If we allowed the correlation in the profitability of each firm to vary then we would obtain a similar result. Equilibria with finite accounting precision will exist when this correlation is below an upper limit.

2.5 Conclusion

This paper provides an explanation of why owners may wish to implement imprecise monitoring of their firm. Accurate monitoring has the benefit of reducing the degree of moral hazard between the manager and the owner but it comes at the cost of reducing speculation on the firm’s value. If the heterogeneity of investor’s beliefs is sufficiently strong then owners will wish to pursue these speculative rents by allowing imprecise monitoring of the firm. The paper argues that episodes of accounting inaccuracy and speculation will occur in waves. To show this I demonstrate that there is a strategic complementarity in accounting accuracy across firms whose profitability is correlated. If all firms are releasing accurate information then there is little scope for speculation and hence it is optimal for each individual firm to release accurate accounts. Conversely, if all other firms are releasing inaccurate information then uncertainty and speculation is strong. In this case it is then optimal for each firm to release noisy accounts.

2.6 Appendix

2.6.1 Proof of Proposition 14.

To begin, I solve for the optimal managerial contract (w_0, w_a, w_s, e) for a given level of accounting precision τ_a . Let $\lambda^{IC}, \lambda^{IR} \geq 0$ be the Lagrange multipliers on incentive compatibility and individual rationality constraints. The problem is

$$\begin{aligned} & \max_{w_0, w_a, w_s, e} (1 - w_a)e - w_0 + \lambda^{IC} (w_a - e\beta) \\ & + \lambda^{IR} \left(w_0 + w_a e - \frac{e^2 \beta}{2} - \frac{r}{2} \left[\frac{(w_a + w_s)^2}{\tau_\theta^M} + \frac{w_a^2}{\tau_a} + \frac{w_s^2}{\tau_s} \right] - \underline{w} \right) \end{aligned}$$

The first order conditions for this problem yield:

$$\begin{aligned}
w_0 : \lambda^{IR} &= 1 \\
w_a : \lambda^{IC} - r \left[\frac{(w_a + w_s)}{\tau_\theta^M} + \frac{w_a}{\tau_a} \right] &= 0 \\
w_s : \frac{(w_a + w_s)}{\tau_\theta^M} + \frac{w_s}{\tau_s} &= 0 \\
e : \lambda^{IC} &= \frac{1 - w_a}{\beta}
\end{aligned}$$

Solving these simultaneously we get

$$\begin{aligned}
w_a^* &= \left[1 + r\beta \left(\frac{1}{\tau_s + \tau_\theta^M} + \frac{1}{\tau_a} \right) \right]^{-1} \\
w_s^* &= - \left(\frac{\tau_s}{\tau_s + \tau_\theta^M} \right) w_a^* \\
e^* &= \frac{w_a^*}{\beta} \\
w_0^* &= \underline{w} - \frac{(w_a^*)^2}{2\beta} \left[1 - r\beta \left(\frac{1}{\tau_s + \tau_\theta^M} + \frac{1}{\tau_a} \right) \right]
\end{aligned}$$

Substituting these into the objective function (and normalizing $\underline{w} = 0$) gives the expression for $V^{\text{Real}}(\tau_a)$ which appears in the Proposition.

2.6.2 Proof of Proposition 15.

For $\hat{\eta} \leq \tau_s$, $\tau_a = \infty$ cannot be optimal if $V^{\text{Spec}}(\tau_a = 0) > V^{\text{Spec}}(\tau_a = \infty)$. For $\hat{\eta} > \tau_s$, $\tau_a = \infty$ cannot be optimal if $V^{\text{Spec}}(\tau_a = \hat{\eta} - \tau_s) > V^{\text{Spec}}(\tau_a = \infty)$. Evaluating these expressions and rearranging gives the conditions in the Proposition.

2.6.3 Proof of Proposition 16.

First observe that

$$\frac{\partial V^{\text{Spec}}(\tau_a)}{\partial \tau_\theta^-} = -V^{\text{Spec}}(\tau_a) \left[(\tau_\theta^+ - \tau_\theta^-)^{-1} + \frac{(\tau_\theta^-)^{-1}}{2} + \frac{(\tau_\theta^- + \eta)^{-1}}{2} \right] \leq 0. \quad (2.8)$$

Next observe that

$$\begin{aligned} \left. \frac{\partial^2 V^{\text{Spec}}(\tau_a)}{\partial \tau_a \partial \tau_\theta^-} \right|_{\tau_a \geq \widehat{\tau}_a} &= V^{\text{Spec}}(\tau_a) \frac{(\tau_\theta^- + \eta)^{-2}}{2} + \\ &\frac{\partial V^{\text{Spec}}(\tau_a)}{\partial \tau_\theta^-} \frac{\partial V^{\text{Spec}}(\tau_a)}{\partial \tau_a} \left(\frac{1}{V^{\text{Spec}}(\tau_a)} \right)^2 \geq 0 \end{aligned} \quad (2.9)$$

where the inequality follows from combining (2.8) with the fact that

$$\frac{\partial V^{\text{Spec}}(\tau_a)}{\partial \tau_a} \leq 0 \text{ for } \tau_a \geq \widehat{\tau}_a.$$

Next, since $\tau_a^{*'}$ and $\tau_a^{*''}$ are optimal choices then we must have that

$$V^{\text{Real}}(\tau_a^{*'}) + V^{\text{Spec}}(\tau_a^{*'}; \tau_\theta^{-'}) \geq V^{\text{Real}}(\tau_a^{*''}) + V^{\text{Spec}}(\tau_a^{*''}; \tau_\theta^{-'}) \quad (2.10)$$

and

$$V^{\text{Real}}(\tau_a^{*''}) + V^{\text{Spec}}(\tau_a^{*''}; \tau_\theta^{-''}) \geq V^{\text{Real}}(\tau_a^{*'}) + V^{\text{Spec}}(\tau_a^{*'}; \tau_\theta^{-''}). \quad (2.11)$$

Subtracting (2.11) from (2.10) gives us that

$$V^{\text{Spec}}(\tau_a^{*'}; \tau_\theta^{-'}) - V^{\text{Spec}}(\tau_a^{*'}; \tau_\theta^{-''}) \geq V^{\text{Spec}}(\tau_a^{*''}; \tau_\theta^{-'}) - V^{\text{Spec}}(\tau_a^{*''}; \tau_\theta^{-''}). \quad (2.12)$$

We know from Proposition 14 that the founder will always select $\tau_a > \widehat{\tau}_a$. Thus we know from (2.9) that (2.12) can only hold if $\tau_a^{*'} \geq \tau_a^{*''}$. This establishes the Proposition.

2.6.4 Proof of Proposition 17.

Combining (2.8) with the fact that

$$\frac{\partial V^{\text{Spec}}(\tau_a)}{\partial \tau_a} \leq 0 \text{ for } \tau_a \geq \widehat{\tau}_a$$

implies that $V^{\text{Spec}}(\tau_a^{*'}; \tau_\theta^{-'}) \leq V^{\text{Spec}}(\tau_a^{*''}; \tau_\theta^{-''})$ because both $\tau_\theta^{-'} > \tau_\theta^{-''}$ (by construction) and $\tau_a^{*'} \geq \tau_a^{*''}$ (from Proposition 16). Next, observe that

$$V^{\text{Real}}(\tau_a^{*'}) + V^{\text{Spec}}(\tau_a^{*'}; \tau_\theta^{-''}) \geq V^{\text{Real}}(\tau_a^{*'}) + V^{\text{Spec}}(\tau_a^{*'}; \tau_\theta^{-'}) \quad (2.13)$$

which follows directly from (2.8). Combining (2.13) with (2.11) establishes that $V(\tau_a^{*''}) \geq V(\tau_a^{*'})$. This establishes the Proposition.

2.6.5 Proof of Proposition 18.

First observe that

$$\frac{\partial V^{\text{Spec}}(\tau_a)}{\partial \tau_\theta^+} = V^{\text{Spec}}(\tau_a) \left[(\tau_\theta^+ - \tau_\theta^-)^{-1} - (\tau_\theta^+ + \eta)^{-1} \right] \geq 0. \quad (2.14)$$

Since $\tau_\theta^-, \eta \geq 0$. Since $\tau_a^{*''}$ is optimal then it must be that

$$V^{\text{Real}}(\tau_a^{*''}) + V^{\text{Spec}}(\tau_a^{*''}; \tau_\theta^{*''}) \geq V^{\text{Real}}(\tau_a^{*'}) + V^{\text{Spec}}(\tau_a^{*'}; \tau_\theta^{*'}). \quad (2.15)$$

Using (2.14) we must have that

$$V^{\text{Real}}(\tau_a^{*'}) + V^{\text{Spec}}(\tau_a^{*'}; \tau_\theta^{*'}) \geq V^{\text{Real}}(\tau_a^{*''}) + V^{\text{Spec}}(\tau_a^{*''}; \tau_\theta^{*''}). \quad (2.16)$$

Finally combining (2.15) and (2.16) we get that $V(\tau_a^{*''}) \geq V(\tau_a^{*'})$. This establishes the proposition.

2.6.6 Proof of Proposition 19.

Since $\tau_a^{*'}$ and $\tau_a^{*''}$ are optimal choices then we must have that

$$\begin{aligned} V^{\text{Real}}(\tau_a^{*'}; \tau_s') + V^{\text{Spec}}(\tau_a^{*'}; \tau_s') &\geq \\ V^{\text{Real}}(\tau_a^{*''} - \Delta; \tau_s') + V^{\text{Spec}}(\tau_a^{*''} - \Delta; \tau_s') &\end{aligned} \quad (2.17)$$

and

$$\begin{aligned} V^{\text{Real}}(\tau_a^{*''}; \tau_s'') + V^{\text{Spec}}(\tau_a^{*''}; \tau_s'') &\geq \\ V^{\text{Real}}(\tau_a^{*'} + \Delta; \tau_s'') + V^{\text{Spec}}(\tau_a^{*'} + \Delta; \tau_s'') &\end{aligned} \quad (2.18)$$

where $\Delta \equiv \tau'_s - \tau''_s > 0$. Subtracting (2.18) from (2.17) gives

$$\begin{aligned} V^{\text{Real}}(\tau_a^*; \tau'_s) - V^{\text{Real}}(\tau_a^* + \Delta; \tau'_s - \Delta) &\geq \\ V^{\text{Real}}(\tau_a^* - \Delta_\eta; \tau'_s) - V^{\text{Real}}(\tau_a^* + \Delta + \Delta_\eta; \tau'_s - \Delta) &\end{aligned} \quad (2.19)$$

where $\Delta_\eta \equiv \tau_a^{*''} - \tau_a^* - \Delta$. Note that this last step makes use of the fact that $V^{\text{Spec}}(\tau_a; \tau_s)$ is simply a function of the sum of $\tau_a + \tau_s$ so that $V^{\text{Spec}}(\tau_a^*; \tau'_s) = V^{\text{Spec}}(\tau_a^* + \Delta; \tau'_s)$ and $V^{\text{Spec}}(\tau_a^{*''} - \Delta; \tau'_s) = V^{\text{Spec}}(\tau_a^{*''}; \tau''_s)$. Next, observe that

$$\frac{\partial}{\partial \Delta_\eta} \left\{ V^{\text{Real}}(\tau_a^* - \Delta_\eta; \tau'_s) - V^{\text{Real}}(\tau_a^* + \Delta + \Delta_\eta; \tau'_s - \Delta) \right\} \geq 0 \quad (2.20)$$

which follows from the fact that $\frac{\partial^2 V^{\text{Real}}}{\partial \tau_a^2} \leq 0$ and $\frac{\partial^2 V^{\text{Real}}}{\partial \tau_a \partial \tau_s} \geq 0$. For (2.19) to be consistent with (2.20) we must have that $\Delta_\eta \leq 0$. This establishes that $\tau_a^* + \tau'_s \geq \tau_a^{*''} + \tau''_s$. We know from Proposition 14 that the founder will always select $\tau_a > \widehat{\tau}_a$ which ensures that V^{Spec} is decreasing in $\tau_a + \tau_s$. It follows then that

$$V^{\text{Spec}}(\tau_a^*; \tau'_s) \geq V^{\text{Spec}}(\tau_a^{*''}; \tau''_s).$$

This establishes the Proposition.

2.6.7 Proof of Proposition 20.

When $\tau_s \rightarrow \infty$ the founder's problem in choosing τ_a becomes

$$\max_{\tau_a \geq 0} \left[2\beta \left(1 + \frac{r\beta}{\tau_a} \right) \right]^{-1}.$$

Since this objective is increasing in $\tau_a \forall \tau_a \geq 0$ then it must be that $\tau_a^* = \infty$. This establishes the Proposition.

Bibliography

- [1] Arya, Anil, Jonathan Glover and Shyam Sunder (1998), "Earnings Management and the Revelation Principle", *Review of Accounting Studies*, 3, 7-34.
- [2] Baskin, Jonathan and Paul Miranti (1997), *A History of Corporate Finance*, Cambridge University Press.
- [3] Bolton, Patrick, Jose Scheinkman and Wei Xiong, (2003), "Executive Compensation and Short-termist Behavior in Speculative Markets," mimeo.
- [4] Brooks, John (1973) *The Go-Go Years*, New York: Weybright and Talley.
- [5] Davis, Angela (2002), "Does Reporting Grossed-up or Barter Revenue Make a Difference?", *Journal of Accounting Research*, 40(2), 445 - 477.
- [6] Galbraith, John K. (1972), *The Great Crash: 1929*, Boston: Houghton Mifflin Company.
- [7] Graham, Carol, Robert Litan and Sandip Sukhtankar (2002), "The Bigger They Are, The Harder They Fall: An Estimate of the Costs of the Crisis in Corporate Governance", Brookings Working Paper.
- [8] Holmstrom, Bengt (1979), "*Moral Hazard and Observability*", *Bell Journal of Economics*, 10, 74-91.
- [9] Kahn, Jeremy (2000), "Presto Chango! Sales Are Huge!", *Fortune*, 141(6).
- [10] Macdonald, Elizabeth (2000), "Fess-Up Time", *Forbes*, September 18, 80-81.

- [11] Mirrlees, James (1974), "Notes on Welfare Economics, Information and, Uncertainty", in *Essays on Economic Behavior and Uncertainty*, ed. by M. Balch, D. Mcfadden and, S.-Y. Wu, Amsterdam: North Holland.
- [12] Kaplan, Steven and Jeremy Stein (1993), "The Evolution of Buyout pricing and Financial Structure in the 1980s", *Quarterly Journal of Economics*, 108(2), 313-357.
- [13] Wu, Min (2002), "Earnings Restatements: A Capital Market Perspective", Ph.D. Thesis, New York University.

Chapter 3

White Vans: Optimal Information Release in the Presence of Decentralized Signal Choice

Information gathering agencies often rely upon reports made to them by agents not directly under their control. The police, for example, rely upon sightings made by members of the public in the course of investigating a crime. If agents have some choice as to what type of information they can look-out for then this gives rise to a principal-agent problem. This chapter considers an environment in which agents choose to look where they think they are most likely to make a positive sighting. The police can influence the beliefs of the agent by the information they release. We show that when the police have weak information (flat prior) then it is optimal to hide this from the public so as to encourage random search. Conversely when the police have strong information which leaves them to choose between a few likely suspects then they will release their findings so as to direct search away from areas they have already ruled out.

3.1 Introduction

Information gathering agencies such as the police and the FBI rely upon reports made to them by members of the general public. The importance of policing informed by reports from the

community has grown in the post 9/11 environment. A key aspect of these relationships is that members of the public are not under the direct control of the police. They are free to look for information wherever they choose. One way the police can affect their decision is through their public announcements. In some instances the police will hide what they have learned so far and in other circumstances they will reveal this information. This paper provides a simple model to study the decision of whether or not the police will choose to reveal their preliminary information to the public.

This paper was originally motivated by a prominent example of this phenomenon: the Washington Sniper. An article in the *New York Times* (Kershaw (2002)) written after the capture of the sniper describes the events:

“Witnesses to some of the first shootings were able to describe only a glimpse at a fleeing white vehicle, information that the police quickly released and eventually used to put together composite sketches. So there was nothing but a white van ... to look for. With no description of the killer, many people, gripped with fear of another attack and seizing on any details in their personal lookout for the sniper, were in a kind of white-van haze.”

This example highlights a number of the key features of our theory. First, the principal-agent relationship between the police and the citizenry is clear. The police were only able to indirectly affect where people looked for clues through their release of preliminary information related to the crime. Second, the example highlights the tendency of people to gather information where they think they are most likely to make a positive sighting. The motivation for this may simply be the utility from seeing something related to a the crime¹ or it may also come from a desire to protect oneself. In this paper we assume this behavior. The white vans example also demonstrates the trade-off faced by the police in deciding whether or not to release their information. Releasing information stops random search and leads citizens to herd on the most likely signal. This is useful if the police have been able to rule out many other alternatives - in which case sightings in those locations would be unproductive. Counterbalancing this, concentrating on “white vans” may not be the most effective way to judge

¹This is similar to the reason why people slow down to look at traffic accidents.

between competing alternatives. The sniper was finally found in a blue sedan. Ex-post it was suggested that the concentration on white vans prolonged the investigation. Kershaw (2002) writes that the focus on white vans may have led people to overlook suspicious activity related to the sedan. In our model we show that under a simple condition looking in the most likely location is not the optimal way to gather information. This condition simply requires that a positive sighting is more informative than not seeing anything at all.

In our paper we consider an environment where the police have some information about a crime. They can either announce this information to the public or not. If they announce, the public will look out for the most likely suspect. Alternately if they hide the information the public will search randomly. We show that it is optimal for the police to reveal their information only when they are faced with a few likely suspects and have ruled out many other possibilities.

The paper proceeds as follows. Section 2 sets out the model. Section 3 solves for the optimal signal choice if the police are able to directly control the actions of the public. Section 4 considers the case where the police are only able to indirectly affect the actions of the public through the information they release. We show the conditions under which releasing the police's prior information is optimal. Section 5 provides a brief conclusion.

3.2 The Model

Consider the following scenario. A crime has been committed and the police need to decide where to investigate. There are N locations and the police can only investigate one of them. One location, denoted by θ , is the true location. The police will solve the crime if and only if they investigate the true location. Information on the true state can be gained through N signals which are labeled s_i for $i \in \{1, 2, \dots, N\}$. Using signal s_i will be referred to as "looking in location i ". The police are unable to collect these signals themselves. Instead they rely upon an agent (a representative member of the public) to collect this information and report it to the police. The agent can collect information from one of the N locations. When an agent looks in location i she either sees something ($s_i = 1$) or sees nothing ($s_i = 0$). A sighting is informative because it is more likely that a sighting is made in spot i if in fact the true state is

i. The model is assumed to be completely symmetric and so we can fully describe the menu of available signals by the following:

$$\begin{aligned}\Pr(s_i = 1|\theta = i) &= p \\ \Pr(s_i = 0|\theta = i) &= 1 - p \\ \Pr(s_i = 1|\theta \neq i) &= q \\ \Pr(s_i = 0|\theta \neq i) &= 1 - q\end{aligned}$$

We assume $p > q$ so that making a sighting in spot i increases our belief that i is the true state. In addition we assume that a positive sighting is more informative than not seeing anything. This requires that $p + q < 1$ ².

We assume that the agent chooses their signal so as to maximize the chance of making a positive sighting (i.e. seeing $s_i = 1$). This assumption is made to capture the idea that following the announcement of a crime people tend to look out for the “likely suspects”. We suppose this is motivated by a personal desire to see something related to the crime. Note also that we assume that the agent does not take into account the objective of the police. This assumption is made to capture the idea that in reality no member of the public thinks their signal choice will be pivotal in the police’s investigation. If we extended the current model to include many agents then (under some conditions) this would be a reasonable supposition.

Ex-ante all states are considered to be equally likely. The timing of the game is as follows. First the police privately receive some information about the true state. Based upon this information the police update their belief. This updated belief can be described by a vector of probabilities $\{\mu_1, \mu_2, \dots, \mu_N\}$ where μ_i denotes the belief that i is the true state and $\sum_{i=1}^N \mu_i = 1$. We place no restrictions upon this distribution. Without loss of generality we can order the states so that $\mu_i \geq \mu_{i+1} \forall i \in [1, N - 1]$ and $i \in J$. The police then choose whether or not to release their private information about θ to the agent. Based upon the announcement of the police the agent updates their belief. Next, the agent selects the signal from the location in

²It is useful to observe that this condition will hold whenever $p \leq \frac{1}{2}$ and so it holds whenever making a sighting is less likely than seeing nothing even in the correct location.

which they believe they are most likely to make a positive sighting³. Subsequently the agent observes the signal they have selected and truthfully reveals it to the police. The police, update their belief based upon the report of the agent and investigate in the location they judge most likely to be the true state of the world.

Observe that if the police reveal their private information then the agent will then also have the same beliefs as the principal. If the police do not reveal their information then the agent's belief will simply be uniform across all states and hence the agent will randomly choose their signal.

3.3 Optimal Signal Choice

We begin by determining which signal maximizes the probability of the police investigating in the correct location. This is the signal that would be chosen if no principal agent problem existed between the police and the public.

Proposition 23 *The signal which maximizes the expected probability of the police investigating the true location is s_2 .*

Proof. See Appendix. ■

To understand Proposition 23 it is useful to define the following constants

$$\bar{\mu} \equiv \mu_1 \frac{(1-p)}{(1-q)} \text{ and } \underline{\mu} \equiv \mu_1 \frac{q}{p}. \quad (3.1)$$

If the belief at location 1 is so strong that $\mu_2 < \underline{\mu}$ then no signal can alter the actions of the police. In this case any signal is optimal (trivially). Next suppose that $\mu_2 \in [\underline{\mu}, \bar{\mu}]$ ⁴. In this case the belief at location 1 is sufficiently strong that a sighting of $s_1 = 0$ will not lead the police to investigate location 2. It follows that s_1 is not a useful signal. In contrast, a positive sighting at location 2 ($s_2 = 1$) will raise the posterior belief at location 2 above that at location 1 and thus change the actions of the police. Thus s_2 is strictly preferred to s_1 . In this region

³If the agent is indifferent between more than one location they randomly choose among them.

⁴Our assumption that $p + q < 1$ ensures that $\bar{\mu} > \underline{\mu}$.

the strict preference for s_2 over s_1 comes directly from our assumption that a positive sighting ($s = 1$) is more informative than no sighting ($s = 0$).

The signal in location 2 is also preferred to any s_i (for $i > 2$). The intuition for this is as follows. Location 2 is the most promising alternative and s_2 is the signal which provides the most useful information for evaluating this alternative. Although signal s_i for $i > 2$ is not optimal it may still be preferred to the signal associated with the most likely location. The conditions under which this are true are given in the following Proposition.

Proposition 24 *Signal s_i is preferred to s_1 if and only if*

$$(i) \mu_2 \leq \bar{\mu}, \text{ or}$$

$$(ii) \mu_2 > \bar{\mu} \text{ and } \mu_i \geq \underline{\mu} + \mu_1 \left(\frac{1-p}{p} \right) \left(\frac{\mu_2 - \bar{\mu}}{\bar{\mu}} \right).$$

Proof. *See Appendix.* ■

The logic for case (i) remains the same as above. If the belief at location 1 is sufficiently strong then s_1 is not useful and so any signal is (at least weakly) preferred to s_1 . In case (ii) s_1 is a useful signal and so s_i must be sufficiently valued if it is to be preferred to s_1 . If $\mu_i \leq \underline{\mu}$ then s_i cannot change the actions of the police and so s_1 must be preferred to s_i . The condition for case (ii) requires that the belief at location i be sufficiently far above this threshold. Note that if $\mu_i = \mu_2$ then this condition is always satisfied. More generally, if the distribution of beliefs is sufficiently flat then this condition will hold for all s_i ($i > 2$).

The intuition for Proposition 24 highlights the two forces which will determine the police's decision to release information to the public or not. By assumption, the agent will always choose s_1 when the police release their information. If the belief at location 1 is sufficiently strong so that $\mu_2 \leq \bar{\mu}$ then this signal is not useful. In this case the police would (at least weakly) prefer to hide their information from the public so as to induce random signal selection. When $\mu_2 > \bar{\mu}$ then s_1 is useful. This opens the possibility that the police may wish to release the information they have. The payoff to random search is diminished by the fact that the agent may choose a signal from a location that has effectively been ruled out ($\mu_i < \underline{\mu}$). If the distribution of beliefs is approximately flat so that s_i is preferred to s_1 for many locations (as a proportion of N) then the police will prefer random search. If however the police's information

has led them to rule out many locations then it may be preferable for the police to release what they know to the public so as to direct search to location 1. Either way, the outcome cannot achieve the unconstrained first best of directing the agent to choose s_1 (as per Proposition 23). We now turn to investigating this trade-off in detail.

3.4 Decentralized Signal Choice and Information Policy

Now return to the constrained model where the signal is chosen by the agent. The agent prefers to choose a the signal which they think is most likely to produce a positive sighting. The police can only influence this signal choice by their decision to release their prior information about the crime or not.

Suppose first that the police chooses to release their information. The agent will then choose signal s_1 . The expected probability that the police will investigate the true state is

$$V^{release} = \left\{ \begin{array}{l} \mu_1 \text{ if } \mu_2 \leq \bar{\mu} \\ \mu_1 p + \mu_2 (1 - q) \text{ if } \mu_2 > \bar{\mu} \end{array} \right\}. \quad (3.2)$$

Next suppose that the police do not release any information so that the agent randomly selects a signal. Let \underline{N} be the number of locations for which $\mu_i \geq \underline{\mu}$. In this case the expected probability that the police will discover the state is

$$V^{hide} = \left\{ \begin{array}{l} \mu_1 \text{ if } \mu_2 \leq \underline{\mu} \\ \frac{1}{\underline{N}} \left[\mu_1 [N - q (\underline{N} - 1)] + p \sum_{i=2}^{\underline{N}} \mu_i \right] \text{ if } \mu_2 \in (\underline{\mu}, \bar{\mu}] \\ \frac{1}{\underline{N}} \left[\mu_1 [p + N - 1 - q (\underline{N} - 1)] + p \sum_{i=2}^{\underline{N}} \mu_i + \mu_2 (1 - q) \right] \text{ if } \mu_2 > \bar{\mu} \end{array} \right\}. \quad (3.3)$$

Comparing these payoffs yields the optimal choice for the police.

Proposition 25 *It is strictly optimal for the police to release their information if and only if*

$$\mu_2 > \bar{\mu} \text{ and } \frac{1}{(N-1)} \sum_{i=2}^{\underline{N}} (\mu_i - \underline{\mu}) < \frac{(1-q)}{p} (\mu_2 - \bar{\mu}) \quad (3.4)$$

Proof. See Appendix. ■

If $\mu_2 \leq \bar{\mu}$ then μ_1 is sufficiently strong so that observing s_1 will not alter the actions of the

police. In this case, releasing the police's information will simply induce the agent to select a worthless signal. Random search can do no worse than this. If $\mu_2 \in (\underline{\mu}, \bar{\mu}]$ then with probability $\frac{N-N-1}{N}$ the agent will draw a signal from location $\{2, 3, \dots, N\}$. Such a signal can potentially change the action of the police (if a positive sighting is made) and thus is valuable.

When $\mu_2 > \bar{\mu}$ an additional force is brought into play. Now, drawing a signal at location 1 is useful because a negative realization of the signal will induce the police to investigate the second location. As a result there may be some benefit to releasing the police's information so as to induce the agent to investigate at this location. This is particularly useful if there are many locations for which $\mu_i < \underline{\mu}$. Signals drawn from these locations have no value since the police's prior belief is sufficiently low that no one realization of the signal will cause the police to investigate them. To understand condition (3.4) suppose that the police's belief is relatively flat so that $\mu_2 = \mu_3 = \dots = \mu_N$. The condition then becomes

$$\mu_1 (1 - p - q) < \mu_2 (1 - p - q)$$

which cannot hold by our assumption that $p + q < 1$. In the case of a flat prior hiding information is desirable since an observation at any location can alter the actions on the police. In contrast suppose that N is large but that $\underline{N} = 2$ so that the principal is effectively choosing between two locations. In this case the condition becomes

$$p (\mu_2 - \underline{\mu}) < (N - 1) (1 - q) (\mu_2 - \bar{\mu})$$

If N is sufficiently large then this condition will hold. In this case random search is not desirable since with probability $\frac{N-2}{N}$ it induces the agent to draw a signal which is of no use to the principal.

3.4.1 Discussion

Proposition 25 allow us to think about when the police would optimally decide to reveal their information. When they have not learned very much (flat prior) then randomized search is desirable and hence the police will hide their information. This may for example be early in the investigation of a crime. When the police has information which rules out many locations

and leaves them to choose between a number of likely suspects then it is optimal to release their information. It is not optimal for the police to release information when they have only one main suspect (when $\mu_2 \leq \bar{\mu}$). Encouraging more search in this location is not useful since it cannot overturn their already strong prior. In this case it is better to hide their information and to allow random search on the chance that the public see something in a location that has not already been ruled out.

3.5 Conclusion

This paper has presented a simple model in which to think about the forces governing the decision of an intelligence gathering agency to release or hide information from the public. There are a number of important extensions which we will pursue in future work. Often the police influence the actions of the public through the use of rewards schemes. If we place reasonable restrictions upon the reward schemes that the police can offer (e.g. only pay positive correct sightings) then of itself these may be unable to avoid the problem of the agent looking in the most likely spot. Again the police will have to be strategic in their decision to release information to the public.

3.6 Appendix

3.6.1 Proof of Proposition 23

Suppose that $\mu_2 \leq \underline{\mu}$. In this case the prior of the police is so strong that no matter what signal is selected the police will investigate location 1. In this case the choice of signal is irrelevant and hence s_2 is (trivially) optimal.

Suppose next that $\mu_2 \in [\underline{\mu}, \bar{\mu}]$. In this case the belief at location 1 is sufficiently strong so that observing s_1 could not alter the action of the police (and hence has no value). However if $s_2 = 1$ is observed this will lead the police to investigate location 2. It follows that s_2 must be preferred to s_1 . Next, to compare s_2 to any s_i for $i \geq 3$. If $\mu_i \leq \underline{\mu}$ then observing s_i cannot alter the action of the police and hence does not have any value. If however $\mu_i > \underline{\mu}$ then s_i

will alter the actions of the police. In this case, the expected payoff from using each signal is

$$\begin{aligned} V(s_2) &= \mu_1(1-q) + \mu_2 p \\ V(s_i) &= \mu_1(1-q) + \mu_i p \end{aligned}$$

and clearly $V(s_2) \geq V(s_i)$. Hence s_2 is optimal whenever $\mu_2 \in [\underline{\mu}, \bar{\mu}]$.

Finally suppose that $\mu_2 > \bar{\mu}$. In this case observing s_1 will alter the actions of the police. Comparing the expected payoff to using s_1 and s_2 we have

$$\begin{aligned} V(s_1) &= \mu_1 p + \mu_2(1-q) \\ V(s_2) &= \mu_1(1-q) + \mu_2 p. \end{aligned}$$

Observe that

$$V(s_2) - V(s_1) = (\mu_1 - \mu_2)(1-q-p) \geq 0$$

and hence s_2 is preferred to s_1 . Next to compare s_2 to any s_i for $i \geq 3$. This analysis is exactly the same as the case where $\mu_2 \in [\underline{\mu}, \bar{\mu}]$. Hence we have that s_2 is optimal. This establishes the Proposition.

3.6.2 Proof of Proposition 24

Observe first that if $\mu_2 \leq \bar{\mu}$ then s_1 cannot change the actions of the police. It follows that s_i must be (weakly) preferred.

Suppose that $\mu_2 > \bar{\mu}$ then observing s_1 can alter the actions of the police. If $\mu_i \leq \underline{\mu}$ then s_i cannot alter the actions of the police and hence cannot be preferred to s_1 . If $\mu_i > \underline{\mu}$ then s_i is also a useful signal. The expected probability of success associated with each signal is

$$\begin{aligned} V(s_1) &= \mu_1 p + \mu_2(1-q) \\ V(s_i) &= \mu_1(1-q) + \mu_i p. \end{aligned}$$

Requiring that $V(s_i) \geq V(s_1)$ yields the expression given in the text. Note that this condition implies that $\mu_i > \underline{\mu}$. This establishes the Proposition.

3.6.3 Proof of Proposition 25

If $\mu_2 \leq \bar{\mu}$ then s_1 is not a useful. In this case if the police release their information they will not receive any useful information from the agent. It follows that hiding information must be weakly preferred.

Suppose first $\mu_2 \leq \underline{\mu}$. Then $V^{hide} = V^{release} = \mu_1$ and hence it is not strictly optimal to release information.

Next suppose that $\mu_2 \in (\underline{\mu}, \bar{\mu}]$. In this case

$$V^{release} - V^{hide} = \mu_1 - \frac{1}{N} \left[\mu_1 [N - q(N - 1)] + p \sum_{i=2}^N \mu_i \right].$$

Upon rearranging we have that

$$V^{release} - V^{hide} = \underline{\mu} - \frac{1}{(N - 1)} \sum_{i=2}^N \mu^i$$

which must be negative since by definition $\mu^i > \underline{\mu}$ for each $i \leq N$. Note also that $N \geq 2$ since $\mu_2 > \underline{\mu}$. Hence it cannot be strictly optimal to release information whenever $\mu_2 \in (\underline{\mu}, \bar{\mu}]$.

Finally $\mu_2 > \bar{\mu}$ then we have that

$$\begin{aligned} V^{release} - V^{hide} &= -\mu_1 p - \mu_2 (1 - q) \\ &+ \frac{1}{N} \left[\mu_1 [p + N - 1 - q(N - 1)] + p \sum_{i=2}^N \mu_i + \mu_2 (1 - q) \right]. \end{aligned}$$

Rearranging this expression and requiring that $V^{hide} - V^{release} > 0$ yields the condition in the Proposition. This establishes the Proposition.

Bibliography

- [1] Kershaw, Sarah (2002), "Those Seeing White Vans Breathe Easier: So Do the Drivers", *New York Times*, October 26.