LIBRARY
Michigan State
University

This is to certify that the
dissertation entitled

ALLOWING OTHERS TO LEARN: ESSAYS ON BAYESIAN UPDATING IN ENVIRONMENTS OF ASYMMETRIC INFORMATION

presented by

Vinit K. Jagdish

has been accepted towards fulfillment
of the requirements for the

Ph.D. degree in Economics

Carl Davidson
Major Professor's Signature

7/22/05

Date

MSU is an Affirmative Action/Equal Opportunity Institution

PLACE IN RETURN BOX to remove this checkout from your record.
TO AVOID FINES return on or before date due.
MAY BE RECALLED with earlier due date if requested.

DATE DUE	DATE DUE	DATE DUE

2/05 c:/CIRC/DateDue.indd-p.15

ALLOWING OTHERS TO LEARN: ESSAYS ON BAYESIAN UPDATING IN ENVIRONMENTS OF ASYMMETRIC INFORMATION

By

Vinit K. Jagdish

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Economics

2005

ABSTRACT

ALLOWING OTHERS TO LEARN: ESSAYS ON BAYESIAN UPDATING IN ENVIRONMENTS OF ASYMMETRIC INFORMATION

By

Vinit K. Jagdish

Allowing Others to Learn: Essays on Bayesian Updating in Environments of Asymmetric Information is composed of three essays: "A Signaling Theory of Managerial Turnover," "Managerial Career Concerns and Termination as a Screening Device," and "A Countersignaling Theory of Advertising and Fads." The three essays have one common thread. In each essay, agents' ability levels are private information. High ability agents signal their private information by placing themselves in positions where Bayesian updating on ability can occur. Thus, in each model, markets disseminate information on agents' abilities in two ways: through signaling and through Bayesian updating. This simple framework provides valuable insights into managerial turnover, managerial termination, and advertising.

"A Signaling Theory of Managerial Turnover" extends the career concerns literature by showing how career concerns can lead to managerial turnover. A model of turnover based on team production and asymmetric information is constructed. Team production makes it difficult, if not impossible, for market participants to learn a manager's individual ability level. Turnover provides an opportunity to further learn managerial ability. As such, high ability managers signal their ability levels by engaging in turnover. The probability of managerial turnover is shown to be decreasing in the cost of turnover, increasing in the variance of firms' perceived abilities, and increasing in the variance of a

manager's perceived ability. The predictions of the model are consistent with recent empirical work on managerial turnover.

"Career Concerns and Termination as a Screening Device" contributes to the literature on optimal contracting in the presence of managerial career concerns by examining the role of termination in contracting. A model of project choice and asymmetric information is constructed. In the absence of contracting, lower ability managers take on excessive risk in their choice of projects. Termination serves as a simple screening device that ensures efficient investment. The model explains the use of the nonlinear termination schedules found in the mutual fund, securities analysis, and hedge fund industries and can explain puzzling stylized facts found in recent empirical work on these industries.

"A Countersignaling Theory of Advertising and Fads" extends the literature on countersignaling by developing a countersignaling theory of advertising. Consumers can learn a firm's innate ability to produce quality in two ways: (1) a firm can signal its ability to produce quality by engaging in costly advertising and (2) consumers can receive information on a firm's quality level through word-of-mouth communication. There exists a countersignaling equilibrium in which the highest and lowest ability firms refrain from advertising while average ability firms advertise. Two testable implications emerge from the analysis: the probability of advertising is not monotonically increasing in advertising and the probability of advertising increases as the cost to advertise increases. A simple extension of the basic model shows how producers fuel and crush fads through their advertising decisions.

To Lisa, Seph, Skittles and Xanadu

Thank you for your love and support

ACKNOWLEDGMENTS

I thank Dale Belman, Jay Pil Choi, Shin Dong Jeung and Lawrence Martin for helpful comments and suggestions. I am especially indebted to Carl Davidson and Thomas Jeitschko for their guidance and advice. The present work could not have been completed without their efforts.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER ONE

A SIGNALING THEORY OF MANAGERIAL TURNOVER

Section I: Introduction

Since the pioneering work of Fama (1980) and Holmstrom (1999), numerous papers have examined how career concerns affect managerial behavior. Career concerns have been shown to alleviate effort-aversion moral hazard, cause managers to hide or ignore information, make young managers overreact to new information, lead managers to delay the divestiture of unproductive assets and induce managers to overinvest in information.[1] The present work extends the literature on implicit incentives by illustrating how career concerns, or concerns over perceived ability, can lead to managerial turnover. It constructs and analyzes a model of turnover based on team production and asymmetric information.

To get at the heart of the model, consider the following scenario: Manager A teams with Firm B to produce output. Output depends on the innate abilities of the manager and the firm. There is incomplete but symmetric information on Firm B's ability level. Manager A, however, has private information. She knows her own ability level; the market only has a common prior. The market observes the joint output of the team but it cannot observe the individual ability levels of A and B. The observation allows the market to update its beliefs on the collective ability of the team as well as the individual ability levels of A and B. Since the market only observes team output, Manager A's perceived ability is linked to Firm B's ability, and vice versa.

[1] See the work of Fama (1980), Holmstrom and Ricart i Costra (1986), Scharfstein and Stein (1990), Boot (1993), Prendergast and Stole (1996), Holmstrom (1999) and Milbourn, Shockley and Thakor (2001).

Suppose Manager A has a high ability level and Firm B has a low ability level. If A and B were to continue their union, A's perceived ability would be dragged down by the low ability of B. Manager A would be undervalued in the market. Manager A, of course, is not forced to work with Firm B. She has other options. A could team with another firm, Firm C. The new output of A and C would be observed by the market. The observation would allow the market to learn about the collective ability of A and C. It would also give the market a better estimate of Manager A's individual ability. While the individual contribution of Manager A would never be observed, it would be a common element in the team output of A and B and the team output of A and C. Thus, the two team outputs could be viewed as noisy signals of Manager A's individual ability. Turnover has provided the market with an opportunity to better learn Manager A's level of ability.

Manager A's turnover is driven by two factors. The first factor driving turnover is the assumption of team production. With the requirement of team production and only team output being observable, the market can never fully learn a manager's ability level when she chooses to remain with her firm. Only collective ability can be fully learned since the market cannot tell what share of output the manager (or the firm) is responsible for producing. Remaining with the same firm, then, offers a manager a sanctuary from statistical inference. In order for the market to update its belief on Manager A's individual ability, Manager A must place herself in a new environment. That is, she must leave her firm, team up with a new firm and produce a new level of output for the market to further learn her ability level.[2]

---

[2] Grossman and Maggi (2000) analyze a model where workers choose between working in an industry with imperfect observability of talent or an industry where compensation is tied heavily to individual talent. In their model, talented workers are drawn to industries that reward individual success. A similar result would arise in the present model if self-employment were allowed.

The second factor driving turnover is asymmetric information. In the example above, turnover is used to let the market learn more about Manager A's ability level. If A's ability level were known to all, there would be no need for turnover. If A's ability were unknown to all, including A, there would still be no need for turnover. This stems from the well-known fact that Manager A's expected perceived ability after turnover would be a martingale with respect to A's expected perceived ability prior to turnover (see Holmstrom, 1999). The martingale property, however, does not hold when managers have private information about their ability levels. Managers with higher ability levels have an incentive to let the market learn their ability. Since turnover allows the market to learn, high ability managers will engage in turnover. Thus, the act of turnover itself signals high ability.

This paper formally analyzes the scenario above. Given past performance, the existence of a unique separating equilibrium in which high ability managers engage in turnover and low ability managers remain with their firms is demonstrated. The probability of managerial turnover is shown to be decreasing in the cost of turnover, increasing in the variance of firms' perceived abilities, and increasing in the variance of a manager's perceived ability. The predictions of the model are consistent with the empirical patterns of managerial turnover in Fee and Hadlock (2003).

In addition to extending the work on implicit incentives, the present work is related to four other strands of the literature. First, it is related to the large body of work that exists on labor and managerial turnover. It is closely related to ability-matching and on-the-job search theories of labor turnover. In ability-matching models, workers are matched with firms and learn about the quality of their match over time. Turnover occurs when workers find they are poorly matched. Johnson (1978) and Jovanovic (1979) provide early examples of this type of model. In on-the-job search models, workers learn about outside job opportunities. Turnover occurs when workers find a more attractive opportunity elsewhere. See Burdett (1978) for an early example. In both types of models, *learning leads to turnover*. In the current work, turnover allows the market to revise its beliefs on a manager's ability level. Here, *turnover leads to learning.*

The present model is also closely related to structural models of managerial turnover. Structural models of turnover focus on the hierarchical management structure of firms as a cause of turnover. In Rosen (1982), for example, there are a limited number of slots available in top management positions. Turnover can occur because managers might be forced to switch firms in order to find open slots in top management. Slot limitations also drive turnover in the "up or out" models of O'Flaherty and Siow (1992) and Demougin and Siow (1994). Turnover models based on slot limitations implicitly assume that a manager's objective is to advance as far as possible in her current firm's hierarchy. Turnover only occurs when a manager's advancement is blocked or denied. The current work also provides a structural model of managerial turnover. The structure leading to turnover in the present work, however, differs from existing models. Here, turnover does not occur because a manager's path in a hierarchy is blocked. In fact, a hierarchical management structure is not modeled. Managers switch firms in the present model when firms can no longer provide managers with an opportunity to let the market further learn their ability.

Next, the current model is related to the literature on team production, task assignment, and job rotation. The model is similar in spirit to Meyer (1994). In Meyer

(1994), production takes place in teams composed of senior and junior members of a firm. As in the present model, only team output is observable. The firm decides how to pair off seniors and juniors in order to learn its employees ability levels as precisely as possible. The present work is also concerned with learning in a team production environment. The main difference here is that managers initiate the learning process by moving across firms rather than waiting to be allocated to different teams within a firm.

An extension of the model has implications for the sorting of managers and firms and relative performance evaluation. Becker (1981) provides early models on sorting. See Lazear and Rosen (1981), Holmstrom (1982) and Nalebuff and Stiglitz (1984) for the usefulness of relative performance evaluation when there is a common element of uncertainty affecting all workers' performances.

The rest of the paper is organized as follows: Section II develops the model. Section III provides equilibrium analysis. A separating equilibrium in which high ability managers engage in turnover is shown to exist. Factors affecting the likelihood of managerial turnover, such as managers' histories, labor market uncertainty on managers' and firms' ability levels, and turnover costs, are also examined. Section IV discusses the results of the model and shows their empirical relevance. Section V concludes.

Section II: The Model

Agents, Production and Turnover

The basic model considers a world in which two types of agents, managers and firms, team up to jointly produce output. Production can only take place if one manager and one firm are present. Output, y, is produced by a linear additive technology and is given by

$$y = (m_i + f_j)\, I(\cdot) \qquad (1)$$

where m_i is the innate ability of manager i, f_j is the innate ability of firm j, and I(·) is an indicator function that equals unity if exactly one manager and one firm are involved in the production process and is zero otherwise.[3] The market can observe output, y, but it cannot observe managerial and firm ability, m_i and f_j. The price of output is normalized to one.

A firm's ability, f_j, is symmetrically unknown but f_j is commonly believed to be an independent draw from a uniform distribution on the interval [c, c + d]. Managers possess private information; manager i knows her own ability, m_i. The market believes m_i is an independent draw from a uniform distribution on the interval [a, a + b]. It is assumed that $b < d$ [4] and a, b, c and d are nonnegative.

After teaming up to produce output, manager i and firm j can choose to dissolve their union. A union is broken if at least one of its participants chooses to leave. After a union is dissolved, a firm can either form a new union with a different manager or work in an alternative industry. If a firm chooses to work in an alternative industry, it receives a wage equal to its perceived ability. A firm incurs no cost when its union dissolves. A manager, on the other hand, incurs a cost of $z > 0$, if her union breaks up. The cost z captures the adjustment cost of having to go back on the market to find a new partner.[5] After a break up, a manager forms a new union with a different firm.[6]

All agents are risk neutral, do not discount the future and are wealth constrained.

---

[3] The qualitative results of the model can still be established with more general forms of the production function.

[4] A relationship between b and d must be assumed for the purposes of Bayesian updating with uniform distributions. The qualitative results of the model would not change if $b > d$ were assumed.

[5] The qualitative results of the model would not change if both firms and managers had to pay adjustment costs.

[6] The manager's participation constraint is assumed to always be satisfied. Providing managers with an attractive outside alternative does not qualitatively affect the results as long as the outside option is independent of managerial ability.

Timing of Events

At time, $t = 0$, nature determines the abilities of firms and managers. At time, $t = 1$, managers and firms meet in the labor market. Managers and firms that form teams agree on how to divide the output they will produce. Output is then produced and observed by the market. After production, output is divided between a firm and manager according to their agreement. At time, $t = 2$, agents decide whether or not to leave their partners. At time, $t = 3$, agents without partners meet in the labor market. The game at $t = 3$ is exactly the same as the game at $t = 1$. After output is divided, the game ends.

Output Agreements

Firms compete to hire managers. Managers are assumed to be on the short side of the market, so competition among firms will drive down a firm's expected payoff to its perceived ability (the value of its outside option). Note that before production takes place at $t = 1$, a firm's ability is believed to be uniformly distributed on the interval $[c, c + d]$. Thus, the value of the firm's outside opportunity at this time is equal to $(2c + d)/2$. After production takes place at $t = 1$, the market can use the output of a firm's team to update its beliefs on a firm's ability. Since a firm's ability is symmetrically unknown the posterior mean on a firm's ability after output is observed is a martingale with respect to the prior mean (Holmstrom, 1999). Ex ante, a firm would be indifferent between receiving a payoff of $(2c + d)/2$ and receiving a payoff equal to the mean of its posterior on ability after output is observed. Accordingly, firms and managers will agree to split the output they produce along the lines of perceived ability for the duration of their union.

That is, each agent's output share will equal the posterior mean of the agent's perceived ability after output is produced.[7]

Market Beliefs

The market uses Bayes' Rule to update its beliefs on firm and managerial ability. Observations on output at t = 1 and t = 3, potentially allow the market to learn about agents' abilities.

The turnover decisions of managers at t = 2, also convey information to the market. Consider the situation of a manager at t = 2: A manager can either stay with her firm or find a new one. From the production function given in (1), if a manager stays with her firm, her team will produce the same amount of output at t = 3 as it did at t = 1.[8] Hence, no learning can take place from observing output at t = 3. Now, consider a manager who switches firms. By switching firms, a manager forms a new team that produces a different level of output than her previous team. The new level of output provides the market with an additional observation on managerial ability.

Since turnover allows the market to learn more about managerial ability, it is reasonable to expect managers of higher ability to switch firms. Accordingly, this paper will examine the market belief that, for a given level of output realized at t = 1, managers with ability greater than or equal to a cutoff value, m`, will switch firms and those with ability level less than m` will remain with their firms. For shorthand, this belief will be referred to simply as market belief, m`.

[7] One might wonder why firms do not screen managers by offering them a menu of contracts from which to choose. Screening would greatly complicate the analysis that follows. As long as firms do not screen managers so that managerial ability is completely learned, the signaling described in the model will occur.
[8] Output for any given team is deterministic. The implications of relaxing this assumption will be discussed later.

Agents' Objectives

All agents seek to maximize the sum of their earnings. The only decision an agent makes is the turnover decision at t = 2. Thus, the agent's objective reduces to making the turnover decision at t = 2, that leads to the greater expected wage at t = 3. Since a firm's ability is symmetrically unknown the posterior mean on a firm's ability after output is observed is a martingale with respect to the prior mean. This implies a firm will be indifferent between staying with its manager and leaving her. As a tie-breaking convention, it is assumed that firms always choose to remain with their managers. Consequently, the analysis here will focus on managers.

Due to private information, the mean of the posterior on ability is not a martingale with respect to the prior mean for managers. Consider the situation of a manager at t = 2: After producing output at t = 1, the market uses Bayes' Rule to update its beliefs on managerial ability. Let y_t denote the output of the manger's team at time t. The posterior on m after observing y_1[9] is:

$$(m \mid y_1) \sim \begin{cases} U[a, y_1 - c] & if \quad y_1 \in [a+c, a+b+c] \\ U[a, a+b] & if \quad y_1 \in (a+b+c, a+c+d) \\ U[y_1 - (c+d), a+b] & if \quad y_1 \in [a+c+d, a+b+c+d] \end{cases} \tag{2}$$

The analysis presented focuses on the low output case of $y_1 \in [a + c, a + b + c]$.[10]

If a manager chooses to remain with her firm, $y_3 = y_1$. In this case, output produced at t = 3 will be useless for making inferences on her ability. However, the fact the manager remained with her firm will convey information to the market. Given the market belief

[9] The posterior can be obtained using procedures in DeGroot (1970).

[10] Analyses of the other two cases are similar to the analysis of the first case. Results from these two cases are provided in the appendix.

that managers with ability, m < m`, remain with their firms, the posterior on m, for a manager who stays with her firm, at t = 3 is[11]

$$(m \mid y_1, \text{no turnover}) \sim U\,[a, m`] \qquad (3)$$

When a manager remains with her firm, the market belief on her ability is right-truncated.

If a manager chooses to leave her firm, the market can update its beliefs on managerial ability based on y_1, the manager's turnover decision, and y_3. Given the market belief that managers with ability, m ≥m`, leave their firms, the posterior on m, for a manager who leaves her firm, at t = 2 is

$$(m \mid y_1, \text{turnover}) \sim U\,[m`, y_1 - c] \qquad (4)$$

When a manager switches firms, the market belief on her ability is left-truncated.

The posterior on m after y_3 is observed is now considered. When a manager switches firms, it is assumed that she teams up with a new firm that has no market history.[12] That is, the new firm's ability is believed to be uniformly distributed on the interval [c, c + d]. Given that a manager leaves her firm at t = 2, the posterior on m after y_3 is observed is[13]

$$(m \mid y_1, \text{turnover}, y_3) \sim \begin{cases} U[a, y_3 - c] & \text{if} \quad y_3 \in [a + c, m'+c) \\ U[m', y_3 - c] & \text{if} \quad y_3 \in [m'+c, y_1] \\ U[m', y_1 - c] & \text{if} \quad y_3 \in (y_1, m'+c+d) \\ U[y_3 - (c+d), y_1 - c] & \text{if} \quad y_3 \in [m'+c+d, y_1 + d] \end{cases} \qquad (5)$$

(5) shows that after y_3 is observed, the reputational states that can occur fall into four different regions. Compared to (4), the first two regions (lower levels of output) lead to downward revisions of managerial ability. The third region does not change the belief on managerial ability. The fourth region (higher levels of output) leads to an upward revision

---

[11] Technically, this is the limiting distribution.

[12] The implication of relaxing this assumption will be discussed in Section IV.

[13] Given the market belief, m`, the first case would be an off-equilibrium event. In this case, it is assumed the market does not use m` in determining the posterior belief on managerial ability as m` is proven to be incorrect. Only y_1 and y_3 are used in determining the posterior.

of managerial ability. Managers with ability, m < m`, can only achieve reputational states in the first three regions. Managers with ability, m ≥m`, can only achieve reputational states in the last three regions.

For a manager making a turnover decision at t = 2, y_3 is a random variable with a uniform distribution on the support [m + c, m + c + d]. Let E_T denote the expected mean of (m | y_1, turnover, y_3). Using (5) and the mean of a uniform distribution, E_T is

$$E_T = \begin{cases} \frac{1}{d}\left[\int_{m+c}^{m'+c} \frac{a+y_3-c}{2}dy_3 + \int_{m'+c}^{y_1} \frac{m'+y_3-c}{2}dy_3 + \int_{y_1}^{m+c+d} \frac{m'+y_1-c}{2}dy_3\right] & \text{when } m < m' \\ \frac{1}{d}\left[\int_{m+c}^{y_1} \frac{m'+y_3-c}{2}dy_3 + \int_{y_1}^{m'+c+d} \frac{m'+y_1-c}{2}dy_3 + \int_{m'+c+d}^{m+c+d} \frac{y_3+y_1-2c-d}{2}dy_3\right] & \text{when } m \geq m' \end{cases} \tag{6}$$

From (3), the expected ability of a manager who remains with her firm is (a + m`)/2. Carrying out the integration in (6) and then subtracting the expected ability of a manager who remains with her firm as well as z, the cost of turnover, yields the expected return of turnover. The expected return of turnover, V(m, m`, y_1, z) is

$$V = \begin{cases} \frac{1}{d}\left[m\left(\frac{y_1-c-a+m'}{2}-\frac{m}{4}\right)+d\left(\frac{y_1-c+m'}{2}\right)-m'\left(\frac{m'-a}{2}\right) - \frac{(y_1-c)^2}{4}\right]-\frac{a+m'}{2}-z & \text{when } m < m' \\ \frac{1}{d}\left[m\left(\frac{y_1-c-m'}{2}\right)+d\left(\frac{y_1-c+m'}{2}\right)+\frac{(m'+y_1-c)(m'-y_1+c)}{4}\right] - \frac{a+m'}{2}-z & \text{when } m \geq m' \end{cases} \tag{7}$$

Section III: Equilibrium Analysis

A manager will leave her firm at $t = 2$ if the expected return of turnover is greater than or equal to zero and will stay with her firm otherwise.[14] The following lemma will help characterize a manager's turnover strategy:

Lemma 1: Given a first period level of output, y_1, and market belief, $\grave{m}$, the expected return of turnover is increasing in managerial ability, $\partial V/\partial m > 0$.

Proof: See Appendix A.

Lemma 1 shows the expected return of turnover in increasing in managerial ability. The intuition behind Lemma 1 is straightforward. An increase in managerial ability increases the probability that the market will revise its belief on managerial ability upwards and allows higher reputational states to be achieved. Lemma 1 implies that, given a value of y_1, $\grave{m}$, and z, there exists a cutoff value of managerial ability, $\tilde{m}$, such that managers with ability, $m \geq \tilde{m}$, leave their firms and managers with ability, $m < \tilde{m}$, remain with their firms. For shorthand, this turnover strategy will be referred to simply as turnover strategy, $\tilde{m}$. Specifically, let $\tilde{m}(\cdot)$, be defined such that $V(m = \tilde{m}, \grave{m}, y_1, z) = 0$.

An equilibrium in this model is a manager's turnover strategy, $\tilde{m}$, and market belief, $\grave{m}$, such that: (1) the manager's turnover strategy, $\tilde{m}$, is optimal given the market belief, $\grave{m}$, and (2) the market belief, $\grave{m}$, coincides with the manager's turnover strategy, $\tilde{m}$.

The following proposition shows existence of a unique separating equilibrium.[15]

---

[14] As a tie-breaking convention, it is assumed that a manager engages in turnover if she is indifferent between moving to a new firm and remaining with her old one.

[15] There also exists a pooling equilibrium such that, for $(a+c) \leq y_1 \leq (a+b+c)$ and $z > (y_1 - c - a)^2/4d$, all managers remain with their firms. This equilibrium is supported by the following reasonable off-equilibrium belief: For any manager who engages in turnover the market updates its belief on managerial ability using Bayes' Rule and observations on y_1 and y_3.

Proposition 1: Given a realization of y_1 such that, $(a + c) \leq y_1 \leq (a + b + c)$, and given $z \in [\frac{y1-c-a}{2} - \frac{(y1-c-a)^2}{4d}, \frac{y1-c-a}{2}]$, there exists a unique separating equilibrium such that managers with ability, $m \geq m^* = (y1-c)-\sqrt{2d(y1-c-a-2z)}$, leave their firms and managers with ability, $m < m^*$, remain with their firms.

Proof: See Appendix A.

The equilibrium cutoff value for turnover, m*, can be used to construct a probability of turnover. Recall from (2) that the posterior on m after y_1 is observed is $(m \mid y_1) \sim U[a, y_1 - c]$. Managers with ability, $m \geq m^*$, will leave their firms. Thus the probability of turnover, P, is

$$P = \frac{y1-c-m^*}{y1-c-a} \tag{8}$$

The probability of turnover depends on first period output, turnover costs, uncertainty on firm ability and uncertainty on managerial ability. The next proposition shows how these variables affect the probability of turnover.

Proposition 2: Given the unique separating equilibrium in Proposition 1, the probability of turnover is: (i) decreasing in the cost of turnover, $\partial P/\partial z < 0$, (ii) increasing in the variance of a firm's perceived ability, $\partial P/\partial d > 0$, and (iii) increasing in the variance of perceived managerial ability, $\partial P/\partial a < 0$, $\partial P/\partial c < 0$ and $\partial P/\partial y_1 > 0$.

Proof: See Appendix A.

Proposition 2 states the probability of turnover is decreasing in z, the cost of turnover. The intuition behind this result is straightforward. An increase in the cost of turnover lowers the expected return of turnover and therefore reduces the number of managers who will find turnover profitable.

The probability of turnover is increasing in d, which is proportional to the variance of perceived firm ability. As the variance of perceived firm ability increases, output at $t = 3$, y_3, becomes less informative for making inferences on managerial ability. This allows managers with lower ability levels to switch firms and receive the signaling benefits of turnover without having to worry as much that y_3 will reveal their low levels of ability.

From (2), the market belief on m after y_1 is observed is $(m \mid y_1) \sim U\ [a, y_1 - c]$. Decreasing a or c, or increasing y_1, increases the variance of perceived managerial ability at $t = 2$ and hence, from Proposition 2, the probability of turnover is increasing in the variance of perceived managerial ability.

To better understand this last result, the expected return of turnover (ignoring z) given in (7) is decomposed into two effects: the net learning value of turnover and the net signaling value of turnover. The analysis will focus solely on the comparative statics result on y_1. Special attention is paid to this result as it is often examined in empirical studies. Analyses of the comparative statics results on a and c are similar to that of y_1.

The net learning value of turnover, LV, is defined as the expected return of turnover if managerial ability were unknown to all. From (2), $(m \mid y_1) \sim U\ [a, y_1 - c]$. The net learning value of turnover depends on the distribution of $(m \mid y_1, y_3)$. Ignoring the effects of signaling and updating solely on output yields:

$$(m \mid y_1, y_3) \sim \begin{cases} U[a, y_3 - c] & if \quad y_3 \in [a + c, y_1] \\ U[a, y_1 - c] & if \quad y_3 \in (y_1, a + c + d) \\ U[y_3 - (c + d), y_1 - c] & if \quad y_3 \in [a + c + d, y_1 + d] \end{cases} \tag{9}$$

Using (9) and the mean of the uniform distribution yields the gross learning value of turnover:

$$\frac{1}{d}[\int_{m+c}^{y_1} \frac{a + y_3 - c}{2} dy_3 + \int_{y_1}^{a+c+d} \frac{a + y_1 - c}{2} dy_3 + \int_{a+c+d}^{m+c+d} \frac{y_3 + y_1 - 2c - d}{2} dy_3] \tag{10}$$

The value of a manager who stays with her firm (ignoring the effects of signaling) is (from (2)) $(a + y_1 - c)/2$. Carrying out the integration in (10) and subtracting off the value of a manager who stays with her firm yields the net learning value of turnover:

$$LV = \frac{1}{d}[m(\frac{y1 - c - a}{2}) + \frac{(a + y1 - c)(a - y1 + c)}{4}] \tag{11}$$

Differentiating (11) with respect to y_1 yields:

$$\frac{\partial LV}{\partial y1} = -\frac{1}{2}(\frac{y_1 - c - m}{d}) \tag{12}$$

The net learning value of turnover is decreasing in y_1. The logic behind this result is simple: A higher value of y_1 indicates a better partner for a manager at $t = 1$. The better the partner a manager has, the less likely she is to gain from switching partners.

To obtain a more complete understanding of the comparative statics result in (12), the reputational states in (9) are reexamined. The reputational states that can arise fall into three different regions. An increase in y_1 increases the posterior mean of m if y_3 is an element of either of the last two regions. An increase in y_1, however, also (from (2)) increases the value of a manager who stays with her firm. These two effects on the net learning value of turnover exactly offset one another. If y_3, however, is an element of the first region, $y_3 \in [a + c, y_1]$, an increase in y_1 has no effect on the posterior mean of m.

The increase in the value of a manager who stays with her firm is not offset and thus, the net learning value of turnover decreases. This case occurs with probability $(y_1 - c - m)/d$.

In addition to Bayesian updating from another observation on output, the market uses the manager's turnover decision to revise its belief on managerial ability. The net signaling value of turnover, SV, is defined as the expected return of turnover (ignoring z) minus the net learning value of turnover. Subtracting (11) from (7) yields:

$$SV = \frac{1}{d}[m(\frac{a-m'}{2}) + d(\frac{m'-a}{2}) + \frac{m'^2}{4} - \frac{a^2}{4}] - (\frac{m'}{2} - \frac{(y1-c)}{2}) \qquad (13)$$

Differentiating (13) with respect to y_1 yields:

$$\frac{\partial SV}{\partial y1} = \frac{1}{2} > 0 \qquad (14)$$

The net signaling value of turnover is increasing in y_1. Recall from (4) that belief on m for a manager who leaves her firm at t = 2 is $(m \mid y_1, \text{turnover}) \sim U\,[m`, y_1 - c]$. An increase in y_1 increases $E\,(m \mid y_1, \text{turnover})$ and therefore makes signaling more attractive.

Overall, the expected return of turnover, V, is increasing in y_1. This occurs because the signaling effect dominates the learning effect. An increase in y_1 has the potential to decrease the net learning value of turnover by the same magnitude as it increases the net signaling value of turnover but the two effects differ in strength because the signaling effect happens with certainty whereas the learning effect only occurs with probability $(y_1 - c - m)/d$.

The signaling effect dominates the learning effect in the comparative statics results on a and c as well. Thus, the key to understanding these comparative statics results lies in examining how turnover (signaling) truncates the distribution of perceived managerial ability. From (2), the market belief on m after y_1 is observed is $(m \mid y_1) \sim U\ [a, y_1 - c]$. From (3), when a manager remains with her firm the market belief on m is right-truncated: $(m \mid y_1, \text{no turnover}) \sim U\ [a, m']$. From (4) the market belief on m for a manager who engages in turnover is left-truncated: $(m \mid y_1, \text{turnover}) \sim U\ [m', y_1 - c]$. Decreasing a or c, or increasing y_1 makes turnover more attractive than remaining with one's firm. Hence, an increase in the variance of perceived managerial ability at $t = 2$, increases the probability of turnover.

It is important to interpret the comparative statics results on a, c, and y_1 collectively as a result on the variance of perceived managerial ability rather than results on the individual variables themselves. The preceding analysis has focused on the low first period output case in (2) where $y_1 \in [a + c, a + b + c]$. In the two other cases, where $y_1 \in (a + b + c, a + c + d)$ or $y_1 \in [a + c + d, a + b + c + d]$, the probability of turnover is increasing in the variance of perceived managerial ability at $t = 2$ as well.[16] However, the individual variables affect the variance of perceived managerial ability differently in these two other cases. Again, special attention is paid to y_1. For intermediate values of first period output, when $y_1 \in (a + b + c, a + c + d)$, an increase in y_1 has no effect on the variance of perceived managerial ability at $t = 2$ and therefore does not affect the probability of turnover. For high levels of first period output, when $y_1 \in [a + c + d, a + b + c + d]$, an increase in y_1 decreases the variance of perceived managerial ability at $t = 2$ and hence, decreases the probability of turnover as well. While increases in the variance of perceived managerial ability increase the probability of turnover, the variance of perceived managerial ability is not monotonically increasing or decreasing in y_1. This

[16] Results are contained in the appendix.

occurs because, ex ante, y_1 is distributed on a finite support. The implications of this result will be discussed in Section IV.

Proposition 2 sheds light on stochastic production. Note in (1) that output produced by a team is deterministic. If a zero mean error term were added to the production function in (1), the probability of turnover will increase for two reasons. First, the added uncertainty in production would lead to a less precise estimate on managerial ability after first period output was observed. That is, the belief on managerial ability given y_1, $(m \mid y_1)$, would have greater variance. Secondly, with added uncertainty in production, if a manager engages in turnover, the output produced by her new team is less informative on managerial ability. Analytically, this is akin to an increase in d, which is proportional to the variance of perceived firm ability. Thus, from Proposition 2, uncertainty in production would increase the probability of turnover.

Section IV: Discussion

It is interesting to compare the current model to the ability-matching and on-the-job search theories of labor turnover of Burdett (1978), Johnson (1978) and Jovanovic (1979). In these models, the arrival of new information causes turnover. That is, *learning leads to turnover*. In the present model, team production guarantees that a manager's ability level cannot be fully learned if she remains with the same firm.[17] Turnover provides the market with another observation on managerial ability and therefore allows the market to revise its beliefs on managerial ability. Thus, in the present work, *turnover leads to learning*. As such, high ability managers engage in turnover to let the market learn their true ability levels.

---

[17] There are two instances when a manager's ability level can be fully learned if she remains with her firm. These are the rare cases where y_1 is an extreme level of output at either the low end or the high end of its support.

The current work and the information based theories of turnover discussed above both predict voluntary turnover should lead to wage increases. The increased wages, however, occur for different reasons. In ability-matching models, workers, on average, should earn higher wages after turnover because of increased productivity resulting from a better match. In on-the-job search models, higher wages are responsible for inducing turnover. In the present model, turnover reveals a manager is of higher ability than her previous work history indicates. Since turnover signals a higher level of ability, managers who switch firms receive higher wages.

Other testable implications of the model come from the comparative statics effects in Proposition 2. The probability of turnover is found to be decreasing in the cost of turnover, increasing in the variance of perceived managerial ability, and increasing in the variance of perceived firm ability. The limited empirical evidence available supports these predictions.

In order to understand the best test of the model it is important to differentiate the present work from existing models of managerial turnover. Most models of managerial turnover focus on the hierarchical management structure of firms as a cause of turnover. These models implicitly assume that a manager's objective is to advance as far as possible in a firm's hierarchy. Turnover only occurs when a manager's advancement is blocked or denied.[18] Here, a hierarchical management structure is not modeled. Managers engage in turnover when firms can no longer provide managers with an opportunity to let the market further learn their ability levels. As such, a true test of the model should control for firms' hierarchical management structures. Turnover due to managers moving up the corporate ladder should not be considered.

[18] See the work of Rosen (1982), O'Flaherty and Siow (1992), and Demougin and Siow (1994).

Fee and Hadlock (2003) empirically investigate the movement of managerial talent across firms. Their work on the lateral moves of managers is of particular interest as it abstracts from the issue of promotion. Their main findings are: (i) there is no statistically significant relationship between a manager's past performance and the probability of turnover, (ii) the probability of turnover increases as the size of a manager's firm increases and (iii) "golden handcuffs" (restricted stock or unvested options, for example) do not affect the probability of turnover. The present model can explain all three results.

As for the first result, recall that after y_1 is observed, the reputational states that can arise fall into the three regions given in (2). For the low output case, it was shown that the probability of turnover is increasing in y_1 (past performance). The results for the cases of intermediate values of output and high values of output are given in the appendix. In the case of intermediate values of y_1, first period output has no effect on the probability of turnover. In the case of high first period output, the probability of turnover is decreasing in y_1. Thus, the probability of turnover is first increasing, then constant, and finally decreasing in y_1. Since the probability of turnover is not monotonically increasing or decreasing in past performance, a statistical test between the two would find no partial correlation.

The key to understanding this first result is to realize that past performance only affects the probability of turnover indirectly through its effect on the variance of perceived managerial ability. For low values of y_1, the variance of perceived managerial ability is increasing in y_1. For intermediate values of y_1, the variance of perceived managerial ability is not effected by y_1. Finally, for high values of first period output, the variance of perceived managerial ability is decreasing in y_1. This occurs because, ex ante,

y_1 is distributed on a finite support. Output levels near the low end or the high end of the support allow the market to make a more precise estimate on managerial ability than intermediate values of output. As seen in the previous section, signaling through turnover becomes more important as the variance of perceived managerial ability increases.

The variance of perceived managerial ability is also vital to explaining why the probability of turnover is increasing in firm size. One can view increased firm size as an expansion of a manager's team. As a manager's team increases in size, the overall output produced by the team is less likely to reflect the individual ability of the manager. Hence, an increase in firm size is analogous to an increase in the variance on perceived managerial ability. As such, the probability of turnover should increase as firm size increases.

The model can also explain why "golden handcuffs" do not affect the probability of turnover. Golden handcuffs might make turnover less attractive, since managers are forfeiting potentially valuable portfolios. However, if a manager does not engage in turnover, the market will not learn her true level of ability. In the long-run, if firms ultimately promote managers to top positions based on ability, then the benefit of signaling high ability through turnover will render the forfeited portfolios insignificant.

Lastly, the implications of relaxing a somewhat restrictive assumption are discussed. It was assumed that a manager who engaged in turnover moved to a new firm with no market history. Given the choice, some managers might prefer to team up with a firm that produced output at $t = 1$.

A manager might want to join a firm with a market history for two reasons. The first reason is that a firm with a market history (relative to a firm with no history) is likely to have less uncertainty on its ability. The firm's past output allows the market to make a more precise estimate on its ability. Undervalued managers (those with high ability given y_1) might want to pair with firms whose perceived abilities have small variances because output produced from these unions will be more likely to reveal true managerial ability.

The second reason a manager might want to switch to a firm with a market history is due to an indirect learning effect. Suppose f_1 is a firm that produced output at t = 1. To produce output at t = 1, f_1 must have been paired with a manager, m_1. Let m_2 be an undervalued manager at t = 2. The undervalued manager might want to team with f_1 since the variance of f_1's perceived ability is likely to be smaller than that of a new firm. There is also an indirect learning effect at work. If m_2 pairs with f_1, f_1's old partner, m_1, must have a new firm. The output produced by m_1 and her new firm potentially yields a better estimate of m_1's ability. This, in turn, allows for more precise estimates to be made on f_1's and m_2's ability levels as well. Thus, an undervalued manager might want to move to a firm with a market history (if available) because it indirectly allows more observations to be made on managerial ability.

An analysis of a market with multiple managers (as few as two) quickly becomes cumbersome. In a working paper, a simple example is presented that ignores the indirect learning effect above. The example is similar to the model presented here with one notable exception: A manager has more than one turnover option.[19] At t = 2, a manager can choose to remain with her firm, switch to a firm whose perceived ability has a relatively large variance, or move to a firm whose perceived ability has a relatively small variance. An equilibrium exists such that, managers who are the most undervalued in the market, at t = 2, choose to move to the small variance firm. Overvalued managers remain

[19] Though tedious, the analysis is a straightforward extension of the current work.

with their firms. Managers with "average" ability conditional on y_1 switch to large variance firms.

Since there is a lower level of uncertainty on its perceived ability, the small variance firm extracts rent from managers. This is vital to establishing separation between managers who switch to the small variance firm and managers who switch to a large variance firm. Managers who move to either type of firm incur the same turnover cost. Rent introduces a cost differential between the two options. The expected return of joining the small variance firm relative to joining a large variance firm is increasing in managerial ability, so only the highest ability managers find it profitable to choose the small variance firm and pay rent.

In practice, a firm's perceived ability is likely to have a small variance for two reasons. The first reason is that a firm might have had a previous output level that was very informative of its ability. That is, a firm might have produced a level of output near either the low end or the high end of the support of y_1. The second reason is that a firm might have a long market history. Many past observations on output would lead to a more precise estimate on firm ability.

Firms that have perceived abilities with small variances due to extreme output levels are of either very low or very high ability. The highest ability managers will flock to these firms, implying that both positive and negative sorting of firms and managers will occur.[20] This result is somewhat sensitive to the linear additive production function given in the model. A high ability manager might be less inclined to switch to a low ability firm if managerial productivity were increasing in firm ability. Nonetheless, it explains that a

[20] See Becker (1981) for early models on sorting.

manager can acquire a good reputation by moving to a high ability firm and keeping performance high or by moving to a low ability firm and raising the firm's performance.

Firms that have perceived abilities with small variances due to long market histories are likely to have teamed with several managers in the past. When a manager teams with a firm that has a long market history, she not only provides the market with another observation on her individual ability, she also allows the market to compare her to the firm's past managers as well. Since turnover provides common benchmarks to compare managers, one would expect to observe relative performance evaluation being more prevalent in industries with high turnover.[21]

Section V: Conclusion

This paper has shown how career concerns can lead to managerial turnover. The assumptions of team production and asymmetric information are vital to the theory. Team production makes it difficult, if not impossible, to learn a manager's individual ability level. Turnover allows the market to further learn a manager's ability. As such, managers who know their own ability level and want to reveal it to the market will engage in turnover. The predictions of the model are supported empirically by Fee and Hadlock (2003).

Extensions of the present model could produce valuable insights into executive retention and turnover. The present model considers a simple contracting environment. It would be interesting to see how the implicit incentives here affect the nature of contracts offered to managers à la Gibbons and Murphy (1992). Also, the relationship between relative performance evaluation and turnover should be empirically explored.

[21] Using the output levels of others in an industry to obtain a more precise estimate on an individual's ability level is explored in Lazear and Rosen (1981), Holmstrom (1982), and Nalebuff and Stiglitz (1984).

CHAPTER TWO

MANAGERIAL CAREER CONCERNS AND TERMINATION AS A SCREENING DEVICE

Section I: Introduction

Recent empirical work on career conscious mutual fund managers, security analysts and hedge fund managers has shown that the use of termination schedules that are nonlinear in performance has caused younger managers to "follow the herd" and take on less risk in their choice of investments.[22] A natural question arises: If nonlinear termination schedules potentially lead to inefficient herd behavior, why do firms use them? This paper argues that nonlinear termination is needed to alleviate another distortion stemming from managerial career concerns. In the current work, career concerns lead lower ability managers to take on excessive risk in their investment choices. Nonlinear termination (where the probability of termination is decreasing in performance up to the industry average and then constant for higher levels of performance) serves as a screening device that ensures efficient investment.[23]

A model of asymmetric information and project choice similar in spirit to Holmstrom and Ricart i Costra (1986) is developed. Managers are endowed with investment ability (their private information) and are hired by firms to invest in either a safe project or a risky project. Efficient investment occurs when managers of lower ability choose the safe project and higher ability managers invest in the risky project. The return on the safe project is independent of managerial ability; the return on the risky project is increasing in managerial ability. Higher ability managers have an incentive to invest in the risky

[22] See Chevalier and Ellison (1999a), Hong, Kubik and Solomon (2000) and Brown, Goetzmann and Park (2001). Scharfstein and Stein (1990) and Banerjee (1992) provide early models of herd behavior.
[23] See Figure 1 in Appendix B for an example of a nonlinear termination schedule.

project because it allows market participants to learn their ability levels. In the absence of contracting, lower ability managers are able to mimic the actions of their higher ability counterparts. All managers invest in the risky project.

Unlike Holmstrom and Ricart i Costra, the current work considers termination as a simple screening device that firms can employ to ensure efficient investment. It is assumed that being terminated is costly to a manager. By firing managers who achieve low returns when undertaking the risky project, firms can make it too costly for lower ability managers to mimic the actions of their higher ability peers.

Surprisingly, termination has received very little attention in the theoretical career concerns literature. Zwiebel (1995) is a notable exception. In Zwiebel, however, firms use termination as a sorting device. Lower ability managers are terminated; higher ability managers are retained. If firms used termination as a tool to retain higher ability managers, performance above the industry average would decrease the probability of termination. This is inconsistent with the nonlinear termination schedules described above. A screening theory of termination can explain this phenomenon. Termination schedules that are nonlinear in performance discourage risk-taking. Bad performance is punished; good performance, however, is not rewarded. The present model predicts termination will be used to deter excessive risk-taking.

A seemingly counterintuitive result arises in the present model. In equilibrium, only higher ability managers invest in the risky project and therefore only higher ability managers are terminated. The termination of higher ability managers in these industries has the potential to explain some puzzling stylized facts.

In particular, the termination of higher ability managers can reconcile the seemingly contradictory evidence in the mutual fund industry that stock-picking ability appears to exist even though actively managed funds have been unable to consistently outperform a set of passive benchmarks.[24] The termination of higher ability managers can also explain the bizarre result in Fee and Hadlock (2004) that managers who are forced from their jobs due to poor performance are more likely to find high level reemployment than managers who are exogenously separated from their positions. Lastly, the screening theory of termination described in the present model can also explain why high powered incentives fail to motivate hedge fund managers.[25]

The rest of the paper is organized as follows: Section II sets up the basic model without termination. In the absence of termination, all managers invest in the risky project. Section III shows how a simple termination policy can restore efficient investment. Section IV discusses the results of the model. Section V concludes.

Section II: The Basic Model without Termination

Agents, Projects and Timing

The basic model without termination is similar to Holmstrom and Ricart i Costra (1986). Firms hire managers to invest in projects. Managers have specific investment ability not possessed by firms so that all firms must hire a manager. Managers are wealth and credit constrained so they are unable to purchase firms. All firms and managers are risk neutral and do not discount the future.

There are two types of projects, a safe project and a risky project. For simplicity, the cost of investment in either project is normalized to zero. The safe project has a return of

[24] See Gruber (1996), Carhart, Carpenter, Lynch and Musto (2002), Chevalier and Ellison (1999b), Chen, Jegadeesh and Wermers (2000) and Wermers (2000).
[25] See Brown, Goetzmann and Park (2001).

S with probability one, where $S \in [a, b]$. Note that the return of the safe project is independent of managerial ability.

The return on the risky project, R, is increasing in managerial ability. Specifically, let R be given by:

$$R = m_i + \varepsilon_j \qquad (1)$$

where m_i is the innate investment ability of manager i and ε_j is a zero mean error term.

A manager's investment ability, m, is her private information; the market has only a known prior. Specifically, the market believes that a manager's investment ability is an independent draw from a uniform distribution with support on [a, b]. The error term, ε_j, is drawn from a uniform distribution on the interval [-c, c]. Error terms are uncorrelated with managerial ability and are uncorrelated with other error terms across time. It is assumed that $b - a < 2c$.[26]

The timing of the model is as follows: At time, t = 0, nature determines the ability of managers. At time, t = 1, firms hire managers to invest in projects. It is assumed that firms make all wage offers and that competition leads to managers earning their perceived marginal product. As is common in the literature, managers are paid before they make investment decisions. Once paid, managers select a project. Project choice and returns are observed by market participants and allow the labor market to revise its beliefs on managers' abilities. At time, t = 2, managers are paid to make the same investment decision. Once again, managers are paid their perceived marginal product before investment decisions are made. After managers invest for the second time, the game ends.

[26] A relationship between (b-a) and 2c must be assumed for the purposes of Bayesian updating with uniform distributions. The main qualitative results of the model would not change if (b-a) > 2c were assumed.

Managerial Objectives and Equilibrium

Efficient investment occurs when managers with ability, $m < S$, select the safe project and managers with ability, $m \geq S$, invest in the risky project. Managerial career concerns will prevent the efficient outcome from occurring.

Managers maximize the expected sum of their earnings. Recall that in each period a manager is paid her perceived marginal product before she invests. Thus, the wage a manager receives at the beginning of $t = 1$ is fixed. A manager's objective therefore reduces to investing in the project at $t = 1$ that maximizes her wage (her perceived marginal product) at the beginning of $t = 2$.

In keeping with the literature, it is assumed that managers will invest efficiently in the last time period as there is no further need to build a reputation. If firms believe managers will invest efficiently at $t = 2$, the expected value of a manager to the firm at the beginning of the time period will be S if a manager is believed to have ability, $m < S$, and m if a manager is believed to have ability, $m \geq S$. As in Holmstrom and Ricart i Costra, the value of a manager's ability has an option-like structure. When a manager's perceived ability falls below S, the manager still earns a wage equal to S.

The following proposition details the unique equilibrium of the game.

Proposition 1: In the absence of contracting, there exist only pooling equilibria in which all managers invest in the risky project.

Proof: See Appendix B.

A similar result arises in Holmstrom and Ricart i Costra. However, the result here occurs for a different reason. In Holmstrom and Ricart i Costra, the option structure on wages leads managers to ignore their private information and invest in the risky project even when their private information tells them they should refrain from investment. Here, higher ability managers have an incentive to invest in the risky project because it allows the market to update its beliefs on managerial ability. In the absence of contracting, lower ability managers are able to mimic the actions of their high ability peers.

Section III: The Model with Termination

Contracting

When managers have career concerns, all managers invest in the risky project. This section examines a simple solution to the managerial career concerns problem, the use of termination.

At the end of $t = 1$, a firm can choose to fire a manager after observing her project choice and returns. As in Zwiebel (1995), if a manager is fired she receives a penalty of F, where $F > 0$. F can represent search costs of finding a new job, loss of firm specific human capital or other reputation costs not considered in this paper. The size of the penalty, F, is fixed and the same for all managers.[27]

Firing penalties and career concerns produce a need for contracting. Managers, due to the firing penalty will seek protection from having their wages lowered below their perceived marginal products. Firms, as already demonstrated, will want to contract to mitigate the problem created by career concerns.

[27] For simplicity, there is no cost to the firm to fire a manager. The qualitative results of the model can still be obtained if firms were to pay a firing cost when terminating managers.

Consider the following contract: Managers are still paid before making investment decisions.[28] At t = 1, firms pay managers their perceived marginal product. If a manager invests in the safe project at t = 1, she receives a wage of S at the beginning of t = 2. If a manager invests in the risky project at t = 1, she receives a wage equal to her perceived marginal product at t = 2 as long as she is retained by her firm. At the end of t = 1, firms terminate managers who invest in the risky project according to the termination schedule T(R), where T(R) provides the probability of termination as a function of the return on the risky project.[29]

In an optimal contract, T(R) will be chosen so that managers invest efficiently. That is, an optimal termination schedule, $T(R)^*$, will induce managers with ability, m < S, to invest in the safe project and managers with ability, m ≥S, to invest in the risky project. The remainder of the paper will focus on the use of an optimal termination schedule.

Managerial Objectives

Managers still make investment choices at t = 1 that maximize their expected wages at t = 2. Consider a manager who invests in the safe project at t = 1: The manager receives a wage of S at the beginning of t = 2. Note that when firms employ an optimal termination schedule, the market belief on managerial ability for a manager who invests in the safe project is $(m \mid safe) \sim U[a, S]$.[30] Due to the option-like structure on project returns, the manager receives a wage equal to her perceived marginal product at the beginning of t = 2.

[28] This practice is maintained for continuity with the previous section and to facilitate comparisons with other career concerns models. The main results of the model can be established without this feature.
[29] For simplicity, T(R) is assumed to be continuous and differentiable.
[30] Technically, this is the limiting distribution.

If a manager invests in the risky project at t = 1 and firms employ an optimal termination schedule, then the market belief on managerial ability (based solely on project choice) is

$$(m \mid \text{risky}) \sim U[S, b] \tag{2}$$

The return on the risky project also allows the market to update its beliefs. The posterior on m after the return on the risky project is observed is[31]

$$(m \mid \text{risky}, R) \sim \begin{cases} U[a, R+c] & \text{if} \quad R \in [a-c, S-c) \\ U[S, R+c] & \text{if} \quad R \in [S-c, b-c] \\ U[S, b] & \text{if} \quad R \in (b-c, S+c) \\ U[R-c, b] & \text{if} \quad R \in [S+c, b+c] \end{cases} \tag{3}$$

(3) shows that after R is observed, the reputational states that can occur fall into four different regions.[32] Compared to (2), the first two regions (lower returns) lead to downward revisions on managerial ability. The third region does not change the belief on managerial ability. The fourth region (higher returns) leads to an upward revision of managerial ability. Managers with ability, m < S, can only achieve reputational states in the first three regions. Managers with ability, m ≥S, can only achieve reputational states in the last three regions.

A manager who invests in the risky project at t = 1 receives a wage equal to her perceived ability at t = 2. This is true whether or not the manager is retained by the firm. If the manager is retained by the firm, she receives a wage at t = 2 equal to her perceived marginal product as stipulated in the contract. If a manager is terminated, competition in

[31] The posterior can be obtained using procedures in DeGroot (1970).
[32] Given the use of an optimal termination schedule, a return on the risky project in the first region would be an off-equilibrium event. In this case, it is assumed the market only uses project returns (and not project choice) to determine the posterior on managerial ability.

the managerial labor market will also lead to her earning her perceived marginal product at t = 2. Note that if a manager achieves a reputational state in the first region given in (3), the option the safe project provides will lead to her earning a wage of S at t = 2.

For a manager investing in the risky project at t = 1, R is a random variable with a uniform distribution on the interval [m – c, m + c]. Let W_R denote a manager's expected perceived marginal product at t = 2 when investing in the risky project. Using (3) and the mean of a uniform distribution, W_R is given by:

$$W_R = \begin{cases} \frac{1}{2c}\left[\int_{m-c}^{S-c} S\,dR + \int_{S-c}^{b-c} \frac{S+R+c}{2}\,dR + \int_{b-c}^{m+c} \frac{S+b}{2}\,dR\right] & when\ \ m < S \\ \frac{1}{2c}\left[\int_{m-c}^{b-c} \frac{S+R+c}{2}\,dR + \int_{b-c}^{S+c} \frac{S+b}{2}\,dR + \int_{S+c}^{m+c} \frac{R-c+b}{2}\,dR\right] & when\ \ m \geq S \end{cases} \tag{4}$$

Let V represent the expected return to the manager of investing in the risky project relative to investing in the safe project. Carrying out the integration in (4), subtracting off S (the value of a manager who invests in the safe project at t = 1), and including the firing penalty, yields V:

$$V = \frac{1}{2c}\left[m\left(\frac{b-S}{2}\right) + c(b+S) + \frac{S^2}{4} - \frac{b^2}{4} - F\int_{m-c}^{m+c} T(R)^* \, dR\right] - S, \text{ for all m} \tag{5}$$

A manager will invest in the risky project if V is greater than or equal to zero.[33] The following lemma will help characterize a manager's investment strategy:

[33] As a tie-breaking convention, it is assumed that managers who are indifferent between investing in the safe project and the risky project will invest in the risky project.

Lemma 1: If the optimal termination schedule, $T(R)^*$, is monotonically nonincreasing in the return on the risky project, R, then the expected return of investing in the risky project relative to investing in the safe project is increasing in managerial ability, $\partial V/\partial m > 0$.

Proof: See Appendix B.

The intuition behind Lemma 1 is straightforward. When investing in the risky project, an increase in managerial ability increases the probability that the market will revise its beliefs upwards and allows higher reputational states to be achieved. Lemma 1 implies that there exists a cutoff level of managerial ability, m^*, such that managers with ability $m \geq m^*$, invest in the risky project and managers with ability, $m < m^*$, invest in the safe project. Specifically, m^* is defined as the value of m that yields V = 0.

Efficient Equilibrium and Optimal Termination

An efficient separating equilibrium in this model is a manager's investment strategy, m^*, and termination schedule, $T(R)^*$, such that: (i) the manager's investment strategy, m^*, maximizes her expected payoff given the termination schedule, $T(R)^*$, and (ii) the termination schedule, $T(R)^*$, is chosen so that managers invest efficiently. That is, $m^* = S$.

The following proposition shows existence of such an equilibrium:

Proposition 2: If (i) the termination schedule, $T(R)^*$, is monotonically nonincreasing in the return on the risky project, R, and (ii) $F > \frac{b-S}{2} - \frac{(b-S)^2}{8c}$, then there exists

an efficient separating equilibrium. In an efficient separating equilibrium, firms choose a termination schedule, such that,

$$\int_{S-c}^{S+c} T(R)^* dR = \frac{-(S^2/4) + S((b/2) - c) + bc - (b^2/4)}{F}.$$

Proof: See Appendix B.

Note that the optimal termination schedule, $T(R)^*$, is not unique. There are an infinite number of termination schedules that satisfy the conditions provided in Proposition 2, including the nonlinear termination schedules discussed at the beginning of the current work. Intuitively, potential termination makes investing in the risky project costly and allows higher ability managers to separate themselves from their lower ability peers. The expected cost of investing in the risky project depends on the size of the firing penalty and on the *overall* probability of termination. The overall required probability mass needed to ensure efficient investment can be distributed over the range of R in an infinite number of ways.

The next proposition provides comparative statics results on the required probability of termination needed to ensure efficient investment, $P = \int_{S-c}^{S+c} T(R)^* dR$:

Proposition 3: The probability of termination needed to ensure efficient investment is (i) increasing in the variance of perceived managerial ability based solely on project choice, $\partial P/\partial S < 0$ and $\partial P/\partial b > 0$, (ii) increasing in the variance of the return on the risky project, $\partial P/\partial c > 0$, and (iii) decreasing in the firing penalty, $\partial P/\partial F < 0$.

Proof: See Appendix B.

Proposition 3 shows that the probability of termination needed to ensure efficient investment, P, is decreasing in the firing penalty, F. The intuition behind this result is straightforward. The expected cost of investing in the risky project needed to ensure efficient investment equals the probability of termination multiplied by the firing penalty. If F increases, then P must be lowered to keep the expected cost constant to ensure efficient investment.

The probability of termination is increasing in c, which is proportional to the variance of the return on the risky project. As the variance of the return on the risky project increases, the return on the project becomes less informative for making inferences on managerial ability. As such, lower ability managers become more likely to invest in the risky project because they can receive the screening benefits of investing in the risky project without having to worry as much that the project's return will reveal their low levels of ability. To prevent managers of lower ability from investing in the risky project, firms must increase the probability of termination when investing in the risky project.

Lastly, the probability of termination is increasing in the variance of perceived managerial ability based solely on project choice. Recall from (2) that the belief on managerial ability when investing in the risky project is $(m \mid \text{risky}) \sim U[S, b]$. An increase in b, increases the screening benefits of investing in the risky project. As such, investing in the risky project becomes more attractive and firms must respond by increasing the probability of termination. Increasing the return on the safe project, S, increases the mean of $(m \mid \text{risky})$ as well. However, increasing S increases the wage of investing in the safe project (which pays S) by a greater magnitude. Thus, when S increases, managers find investing in the safe project more attractive and firms must lower the expected cost of investing in the risky project to ensure efficient investment.

Section IV: Discussion

Risk Deterrence

The empirical studies of Chevalier and Ellison (1999a), Hong, Kubik and Solomon (2000) and Brown, Goetzmann and Park (2001) on young mutual fund managers, security analysts and hedge fund managers, respectively, find that the probability of termination is nonlinear in performance. Specifically, the probability of termination is decreasing in performance up to the industry average and then constant for higher levels of performance. A sorting theory of termination cannot explain this phenomenon. If firms were trying to sort managers based on ability, performance above the industry average would reduce the probability of termination.

The present model can explain the use of nonlinear termination schedules. Any termination schedule satisfying the conditions in Proposition 2 deters risk-taking by lower ability managers and ensures efficient investment. The nonlinear termination schedules above fall into this category.

Another result found in Chevalier and Ellison (1999a) is that managers are more likely to be terminated, all else constant, when they take on more risk. This is akin to increasing c in the current model. As shown in Proposition 3, increasing c forces firms to respond by increasing the probability of termination.

The screening theory of termination described in the present model can also explain why high powered incentives fail to motivate hedge fund managers. Hedge fund managers receive bonuses for surpassing a high benchmark on annual returns. Brown, Goetzmann and Park (2001) find that these bonuses do little to change the conservative nature of hedge fund managers. Fears of termination play a bigger role in driving managerial behavior. The current work suggests that managers are naturally risk-takers and that termination is responsible for conservative behavior. If firms wanted their managers to take on more risk, lowering the probability of termination would be more effective than utilizing high powered incentives.

Termination of Higher Ability Managers

In equilibrium, only higher ability managers choose the risky project. Therefore, only higher ability managers are fired. Two natural questions arise: (i) Ex post, why would firms fire managers they know are of high ability? and (ii) Why don't firms replace known lower ability managers with higher ability managers in the second period?

Adding a longer timeframe to the model answers both questions. Consider the situation of an infinitely-lived firm dealing with overlapping generations of managers who live for two periods. The classic dynamic moral hazard argument applies. By not firing poor-performing higher ability managers, firms would run the risk of having managers not invest efficiently in future time periods. The same argument applies to why firms would not want to replace known lower ability managers with known higher ability managers.[34]

In addition, the highest ability managers may prefer the use of the contract in the present model. Investing in the risky project is only costly if a manager is terminated. If the optimal termination schedule is monotonically nonincreasing in the return on the

[34] In the context of the current model, such a strategy would not be optimal if firms faced positive firing costs.

risky project, the highest ability managers are the least likely to be terminated. They are able to separate themselves from lower ability managers at little cost. Other managers, as in a rat race (Akerlof, 1976), may be forced to follow suit.

The termination of higher ability managers can explain some puzzling empirical evidence. Fee and Hadlock (2004) find that managers who are fired for poor performance are more likely to find high level reemployment than managers who are exogenously separated from their positions. The current model can explain this result. Managers who have been terminated are of higher ability; managers retained by firms have ability levels across the entire range of managerial ability. When managers, who firms would otherwise have retained, lose their positions they find it harder to find high level reemployment because they are, on average, of lower ability than those managers who were terminated due to poor performance.[35]

The termination of higher ability managers also sheds light on the debate in finance on whether or not investment ability exists. The debate has mostly focused on the actively managed mutual fund industry. Those who believe investment ability does exist point out that fund companies appear to sort managers as if they are trying to learn about ability. Empirically, Chevalier and Ellison (1999b), Chen, Jegadeesh and Wermers (2000) and Wermers (2000) find some evidence of investment ability at the managerial level. Those who believe investment ability does not exist look at the evidence on the persistence of mutual fund performance through time. Hendricks, Patel and Zeckhauser (1993), Gruber (1996) and Carhart, Carpenter, Lynch and Musto (2002) find evidence of persistence in mutual fund performance over short horizons but find no long run persistence at the fund

[35] It should be noted that Fee and Hadlock argue that lower ability, not higher ability, managers are terminated. They base their findings on the wage and prestige of new jobs found by managers. A story of firm-specific human capital could also explain their results.

level. If investment ability were to exist, fund companies would eventually identify and stockpile high ability managers. If this were the case, actively managed funds would consistently be able to outperform a set of passive benchmarks. No funds have performed at such a high level.

The model presented here can explain the seemingly contradictory evidence. Investment ability is assumed to exist. Firms, however, do not sort managers in an attempt to identify and retain high ability types. The firm's objective is to discourage excessive risk-taking by its lower ability managers. As a result, in equilibrium, only higher ability managers are terminated. Since fund companies only terminate higher ability managers, actively managed funds should not be expected to consistently outperform passive benchmarks.

Section V: Conclusion

This paper has considered the role of termination in contracting in the presence of career concerns. Surprisingly, termination has received little attention in the theoretical career concerns literature. Here, termination is used as a mechanism to keep lower ability managers from taking on excessive risk in their choice of projects. The model can explain the nonlinear termination schedules used by firms in the mutual fund, securities analysis, and hedge fund industries. When termination is used to deter risk-taking, important testable implications emerge: (i) mangers that are terminated are, on average, of higher ability than managers who are retained by firms, (ii) firms that use termination to deter risk taking (and therefore retain lower ability mangers on average) should be more susceptible to takeovers and (iii) high powered incentives (e.g. stock grants or options) should be ineffective at motivating managers.

CHAPTER THREE

A COUNTERSIGNALING THEORY OF ADVERTISING AND FADS

Section I: Introduction

Since the pioneering work of Spence (1973), the standard signaling framework has yielded key insights into economic, social and biological behavior.[36] The framework is well-known. High ability agents take costly wasteful actions that cannot be imitated by their low ability counterparts in order to reveal their type to the market. In the standard framework, the signal serves as the only source of information on an agent's unobservable type.

By adding an additional noisy source of information on an agent's type to a standard signaling framework, Feltovich, Harbaugh and To (2002) show that a countersignaling equilibrium can exist. In a countersignaling equilibrium, medium ability agents take costly wasteful action to separate themselves from low ability agents. High ability agents, like their low ability counterparts, choose not to signal. By not signaling, high types are able to separate themselves from medium types. High types then rely on the additional noisy information to distinguish themselves from low types.

Feltovich, Harbaugh and To use countersignaling to explain a wide variety of social behavior, such as why the nouveau riche flaunt their wealth while old money does not or why mediocre students are quick to answer a teacher's easy questions while the best students choose not to display their knowledge. Harbaugh and To (2004) offer a countersignaling explanation as to why mediocre types might boast and disclose good news while high types refrain from such activity.

---

[36] See Riley (2001) for a survey of the literature.

The present work extends the literature on countersignaling by developing a countersignaling theory of advertising. In the current framework, firms can signal ability through costly advertising. In addition to advertising, consumers also obtain coarser information on a firm's ability level through positive word-of-mouth communication. A countersignaling equilibrium exists in which the highest and lowest ability firms choose not to advertise while "average" ability firms do advertise.

To fix ideas, consider a firm endowed with an innate ability to produce quality. The firm's ability level is its private information. At time t, the firm produces a good with an actual quality level that is a noisy signal of its innate ability. The firm must decide whether or not to undertake costly advertising; the decision, however, is made before actual quality is realized.

The firm produces a durable experience good so consumers can only discern quality after purchase. Consider a consumer who purchases the good at time t. Purchasing the good allows the consumer to observe quality and update her beliefs on the firm's ability level. The extra information on the firm's ability level, however, is useless to the consumer since she does not purchase the (durable) good again. In the absence of communication between consumers, the next generation of consumers who will purchase the good at time, t + 1, learn nothing from the previous generation of consumers.

Information is spread by consumers through casual interactions with one another. Two events enable the spread of information: (1) Unusual or noteworthy events are often topics of conversation. If a consumer has knowledge of a good that is of exceptionally high quality, she is assumed to pass on the information to other consumers. This event will be referred to as "positive word-of-mouth" communication. Positive word-of-mouth communication can occur regardless of whether or not a firm advertises. The highest ability firms are the ones most likely to receive positive word-of-mouth. (2) "Shared experiences" and commonalities are also topics of conversation. Observing an advertisement creates a commonality for consumers. Thus, if a firm advertises, it is more likely that the information learned by a consumer who purchases the good at time t will be passed on to future consumers through social interaction. That is, advertising enables consumers, as a whole, to better learn a firm's ability to produce quality.

As is common in models of asymmetric information, multiple equilibria exist. Like other models of advertising as a signal of product quality and in keeping with the standard signaling framework, there exists a separating equilibrium in which higher ability firms choose to advertise and lower ability firms do not advertise. Advertising allows consumers to better learn a firm's ability level, therefore higher ability firms have a greater incentive to advertise compared to their lower ability peers. Existence of this equilibrium, however, is jeopardized if firms can easily obtain positive word-of-mouth.

As mentioned above, a countersignaling equilibrium also exists in which the highest and lowest quality firms do not advertise while average ability firms do advertise. This result is driven by the positive word–of–mouth mechanism.[37] In this equilibrium, low quality firms do not advertise because they do not want their low ability levels to be learned. Average ability firms advertise to separate themselves from the lowest ability firms. The highest ability firms do not advertise to separate themselves from average

[37] In a true countersignaling framework, word-of-mouth would provide consumers with different information than advertising. Here, word-of-mouth and advertising potentially provide consumers with the same information. The term countersignaling is used here because of the many similarities in the present model to the general framework of Feltovich, Harbaugh and To.

ability firms. After pooling with low types, high types then rely on positive word-of-mouth to separate themselves from the lowest ability firms.

It is interesting to compare the treatment of advertising in the current work to the role it plays in the existing literature. In Nelson (1974), Kihlstrom and Riordan (1984), Milgrom and Roberts (1986), and Horstmann and MacDonald (1994), advertising signals product quality through a repeat purchase mechanism. The more a firm advertises, the more likely it produces a high quality product and will be able to recoup costly advertising expenditures through repeat purchases of the good.

This treatment of advertising leaves a lot to be desired. First, it is not truly a theory of advertising. As the authors note, advertising could be replaced in their models by any conspicuous wasteful expenditure. As such, these models provide no insight into common advertising practices. For example, they are unable to explain the common practice of not referring to competitors by name. Second, as Becker and Murphy (1993)[38] argue, if wasteful advertising expenditures signal high quality then firms would want to advertise their level of advertising expenditures. Casual observation suggests that firms do not engage in this practice.[39] Lastly, since the models rely on a repeat purchase mechanism, they only apply to goods which consumers purchase frequently.

The difference between the treatment of advertising in the current model and the conspicuous wasteful expenditure approach is best illustrated by example. Consider the advertising campaign for the 1984 Ford Ranger truck discussed in Milgrom and Roberts. The ads feature trucks being thrown out of airplanes and be driven off high cliffs. Milgrom and Roberts argue the clearest message being sent to consumers is, "We are

[38] Becker and Murphy treat advertising as an input into a consumer's stable utility function.
[39] See Hertzendorf (1993) for a model in which a consumer must infer a firm's expenditures on advertising from the number of times she witnesses an advertisement.

spending an astronomical amount of money on this ad campaign." In their model, consumers realize throwing trucks out of airplanes is costly; therefore the trucks must be high quality. In the current framework, advertising is not a conspicuous wasteful expenditure; it facilitates learning. In the present model, when consumers witness a truck being thrown out of an airplane, it leads to general conversations on the truck. In these discussions, consumers who have information on the product's quality pass on that information to uninformed consumers. The present work maintains that the latter framework is a better description of how consumers respond to advertising.

Note that in the current model, firms still have an incentive to advertise their levels of advertising. However, since advertising facilitates information exchange among consumers, firms have more of an incentive to focus on making their advertisements entertaining and noteworthy. The present work can also explain the "Brand X" advertising practice where competing products are not mentioned by name. In the present framework, mentioning a competitor by name would provide free advertising to a rival. Lastly, since advertising facilitates learning by consumers as a whole, the present model applies to any good in which consumers imperfectly learn a firm's ability to produce quality. That is, the model can apply to both frequently purchased goods and durable goods.[40]

The current model might also better explain empirical evidence on quality and advertising. In a cross-sectional analysis, Caves and Greene (1996) find advertising does not serve as a quality signal. This contradicts the standard signaling theory of advertising

[40] Though the present model can apply to goods in which consumers make repeat purchases, only durable goods are considered in the present work. This assumption simplifies the analysis by ensuring that all relevant consumers have the same information set. As such, firms do not have to consider the issue of how to price their goods in the presence of better informed and less informed consumers. See Linnemer (2002) for a model of pricing and advertising with informed and uninformed consumers.

(where advertising and quality are positively correlated), but it is not inconsistent with the current model. In a countersignaling equilibrium, the decision to advertise is not monotonically increasing in a firm's ability to produce quality. Thus, it is not surprising to find a zero correlation between advertising and quality.[41] In time series analyses, Thomas, Shane and Weigelt (1998) and Horstmann and MacDonald (2003) find an inverted "U" relationship between advertising intensity and a product's time on the market. The present model can explain this relationship if older high ability firms find themselves unable to obtain positive word-of-mouth over time.

A simple one period extension of the current framework can also explain the social phenomena of fads. The existing literature on fads has focused solely on consumer behavior. Explanations have included sanctions on deviants, positive payoff externalities, conformity preference, communication, and informational cascades.[42] The literature seems to miss one key aspect of fads; they are ubiquitous. In the current model, a fad occurs when information on a product spreads both through advertising and positive word-of-mouth. After a fad hits the market, it is shown that in the following period, an appealing equilibrium has the firm cut its advertising, thereby crushing the fad.

The remainder of the paper is organized as follows. The next section develops the basic model. Section III demonstrates existence of a standard separating equilibrium and a countersignaling equilibrium. Section IV discusses the results and provides the extension on fads. Section V concludes.

---

[41] Orzach, Overgaard and Tauman (2002) develop a model in which low types advertise more than high types. In their model, advertising is not a wasteful expenditure; it informs consumers of the product's existence and thus determines the upper bound on sales.

[42] See Bikhchandani, Hirshleifer and Welch (1992) for a theory of informational cascades as well as a summary of alternative theories of fads and conformity.

Section II: The Basic Model

Agents, Production and Information

The basic model is constructed to offer the simplest nontrivial analysis possible. At time, t = 0, nature endows a firm with an innate ability to produce quality, η. The firm's ability level is its private information. The market's known prior on the firm's ability is $\eta \sim U\,[a, b]$. It is assumed that a and b are nonnegative. The firm is risk neutral and does not discount the future.

At time, t = 1, the firm produces a durable good of variable quality. Let the quality of the good produced by the firm at any time t be given by:

$$q_t = \eta + \varepsilon_t \qquad (1)$$

The ε_t term is an unobservable random shock to quality that is intended to capture uncertainty in the consumption experience. Thus, immediately after producing a good, the firm cannot observe ε_t or q_t. It is assumed that ε_t is a draw from a uniform distribution with support on the interval [c, c + d]. The random shocks to quality are uncorrelated with the firm's ability level and uncorrelated with shocks over time. It is assumed that c and d are nonnegative and that d > (b – a)[43].

The firm produces the good at a constant unit cost, C = a + (2c + d)/2. Note that the unit cost of production is independent of quality and that firms do not choose quality in the present framework. As in Horstmann and MacDonald (1994), these assumptions are maintained to focus on consumer learning.

After production occurs at t = 1, the firm sells its output to consumers. The good being produced is an experience good so consumers cannot discern quality by inspection. At

[43] A relationship between d and (b – a) must be assumed for the purpose of Bayesian updating with uniform distributions.

any time, t, there are N risk neutral consumers who are willing to pay a maximum price, $p_{max} = E(q_t) = E(\eta) + (2c + d)/2$, for exactly one unit of the good. A consumer who purchases the good at t = 1 is able to observe quality and update her beliefs on the firm's ability level. Once a consumer has purchased the (durable) good, she never repurchases.

Advertising, Word-of-Mouth and Observability

At time, t = 2, the firm chooses whether or not to advertise. If the firm advertises, it incurs a fixed cost, A > 0.[44] The benefits of advertising to the firm come from consumer communication as follows:

A consumer who purchases the good at t = 1 observes quality and updates her beliefs on the firm's ability level. The new information on the firm's ability level, however, is useless to a consumer who purchases the good at t = 1 since she never repurchases the (durable) good. In the absence of communication between consumers, the next generation of consumers who will purchase the good learn nothing from the consumers who made purchases at t = 1.

It is assumed that advertising facilitates the spread of information during casual interactions by consumers. Observing an advertisement creates a "shared experience" for consumers. Consumers are more likely to discuss shared experiences and commonalities in every day conversations. Thus, if a firm advertises, it is more likely that the information learned by consumers who purchase the good at t = 1 will be passed on to future consumers of the good. For simplicity, it is assumed that if a firm advertises, information on quality (q_1) reaches all market participants. That is, the market can

[44] Kihlstrom and Riordan (1984) and Horstmann and MacDonald (1994) also treat advertising expenditures as a binary variable. A countersignaling equilibrium can still be established by allowing advertising expenditures to be a continuous choice variable. See Feltovich, Harbaugh and To (2002) for an example.

collectively update its beliefs on the firm's ability, η. Thus, the benefit of advertising is that it allows the market to learn a firm's ability level.

In addition to advertising, there is another source of information for future consumers. Unusual or noteworthy events are also discussed in every day conversations. If consumers have knowledge of a good of exceptionally high quality, they are assumed to pass on the information to other consumers. This positive word-of-mouth communication occurs whether or not the firm chooses to advertise. Analytically, if a firm produces a good with a quality level, $q_1 \geq W$, positive word-of-mouth occurs and information on q_1 spreads to all market participants. It is assumed that $(a + c + d) < W$. That is, the lowest quality firms are unable to produce a quality level such that positive word-of-mouth occurs.[45]

It is interesting to compare the information consumers obtain from advertising to the information obtained through the positive word-of-mouth mechanism. For higher quality levels, $q_1 \geq W$, advertising and positive word-of-mouth provide consumers with the same information, an exact observation on q_1. For lower quality levels, $q_1 < W$, advertising still provides consumers with an exact observation on q_1. The word-of-mouth mechanism, however, does not provide consumers with an exact observation. The only information consumers receive in this situation is that they did not receive positive word-of-mouth. This enables consumers to learn that $q_1 < W$, but does not provide any more precise information. Thus, for lower levels of quality, the positive word-of-mouth mechanism provides consumers with coarser information than advertising.

[45] The case where all firms can obtain positive word-of-mouth is analyzed in the next section.

Firms, at t = 2, make their advertising decisions before consumers communicate with one another. Thus, a firm does not know if it will benefit from positive word-of-mouth when it chooses whether or not to advertise.[46] At time, t = 3, firms produce a new run of the good and sell to the next generation of consumers.

The Firm's Objective

The firm maximizes the sum of expected profits. Note that the only decision the firm makes is whether or not to advertise at t = 2. Thus, the firm's objective reduces to making the advertising decision at t = 2 that leads to the greater expected profit at t = 3. The firm's expected profit at t =3 is given by:

$$E(\pi) = \begin{cases} (E(\eta \mid \text{advertise}, q1) - a)N - A, \text{if the firm advertises} \\ (E(\eta \mid \text{no advertise}, q_1) - a)N, \text{if the firm does not advertise and } q1 \geq W \\ (E(\eta \mid \text{no advertise}, q_1 < W) - a)N, \text{if the firm does not advertise and } q_1 < W \end{cases}$$

Since the lower bound on the firm's ability level is a, the firm could always choose not to advertise and would be guaranteed positive expected profits. Thus, the firm's participation constraint is always satisfied. Also, note that the firm's expected profits are linear in E(η). Let z = A/N, denote advertising expenditures per unit of output. Profit maximization implies the firm will maximize E(η) by deciding whether or not to advertise and incur cost, z. For the remainder of the paper, E(π) will denote the expected value of E(η) less advertising expenditures.

Section III: Equilibria

Separating Equilibrium

In the standard signaling framework, there exists a separating equilibrium in which higher ability firms advertise and lower ability firms do not advertise. It is not readily

[46] This assumption is crucial for the existence of a countersignaling equilibrium.

apparent that such an equilibrium exists in the current model. On the one hand, advertising allows consumers to learn a firm's ability level. Thus, firms with higher levels of ability have a greater incentive to advertise than their lower ability peers. On the other hand, higher ability firms are more likely to benefit from positive word-of-mouth. If positive word-of-mouth occurs, a firm can have consumers learn its ability level without incurring advertising costs. As shown below, the ease of obtaining positive word-of-mouth plays a key role in whether or not a separating equilibrium exists.

Let the market have the following belief: Firms with ability, $\eta \geq \eta^*$, engage in costly advertising; firms with ability, $\eta < \eta^*$, do not advertise. In a separating equilibrium: (i) firms make the advertising decision that maximizes their expected profits given market beliefs, (ii) market beliefs are correct and (iii) the market updates its beliefs according to Bayes' Rule whenever possible.

It is assumed that $\eta^* < (W - c - d)$. That is, firms with an ability level less than or equal to η^* cannot have positive word-of-mouth occur.[47] If a firm does not advertise, the following two reputational states can occur:

$$(\eta|\text{ no advertise}) \sim \begin{cases} U[a,\eta^*] \text{ if } q_1 < W \\ U[q_1 - c - d, b] \text{ if } q_1 \geq W \end{cases} \tag{2}$$

The first state given in (2) is on the equilibrium path. Given market beliefs, if a firm does not advertise, the prior belief on a firm's ability is right-truncated. If $q_1 < W$, future consumers do not observe q_1 and no further learning occurs. Firms with ability, $\eta < \eta^*$, can never benefit from positive word-of-mouth. Given market beliefs, the second state in (2) is off the equilibrium path; a firm must have ability, $\eta \geq (W - c - d) > \eta^*$, to

[47] This assumption, as will be shown below, is necessary for existence of a separating equilibrium.

have positive word-of-mouth occur. Recall that q_1 is observable when positive word-of-mouth is obtained. In this case, market beliefs are simply revised using Bayes' Rule and the observation on q_1.

A firm with an ability level, $\eta < W - c - d$, can only achieve the first reputational state given in (2). For these firms, the expected payoff from not advertising is $(a + \eta^*)/2$. A firm with an ability level, $\eta \geq W - c - d$, can achieve both reputational states given in (2). From a firm's point of view, q_1 is uniformly distributed on the interval $[\eta + c, \eta + c + d]$. Thus, expected profits from not advertising are given by:

$$E(\pi | \eta < W\text{-}c\text{-}d, \text{ no advertise}) = \frac{a+\eta^*}{2} \tag{3}$$

$$E(\pi | \eta \geq W\text{-}c\text{-}d, \text{ no advertise}) = \frac{1}{d}\left[\int_{\eta+c}^{W} \frac{a+\eta^*}{2} dq_1 + \int_{W}^{\eta+c+d} \frac{q_1 - c - d + b}{2} dq_1\right] \tag{4}$$

Given market beliefs, the prior belief on a firm's ability is left-truncated when a firm engages in costly advertising. That is, $(\eta \mid \text{advertise}) \sim U[\eta^*, b]$. Advertising allows consumers to observe q_1 and further update their beliefs on a firm's ability level. If a firm advertises, the following reputational states can occur:

$$(\eta | \text{ advertise}, q_1) \sim \begin{cases} U[a, q_1 - c] \text{ if } q_1 < \eta^* + c \\ U[\eta^*, q_1 - c] \text{ if } q_1 \in [\eta^* + c, b + c) \\ U[\eta^*, b] \text{ if } q_1 \in [b + c, \eta^* + c + d] \\ U[q_1 - c - d, b] \text{ if } q_1 > \eta^* + c + d \end{cases} \tag{5}$$

The first state given in (5) is off the equilibrium path and can only be obtained by firms with ability levels, $\eta < \eta^*$. In this case, beliefs are updated using Bayes' Rule and the observation q_1. Signaling by advertising is not factored into the belief as the low level of quality has proven the firm to have an ability level less than η^*.

Expected profits from advertising are given by:

$E(\pi|$ advertise, $q_1, \eta < \eta^*) =$

$$\frac{1}{d}\left[\int_{\eta+c}^{\eta^*+c}\frac{a+q_1-c}{2}dq_1+\int_{\eta^*+c}^{b+c}\frac{\eta^*+q_1-c}{2}dq_1+\int_{b+c}^{\eta+c+d}\frac{\eta^*+b}{2}dq_1\right]-z \tag{6}$$

$E(\pi|$ advertise, $q_1, \eta \geq \eta^*) =$

$$\frac{1}{d}\left[\int_{\eta+c}^{b+c}\frac{\eta^*+q_1-c}{2}dq_1+\int_{b+c}^{\eta^*+c+d}\frac{\eta^*+b}{2}dq_1+\int_{\eta^*+c+d}^{\eta+c+d}\frac{q_1-c-d+b}{2}dq_1\right]-z \tag{7}$$

Let V denote the expected net return to advertising. That is, let V = $E(\pi|$ advertise, $q_1)$ – $E(\pi|$ no advertise). The following lemma will be helpful in demonstrating the existence of a separating equilibrium.

Lemma 1: Given the market beliefs above, the expected net return to advertising, V, is increasing in ability for firms unable to achieve positive word-of-mouth, $\eta < W-c-d$, and is decreasing in ability for firms that can achieve positive word-of-mouth, $\eta \geq W-c-d$.

Proof: See Appendix C.

The intuition behind Lemma 1 is as follows: Advertising allows market participants to observe quality and revise their beliefs on a firm's ability level. Since advertising allows learning to occur, the benefits from advertising are greater for higher ability firms. That is, $\partial E(\pi|$ advertise, $q_1)/\partial\eta > 0$. Now consider how $\partial E(\pi|$ no advertise$)/\partial\eta$ varies with ability. If a firm has an ability level, $\eta < W - c - d$, it cannot have positive word-of-mouth occur. If a firm that cannot achieve positive word-of-mouth does not advertise, q_1 is never observed by the next generation of consumers. In this case, higher values of q_1 (due to higher levels of ability) do not benefit the firm since q_1 is effectively unobserved by the market. That is, $\partial E(\pi|$ no advertise$)/\partial\eta = 0$. Thus, overall the expected net return to advertising is increasing in ability for lower ability firms. Firms with higher levels of ability, $\eta \geq W - c - d$, can have positive word-of-mouth occur. Like advertising, when a firm achieves positive word-of-mouth, quality is observed and market participants can revise their beliefs on the firm's ability level. The same argument applies: since positive word-of-mouth allows learning to occur, the benefits of positive word-of-mouth are greater for higher ability firms. Unlike advertising, where quality is observed with probability one, learning through positive word-of-mouth is not guaranteed. The higher the ability level, the more likely positive word-of-mouth (and thus learning) will occur and the less likely a firm will suffer the negative consequences from not advertising. Since ability also affects the probability learning will occur when a firm does not advertise, the returns to ability are greater when a firm chooses not to advertise than when it advertises. That is, for higher ability firms, $\partial E(\pi|$ advertise, $q_1)/\partial\eta < \partial E(\pi|$ no advertise$)/\partial\eta$. Overall, the net return to advertising is decreasing in ability for higher ability firms.

Firms will advertise if V is greater than or equal to zero; firms will refrain from advertising if V is negative.[48] In a separating equilibrium, it must be the case that the expected net return to advertising is greater than or equal to zero for firms with ability, $\eta \geq \eta^*$, and the expected net return to advertising is less than zero for firms with ability, $\eta < \eta^*$. The following proposition provides necessary conditions for the existence of a separating equilibrium.

Proposition 1: If $z \in [\frac{2d(b-a)-(b-a)^2}{4d}, \frac{2d(b-a)-(b-\eta_{\max})^2}{4d}]$**, there exists a separating equilibrium in which firms with ability,** $\eta \geq \eta^*$**, advertise and firms with ability,** $\eta < \eta^*$**, do not advertise. The values of** η^* **and** $\eta_{\max}$ **are given by** $\eta^* = b - \sqrt{2d(b-a-2z)} < W - c - d$ **and**

$$\eta_{\max} = \frac{(b+W-c-d) - \sqrt{b^2 + 2b(b-(W-c-d)) - 4a(b-(W-c-d)) - (W-c-d)^2}}{2}$$

Proof: See Appendix C.

A firm with ability, $\eta = \eta^*$, will be indifferent between advertising and not advertising. From Lemma 1, it follows that firms with ability, $\eta < \eta^*$, will refrain from advertising and firms with ability, $\eta^* \leq \eta < W - c - d$, will choose to advertise. However, it is not guaranteed that firms with ability, $\eta \geq W - c - d$, will want to advertise since V is decreasing in ability for these firms.

The $\eta_{\max}$ term above provides the maximum value of η^* that ensures the highest ability firm will find it profitable to advertise. A maximum cutoff value of η^* exists

[48] As a tie-breaking convention it is assumed when firms are indifferent between advertising and not advertising (V = 0), they advertise.

because $\partial V/\partial\eta^* < 0$ for firms with ability, $\eta > W - c - d$. From (2) and (4) the payoff from not advertising and not having positive word-of-mouth occur is increasing in η^*. For firms with ability, $\eta > W - c - d$, the absence of positive word-of-mouth occurs with probability $(W - c - \eta)/d$. The payoff to a firm that advertises is also increasing in η^*. However, increasing η^* only benefits a firm that advertises if the second or third states in (5) occur. That is, firms only benefit if low or intermediate quality levels arise. If a firm advertises, the second or third states in (5) occur with probability $(\eta^* + d - \eta)/d$. Since $\eta^* < W - c - d$, the probability of not obtaining word-of-mouth (a very high quality level) is greater than the probability of obtaining low or intermediate quality levels when advertising. Thus, the expected net return to advertising is decreasing in η^* for firms with ability, $\eta \geq W - c - d$. If η^* increases, it is possible that the highest ability firm will find it more profitable to refrain from advertising. As long as $\eta^* < \eta_{max}$, the highest ability firm will still want to advertise and a separating equilibrium will exist.

Similar to other models where advertising signals product quality, the separating equilibrium here predicts advertising and quality will be positively correlated.[49] Note that $\partial\eta^*/\partial z > 0$. If advertising costs increase, lower ability firms will find advertising less profitable and thus fewer firms will advertise. Also note that η_{max} is increasing in W. That is, as positive word-of-mouth becomes easier to obtain, the maximum possible value of η^* decreases. Thus, positive word-of-mouth decreases the parameter space in which a separating equilibrium exists. These comparative statics results will be reversed in a countersignaling equilibrium.

[49] See Nelson (1974), Kihlstrom and Riordan (1984), Milgrom and Roberts (1986) and Horstmann and MacDonald (1994) for models where advertising signals quality.

Countersignaling Equilibrium

Lemma 1 showed that due to the positive word-of-mouth mechanism, the expected net return to advertising is not monotonically increasing in ability. Although Lemma 1 was derived for a specific market belief, a result similar in spirit holds for a variety of market beliefs.[50] As such, it is possible for a countersignaling equilibrium to arise in which the highest and lowest ability firms refrain from advertising while average ability firms advertise.

Let the market have the following beliefs: Firms with ability, $\eta > H_1$, and firms with ability, $\eta < L_1$, do not advertise; firms with ability, $L_1 \leq \eta \leq H_1$, advertise. In a countersignaling equilibrium: (i) firms make the advertising decision that maximizes their expected profits given market beliefs, (ii) market beliefs are correct and (iii) the market updates its beliefs using Bayes' Rule whenever possible.

It is assumed that $L_1 < W - c - d < H_1$. This assumption guarantees that firms with ability, $\eta \leq L_1$, cannot obtain positive word-of-mouth. In a countersignaling equilibrium, medium ability firms advertise to separate themselves from the lowest ability firms. The next section will show that it is necessary for the lowest ability firms not to be able to obtain positive word-of-mouth in order for separation to occur.

Consider a firm that advertises. Given market beliefs, advertising both left and right truncates the prior on the firm's ability level. Thus, $(\eta |$ advertise$) \sim U[L_1, H_1]$. Advertising allows consumers to observe q_1 and further update their beliefs on the firm's ability level. When a firm advertises, the following reputational states can occur:

[50] A result similar in spirit to Lemma 1 also arises in a pooling equilibrium in which all firms advertise and in a pooling equilibrium in which no firm advertises. Pooling equilibria will be discussed in the next section.

$$(\eta|\text{ advertise}, q_1) \sim \begin{cases} U[a, q_1 - c] \text{ if } q_1 < L_1 + c \\ U[L_1, q_1 - c] \text{ if } q_1 \in [L_1 + c, H_1 + c) \\ U[L_1, H_1] \text{ if } q_1 \in [H_1 + c, L_1 + c + d] \\ U[q_1 - c - d, H_1] \text{ if } q_1 \in (L_1 + c + d, H_1 + c + d] \\ U[q_1 - c - d, b] \text{ if } q_1 > H_1 + c + d \end{cases} \qquad (8)$$

The first and last states given in (8) are off the equilibrium path. Off equilibrium beliefs are treated in the same manner as in the separating equilibrium: Beliefs are updated using Bayes' Rule and the observation on q_1. Signaling by advertising is not factored into the belief.

Firms with ability, $\eta < L_1$, can only obtain the first three states given in (8), firms with ability, $L_1 \leq \eta \leq H_1$, can only achieve reputations in the middle three states and firms with ability, $\eta > H_1$, can only have outcomes in the last three states occur. From a firm's perspective, q_1 is uniformly distributed on the interval $[\eta + c, \eta + c + d]$. Expected profits from advertising are given by:

$E(\pi|\text{ advertise}, q_1, \eta < L_1) =$

$$\frac{1}{d}[\int_{\eta+c}^{L_1+c} \frac{a + q_1 - c}{2} dq_1 + \int_{L_1+c}^{H_1+c} \frac{L_1 + q_1 - c}{2} dq_1 + \int_{H_1+c}^{\eta+c+d} \frac{L_1 + H_1}{2} dq_1] - z \qquad (9)$$

$E(\pi|\text{ advertise}, q_1, L_1 \leq \eta \leq H_1) =$

$$\frac{1}{d}[\int_{\eta+c}^{H_1+c} \frac{L_1 + q_1 - c}{2} dq_1 + \int_{H_1+c}^{L_1+c+d} \frac{L_1 + H_1}{2} dq_1 + \int_{L_1+c+d}^{\eta+c+d} \frac{q_1 - c - d + H_1}{2} dq_1] - z$$

(10)

E(π| advertise, q_1, $\eta > H_1$) =

$$\frac{1}{d}\left[\int_{\eta+c}^{L_1+c+d}\frac{L_1+H_1}{2}dq_1+\int_{L_1+c+d}^{H_1+c+d}\frac{q_1-c-d+H_1}{2}dq_1+\int_{H_1+c+d}^{\eta+c+d}\frac{q_1-c-d+b}{2}dq_1\right]-z \tag{11}$$

Now consider a firm that does not advertise. Given market beliefs, if a firm refrains from advertising, the posterior on the firm's ability is given by:

$$f(\eta|\text{ no advertise}) = \begin{cases}\dfrac{1}{(L_1-a)+(b-H_1)} & \text{when } x\in[a,L_1) \text{ or } x\in(H_1,b] \\ 0 & \text{otherwise}\end{cases} \tag{12}$$

When a firm does not advertise, the posterior on the firm's ability puts zero probability on the firm having an ability level on the interval [L_1, H_1]. The probability mass on this interval in the prior distribution is evenly spread out over [a, L_1) and (H_1, b] in the posterior distribution. That is, the posterior on η is a proper uniform distribution on the interval [a, b] with support missing on [L_1, H_1].

When a firm does not advertise, two distinct events can occur: the firm obtains positive word-of-mouth or it does not obtain positive word-of-mouth. Given market beliefs, if positive word-of-mouth occurs it must be the case that (η| no advertise, positive word-of-mouth) ~ U [H_1, b], since the lowest ability firms that refrain from advertising cannot obtain positive word-of-mouth. When positive word-of-mouth occurs, consumers are provided with an exact observation on q_1 and can further update their beliefs. If positive word-of-mouth occurs the following two reputational states can arise:

$$(\eta|\text{ no advertise, } q_1 \geq W) \sim \begin{cases} U[H_1,b] \text{ if } q_1\in[W,H_1+c+d] \\ U[q_1-c-d,b] \text{ if } q_1 > H_1+c+d\end{cases} \tag{13}$$

If positive word-of-mouth does not occur, consumers do not receive an exact observation on q_1. The only extra information consumers obtain is that they did not hear any positive word-of-mouth. Hence, consumers learn that $q_1 < W$. The derivation of expected profits when a firm does not advertise and does not obtain positive word-of-mouth is cumbersome. For now, let f (η| no advertise, q_1) be the posterior distribution on η given an exact observation on q_1. Let $g(q_1)$ be the probability distribution of q_1 given the prior on η in (12). Let G denote the cumulative density function of $g(q_1)$. Let R denote profits when a firm does not advertise and does not receive positive word-of-mouth. Given the notation above,

$$R = \frac{\int_{a+c}^{W} g(q_1)(\int(x) f(\eta \mid \text{no advertise}, q_1)dx)dq_1}{G(q_1 < W)}$$

Lemma 2: Expected profits when a firm does not advertise and does not obtain positive word-of-mouth are given by,

$$R = \frac{3(b^2 - H_1^2)(W - c - d) + 3d(L_1^2 - a^2 + b^2 - H_1^2) - 2(b^3 - H_1^3)}{6(b - H_1)(W - c - d) + 6d(L_1 - a + b - H_1) - 3(b^2 - H_1^2)}$$

Proof: See Appendix C.

Again, from a firm's perspective, q_1 is uniformly distributed on the interval $[\eta + c, \eta + c + d]$. Expected profits from not advertising are given by:

$$E(\pi \mid \text{no advertise}, \eta < W - c - d) = R \quad (14)$$

$$E(\pi \mid \text{no advertise}, W - c - d \leq \eta \leq H_1) = \frac{1}{d}\left[\int_{\eta+c}^{W} R\,dq_1 + \int_{W}^{\eta+c+d} \frac{H_1 + b}{2} dq_1\right] \quad (15)$$

$E(\pi \mid \text{no advertise}, \eta > H_1) =$

$$\frac{1}{d}\left[\int_{\eta+c}^{W} R dq_1 + \int_{W}^{H_1+c+d} \frac{H_1+b}{2} dq_1 + \int_{H_1+c+d}^{\eta+c+d} \frac{q_1-c-d+b}{2} dq_1\right] \tag{16}$$

The expressions in (9)-(11) and (14)-(16) can be used to determine V, the expected net return to advertising. A result similar to Lemma 1 arises:

Lemma 3: Given market beliefs, if R $<\frac{L_1+H_1}{2}$, the expected net return to advertising is increasing in ability for firms unable to achieve positive word-of-mouth, $\eta < W - c - d$, and is decreasing in ability for firms that can achieve positive word-of-mouth, $\eta \geq W - c - d$.

Proof: See Appendix C.

Lemma 3 is vital to the existence of a countersignaling equilibrium. In a countersignaling equilibrium, the expected net return to advertising must be greater than or equal to zero for firms with ability, $\eta \in [L_1, H_1]$. The expected net return to advertising is continuous in ability. The implication of Lemma 3 is if V = 0 when $\eta = L_1$ and when $\eta = H_1$, then there exists a countersignaling equilibrium. The following lemma provides necessary conditions for the existence of a countersignaling equilibrium.

Lemma 4: There exists a countersignaling equilibrium in which firms with ability, $\eta < L_1$ or $\eta > H_1$, refrain from advertising and firms with ability, $L_1 \leq \eta \leq H_1$, advertise, if the following three conditions hold:

(i) $\quad R < \frac{L_1+H_1}{2}$

(ii) $\quad \frac{1}{d}\left[d\frac{L_1+H_1}{2} - \frac{(H_1-L_1)^2}{4}\right] - z = R$

(iii) $$\frac{1}{d}[d(\frac{L_1+H_1}{2})+\frac{(H_1-L_1)^2}{4}]-z=\frac{1}{d}[(W-H_1-c)(R-\frac{H_1+b}{2})+d(\frac{H_1+b}{2})]$$

Proof: See Appendix C.

Condition (i) is needed for Lemma 1 to hold. Condition (ii) ensures that V = 0 when $\eta = L_l$ and condition (iii) guarantees V = 0 when $\eta = H_l$. Analytical solutions for L_l and H_l as well as parameter restrictions on z necessary for existence of a countersignaling equilibrium can be obtained, but the expressions are abstruse and unintuitive. Instead, existence of a countersignaling equilibrium will be demonstrated through a numerical example.

Result 1: Let a = 0, b = 1, c = 0, d = 2, W = 2.75 and z = 0.35. There exists a countersignaling equilibrium in which $L_l \approx 0.440858$ and $H_l \approx 0.934846$.

The main implication of Result 1 is that there exists a countersignaling equilibrium. The intuition is as follows: Consumers can learn a firm's ability level from advertising and from the word-of-mouth mechanism. For high quality levels, $q_l \geq W$, the two sources of information are equivalent. For lower quality levels, $q_l < W$, advertising provides finer information than the word-of-mouth mechanism. In this case, advertising provides an exact observation on q_l whereas with the word-of-mouth mechanism consumers only learn that $q_l < W$. The lowest ability firms do not advertise because they do not want consumers to learn they are of low ability. In this case, the only information they send to consumers is the coarse information of not obtaining positive word-of-mouth. A firm of average ability is likely to have a quality level less than W. If the firm does not advertise and does not receive positive word-of-mouth, the coarse information sent to consumers will pool the average ability firm with the lowest ability firms. Average ability firms will advertise to send finer information on their ability levels to consumers. That is, they will advertise to separate themselves from the lowest ability firms. The highest ability firms separate themselves from average ability firms by not advertising. By not advertising, the highest ability firms pool with the lowest ability firms. High ability firms then rely on positive word-of-mouth to stochastically separate themselves from the lowest ability firms.

The next result analyzes a change in z, the cost of advertising. Increasing the cost of advertising decreases the lower cutoff ability level for advertising, L_1, and increases the upper cutoff ability level for advertising, H_1. That is, an increase in advertising costs actually increases the probability that firms will advertise.

Result 2: Let a = 0, b = 1, c = 0, d = 2, W = 2.75 and z = 0.4. There exists a countersignaling equilibrium in which $L_1 \approx 0.397627$ and $H_1 \approx 0.972671$.

In a countersignaling equilibrium, it is important to realize that the highest ability firms refrain from advertising to separate themselves from medium ability firms, not to save on advertising costs. If advertising costs increase, the expected net return to advertising decreases. When the expected net return to advertising decreases, high ability firms expect average ability firms to advertise less. High ability firms counter by advertising more. That is, H_1 increases. If more high ability firms start to advertise, however, the payoff from not advertising and not having positive word of mouth occur decreases since it is less likely that a high ability firm will have this outcome. As such, the return to not advertising decreases and leads to more average types advertising. That is, L_1 decreases.

A seemingly unintuitive result also occurs if it becomes easier to obtain positive word-of-mouth. If positive word-of-mouth becomes easier to obtain, the probability of advertising increases. The logic behind this result is the same as the intuition behind an increase in the cost of advertising. If obtaining positive word-of-mouth becomes easier, high types expect medium types to advertise less. To counter, high ability firms start to advertise, thus increasing H_1. As explained above, this leads to a decrease in L_1 and an overall increase in the probability a firm advertises.

Section IV: Discussion and Extensions

Discussion

The empirical literature on advertising as a signal of product quality offers a mixed bag of results. The work of Caves and Greene (1996) is representative of the cross sectional studies in the literature. In an analysis of nearly two hundred product categories, Caves and Greene find a positive correlation between advertising and quality for many products. However, they also find many cases of no correlation and even negative correlation between the two variables. While the current model cannot explain negative correlations between advertising and quality, it can explain both zero and positive correlations between the two variables. [51] Positive correlations between advertising and quality are predicted by the standard signaling framework. As seen above, a standard separating equilibrium exists in the current model. Countersignaling can explain a zero correlation between advertising and quality. In a countersignaling equilibrium, advertising is not monotonically increasing in quality. Thus, it is not surprising to find a zero correlation between the two variables. Industries with zero or near zero correlations between advertising and quality should be examined further for properties of countersignaling. Quadratic terms should be included in regressions to capture non-

[51] See Orzach, Overgaard and Tauman (2002) for a model that predicts a negative correlation between advertising and quality.

monotonicities. These industries could also be examined to determine if increased advertising costs increase the probability of advertising.

The model considered so far is static in the sense that a firm makes a one-time advertising decision. Nevertheless, it can still provide insight into advertising dynamics. In time series analyses, Thomas, Shane and Weigelt (1998) and Horstmann and MacDonald (2003) find that advertising intensities for high quality cars and compact disc players, respectively, exhibit an inverted "U" shape over time. That is, advertising expenditures are initially low, increase over time, reach a peak and then decline. A "static" countersignaling theory of advertising can explain this result if the cutoff point for obtaining positive word-of-mouth increase over time.

Currently a firm's quality level is given by (1) where $\varepsilon_t \sim U\,[c, c + d]$. Let a first generation firm continue to produce with this technology. The next time a firm produces, let a new generation of firms enter the market with the technology, $\bar{q}_t = \eta + \nu_t$. Let the prior on η be the same as first generation of firms, but assume $\nu_t \sim U\,[\bar{c}, \bar{c} + d]$, where $\bar{c} > c$. Note that for the same level of ability, ex ante quality produced by the new generation of firms first order stochastically dominates ex ante quality produced by the initial generation of firms. This assumption can be interpreted as the quality of newly introduced products improving over time.

On its own, the assumption above is not enough to change the behavior of a first generation firm. First generation firms are rewarded based on $E(\eta)$, not on relative quality. However, it is likely the case that as quality improvements occur over time, consumers will revise their notions of exceptionally high quality products upwards. That is, the cutoff for obtaining positive word-of-mouth is likely to increase. If this occurs, it becomes more difficult for a first generation firm (with its primitive technology) to obtain positive word-of-mouth.

Consider the case of a high quality first generation firm that initially countersignals and does not advertise. If the firm becomes *unable* to obtain word-of-mouth in subsequent time periods, it will no longer countersignal because it has no way of stochastically separating itself from low types. If the firm still wants consumers to further learn its ability level it will be forced to advertise in subsequent time periods. Over time, once ability is sufficiently learned, the firm will no longer find it profitable to advertise. Thus, for high quality firms in an environment where word-of-mouth becomes unobtainable, advertising intensity will exhibit an inverted "U" shaped pattern over time.

Fads

Consider a simple one round extension to the basic model outlined in Section II. Instead of the game ending at t = 3, let firms make another advertising decision at t = 4 (before q_3 is realized) and produce and sell another run of output at t = 5. Finding equilibria using backwards induction quickly leads to an intractable analysis. Instead the behavior of myopic firms that make advertising decisions in order to maximize expected profits in the next time period is considered. This simple extension can provide insight into the phenomena of fads.

The existing literature on fads has focused solely on consumer behavior. Explanations of fads have included sanctions on deviants, positive payoff externalities, conformity preference, communication and information cascades. The existing literature misses one key aspect of fads; they are ubiquitous. By definition, a fad occurs in the current model when information on a good spreads both through advertising and positive word-of-mouth. Thus, in the present model, firms fuel fads by advertising. Fads, as emphasized by Bikhchandani, Hirshleifer, and Welch (1992), are temporary.[52] It is now shown that after a fad occurs, an appealing equilibrium has the firm cut its advertising, thereby crushing the fad.

Fads can arise in both the separating and countersignaling equilibria described in the previous section. The analysis here proceeds on the assumption that a countersignaling equilibrium existed at t = 2 and that a fad occurred. In a countersignaling equilibrium, if a firm advertises and has positive word-of-mouth occur, the posterior on its ability (given in (8)) is $\eta \sim U\,[q_1 - c - d, H_1]$ where $q_1 \geq W$.

The firm produces another run of the good at t = 3 and must decide whether or not to advertise at t = 4. The analysis is a straightforward extension of the previous section and is contained in the appendix. Results and intuition are presented here:

Lemma 5: After a fad occurs at t = 2, there does not exist a separating equilibrium at t = 4.

Proof: See Appendix C.

In a standard signaling model, the expected net return to advertising would be greater for higher ability firms, allowing separation to occur. As seen in the previous section, the positive word-of-mouth mechanism decreases the net returns to ability in advertising and limits the parameter space for which a separating equilibrium exists. After a fad occurs, the posterior on a firm's ability is $\eta \sim U\,[q_1 - c - d, H_1]$ where $q_1 \geq W$. Given this belief,

---

[52] Bikhchandani, Hirshleifer, and Welch develop a theory of information cascades and summarize the literature on fads.

even the lowest ability firm ($\eta = q_1 - c - d$) can obtain positive word-of-mouth. This sufficiently lowers the net returns to ability in advertising so that separation cannot occur.

Lemma 6: After a fad occurs at t = 2, there does not exist a countersignaling equilibrium at t = 4.

Proof: See Appendix C.

In a countersignaling equilibrium, medium ability firms separate themselves from low ability firms by advertising. The ease of obtaining positive word-of-mouth lowers the net returns to ability in advertising to the point that medium ability firms do not find it profitable to advertise.

Lemma 7: After a fad occurs at t = 2, there exists a pooling equilibrium in which all firms advertise provided that advertising costs are sufficiently low, z

$$\leq \frac{1}{d}\left[-\frac{H_1 - (q_1 - c - d))^2}{4} + d\left(\frac{H_1 - (q_1 - c - d)}{2}\right) + \frac{(q_1 - W)^2}{2}\right]$$

Proof: See Appendix C.

The pooling equilibrium above is supported by the following off-equilibrium belief: If a firm refrains from advertising, it is believed to be the lowest ability firm that could produce q_3. The intuition behind Lemma 7 is straightforward. If advertising costs are sufficiently low, all firms will find it profitable to advertise.

Proposition 2: After a fad occurs at t = 2, there exists a pooling equilibrium in which no firm advertises provided that advertising costs are sufficiently high, z > $\frac{(2M - (q_1 - c - d))^2 + 2(W - c)(H_1 + (q_1 - c - d) - 2M) - H_1^2}{4d}$**, where**

$$M = \frac{3W(H_1 + (q_1 - c - d) - c(H_1 - (q_1 - c - d)) - 2(q_1 - d)^2 - 2H_1(H_1 + q_1 - d) + 2c^2}{3(2W - (H_1 + c) - (q_1 - d))}$$

Proof: See Appendix C.

The pooling equilibrium in which no firms advertise is supported by the following reasonable off-equilibrium belief: If a firm advertises, the market uses Bayes' Rule and the observation on q_3 to update its beliefs. If advertising costs are sufficiently high, no firm will find it profitable to advertise.

The pooling equilibrium that arises at $t = 4$ depends on z, the cost of advertising. The analysis here has assumed that a countersignaling equilibrium existed at $t = 2$. Thus, if advertising costs remain constant over time, the value of z at $t = 4$ must be a value that permitted a countersignaling equilibrium at $t = 2$. Using z from the countersignaling equilibria discussed in Result 1 and Result 2 (as well as in other numerical examples not presented in the current work), the unique equilibrium that occurs at $t = 4$ is the pooling equilibrium in which no firm advertises.

The intuition is as follows: In a countersignaling equilibrium, a firm reveals itself to be of "average" ability by advertising. The fact that the firm also benefits from positive word-of-mouth reveals that the firm is at the high-end of "average" ability firms. The two sources of information greatly reduce consumer uncertainty on the firm's ability level. With less uncertainty on the firm's ability level, it no longer needs to engage in costly advertising to allow consumers to further learn its level of ability. By definition, once advertising is stopped, the fad comes to an end.

Section V: Conclusion

The present work has extended the literature on countersignaling by developing a countersignaling theory of advertising. Advertising in the current framework is not just a conspicuous wasteful expenditure; it allows consumers as a whole to learn a firm's ability level. In addition to advertising, consumers also receive coarser information on a firm's ability to produce quality through a word-of-mouth mechanism. A countersignaling equilibrium exists in which average ability firms advertise while the highest and lowest ability firms refrain from advertising. High ability firms pool with low ability firms because they expect positive word-of-mouth to separate them from low types.

The present work opens two avenues for future research. First, the model could be extended to include repeat purchases. The main implication of repeat purchases is that if a firm does not advertise and does not obtain positive word-of-mouth, consumers can have different information sets. If consumers have different information sets, pricing becomes an issue. As in Linnemer (2002), optimal price and advertising decisions in an environment of informed and uniformed consumers could be examined. The key difference is that, in the current framework, the decision to advertise would determine whether or not consumers are informed or uninformed.

Second, a simple one round extension of the current model (similar to the fads extension) can explain the phenomena of "selling out." In a countersignaling equilibrium, the highest quality firms do not advertise, hoping instead to benefit from positive word–of–mouth. Some high ability firms will be unlucky and not receive positive word–of–mouth. In the subsequent time period, these unlucky high ability firms will be pooled with the lowest ability firms. It is likely that firms in the low-end of the high ability group will advertise in the following time period in order to separate themselves from the lowest ability firms. These firms "sell out" by reaching consumers through mainstream publicity. Note that this framework would provide the true dynamics to explain the time series results on advertising and quality found in Thomas, Shane and Weigelt (1998) and Horstmann and MacDonald (2003).

APPENDIX A

Proof of Lemma 1

The expected return of turnover, V, is given in (7). Taking the limit in (7) as $m \rightarrow m`$, it is easy to verify that V is continuous in m. Differentiating (7) with respect to m yields:

$$\frac{\partial V}{\partial m} = \begin{cases} \dfrac{y_1 - c - a + m' - m}{2d} & when \;\; m < m' \\ \\ \dfrac{y_1 - c - m'}{2d} & when \;\; m \geq m' \end{cases}$$

Both expressions above are positive and hence, V is increasing in m. ■

Proof of Proposition 1

The manager's turnover strategy, $m^{\sim}(\cdot)$, is defined by V ($m = m^{\sim}$, $m`$, $y_1$, z) = 0, where V is the expected return of turnover. Ultimately the analysis will focus on a fixed point where $m^{\sim} = m`$. As such, V is only examined for the case where $m \geq m`$. From (7), setting V = 0, letting $m = m^{\sim}$, and solving for $m^{\sim}$ yields:

$$m^{\sim} = \frac{2d}{y_1 - c - m'}[\, z + \frac{a + m'}{2} - \frac{(m' + y_1 - c)(m' - y_1 + c)}{4d} - \frac{y_1 - c + m'}{2}] \qquad \text{(A1)}$$

The goal of the analysis is to find a fixed point m*, where $m^{\sim} = m^* = m`$, that solves V = 0 such that $m^* \in [a, y_1 - c]$. The next expression provides fixed points that satisfy V = 0 (that is, satisfy (A1)):

$$m^* = \frac{2d}{y_1 - c - m^*}[\, z + \frac{a + m^*}{2} - \frac{(m^* + y_1 - c)(m^* - y_1 + c)}{4d} - \frac{y_1 - c + m^*}{2}]$$

Rearranging terms,

$$-(m^*)^2 + 2m^*(y_1 - c) - 4dz - y_1^2 - c^2 - 2ad + 2cy_1 + 2d(y_1 - c) = 0 \qquad \text{(A2)}$$

Using the quadratic formula to solve for m* yields:

$$m^* = (y_1 - c) - \sqrt{2d(y_1 - c - a - 2z)} \quad \text{(A3)}$$

When $y_1 \in [a + c, a + b + c]$, from (2) it must be the case that $m' \in [a, y_1 - c]$. Therefore, m^* must also be an element of $[a, y_1 - c]$. From (A3), if $z \in [\frac{y1 - c - a}{2} - \frac{(y1 - c - a)^2}{4d}, \frac{y1 - c - a}{2}]$, then $m^* \in [a, y_1 - c]$. Thus, for any $y_1 \in [a + c, a + b + c]$, if $z \in [\frac{y1 - c - a}{2} - \frac{(y1 - c - a)^2}{4d}, \frac{y1 - c - a}{2}]$, managers with ability equal to m^* will be indifferent between remaining with their firms and leaving their firms when the market belief is $m' = m^*$. From Lemma 1, managers with ability, $m \geq m^*$, will leave their firms and managers with ability, $m < m^*$, will stay with their firms. Hence, for a range of parameter values of z, a separating equilibrium exists.

Uniqueness of the equilibrium can be seen by differentiating the left-hand side of (A2) with respect to m^*. $\frac{\partial(A2)}{\partial m^*} = 2\,(y_1 - c - m^*) \geq 0$. Since the left-hand side of (A2) is monotone in m^*, the equilibrium is unique. ■

Proof of Proposition 2

The probability of turnover, P, is given in (8). Differentiating P with respect to z, d, a, c and y_1 provides the comparative statics results given in the proposition:

$$\text{(i): } \frac{\partial P}{\partial z} = -\frac{\partial m^*/\partial z}{y_1 - c - a} = -\frac{2d}{\sqrt{2d(y_1 - c - a - 2z)}}\ \frac{1}{y_1 - c - a} < 0$$

$$\text{(ii) } \frac{\partial P}{\partial d} = -\frac{\partial m^*/\partial d}{y_1 - c - a} = \frac{y_1 - c - a - 2z}{\sqrt{2d(y_1 - c - a - 2z)}}\ \frac{1}{y_1 - c - a} > 0$$

$$\text{(iii) } \frac{\partial P}{\partial a} = -\frac{\partial m^*/\partial a}{y_1 - c - a} + \frac{y_1 - c - m^*}{(y_1 - c - a)^2} = -\frac{d}{\sqrt{2d(y_1 - c - a - 2z)}}\ \frac{1}{y_1 - c - a}$$

$$+ \frac{y_1 - c - m^*}{(y_1 - c - a)^2} < 0$$

$$\text{(iv)} \ \frac{\partial P}{\partial c} = - \frac{1 + (\partial m^* / \partial c)}{y_1 - c - a} + \frac{y_1 - c - m^*}{(y_1 - c - a)^2} = - \frac{d}{\sqrt{2d(y_1 - c - a - 2z)}} \frac{1}{y_1 - c - a}$$

$$+ \frac{y_1 - c - m^*}{(y_1 - c - a)^2} < 0$$

$$\text{(v)} \ \frac{\partial P}{\partial y_1} = \frac{1 - (\partial m^* / \partial y_1)}{y_1 - c - a} - \frac{y_1 - c - m^*}{(y_1 - c - a)^2} = \frac{d}{\sqrt{2d(y_1 - c - a - 2z)}} \frac{1}{y_1 - c - a}$$

$$- \frac{y_1 - c - m^*}{(y_1 - c - a)^2} > 0$$

The sign of the comparative statics effects in (i) and (ii) are readily apparent. The sign of the comparative statics effects in (iii), (iv), and (v) depend on the magnitude of the following two terms:

$$\frac{d}{\sqrt{2d(y_1 - c - a - 2z)}} \frac{1}{y_1 - c - a} \text{ and } \frac{y_1 - c - m^*}{(y_1 - c - a)^2}$$

It is now shown that

$$\frac{d}{\sqrt{2d(y_1 - c - a - 2z)}} \frac{1}{y_1 - c - a} > \frac{y_1 - c - m^*}{(y_1 - c - a)^2} \quad \text{(A4)}$$

Obtaining like denominators and plugging in the value of m* reveals that the inequality in (A4) is true if

$$d\,(y_1 - c - a) > 2d\,(y_1 - c - a - 2z) \quad \text{(A5)}$$

Simplifying (A5), the inequality in (A4) holds if

$$(y_1 - c - a) < 4z \quad \text{(A6)}$$

The smallest possible value of z in equilibrium is $\frac{y1-c-a}{2}$- $\frac{(y1-c-a)^2}{4d}$. If the inequality in (A6) holds for the smallest possible value of z, then it will hold for all values of z. Substituting in the smallest possible value of z into (A6) and simplifying, the inequality in (A4) holds if

$$0 < (y_1 - c - a)\,\frac{d-(y_1-c-a)}{d} \qquad \text{(A7)}$$

The inequality in (A7) holds, making the inequality in (A4) true. Both terms on the right hand side are positive. The first term is the range of $(m \mid y_1)$. The second term is positive under the assumption of $d > b$. The term b is the range of the prior distribution on managerial ability. Since y_1 allows the market to get a more precise estimate on managerial ability, $b \geq (y_1 - c - a)$. If y_1 reaches its greatest value of $a + b + c$, the two expressions are equal. Thus, by transitivity, $d > (y_1 - c - a)$. ■

The Cases of Intermediate and High Values of First Period Output

The analysis so far cover has covered the low first period output case where $y_1 \in [a + c, a + b + c]$. The next propositions cover the cases when $y_1 \in [a + b + c, a + c + d]$ and when $\in [a + c + d, a + b + c + d]$. The analyses of these two cases are along the same lines as the case of low first period output and are not particularly instructive. As such, the equilibrium and comparative statics results are provided below, but no proofs are given.

For intermediate values of first period output:

Proposition 1a: Given a realization of y_1 such that, $(a + b + c) < y_1 < (a + c + d)$, and given $z \in [\frac{b}{2}-\frac{b^2}{4d}, \frac{b}{2}]$, there exists a unique separating equilibrium such that managers with ability, m $\geq$ m* = $(a+b)-\sqrt{2d(b-2z)}$, leave their firms and managers with ability, m < m*, remain with their firms.

Proposition 2a: Given the equilibrium in Proposition 1a, the probability of turnover, P = $\frac{a+b-m^*}{b}$, is (i) decreasing in z, (ii) increasing in d and (iii) increasing in b.

For high values of first period output:

Proposition 1b: Given a realization of y_1 such that, (a + c + d) $\leq y_1 \leq$(a + b + c + d), and given $z \in [\frac{(a+b)-(y_1-(c+d))}{2} - \frac{((a+b)-(y_1-(c+d)))^2}{4d}, \frac{(a+b)-(y_1-(c+d))}{2}]$, there exists a unique separating equilibrium such that managers with ability, m $\geq$m* = $(a+b)-\sqrt{2d((a+b)-(y_1-(c+d))-2z)}$, leave their firms and managers with ability, m < m*, remain with their firms.

Proposition 2b: Given the equilibrium in Proposition 1b, the probability of turnover, P = $\frac{a+b-m^*}{(a+b)-(y_1-(c+d))}$, is (i) decreasing in z, (ii) increasing in d, (iii) increasing in a, (iv) increasing in b, (v) increasing in c and (vi) decreasing in y_1.

APPENDIX B

Proof of Proposition 1

The proof of this proposition will be in two parts. First, it will be demonstrated that there exist pooling equilibria in which all managers invest in the risky project. Next, uniqueness of this class of equilibria will be argued.

Part One:

Consider a pooling equilibrium in which all managers invest in the risky project supported by the following off-equilibrium belief: Any manager investing in the safe project is believed to have ability, m < S.[53]

Due to the option the safe project provides, managers with ability, m < S, are worth S to the firm and managers with ability, m ≥S, are worth m to the firm. Thus, a manager who invests in the safe project earns a wage of S at t = 2.

Now consider a manager who invests in the risky project: The posterior on m after the return on the risky project is observed is

$$(m \mid R) \sim \begin{cases} U[a, R+c] & \text{if} \quad R \in [a-c, b-c] \\ U[a, b] & \text{if} \quad R \in (b-c, a+c) \\ U[R-c, b] & \text{if} \quad R \in [a+c, b+c] \end{cases} \qquad \text{(B1)}$$

Given the option-like nature of the projects, the value (W) of the reputational states given in (B1) is as follows:

(i) If (m | R) ~ U[a, x], where x ≤S, then

W = S

(ii) If (m | R) ~ U[a, x], where x > S, then

[53] More specific off-equilibrium beliefs that are consistent with m < S, such as, m = a, will also support the equilibrium of all managers investing in the risky project.

$$W = \frac{S-a}{x-a}S + \frac{1}{x-a}\int_S^x m\,dm = \frac{S^2 - 2aS + x^2}{2(x-a)} > S \tag{B2}$$

(iii) If (m | R) ~ U[x, b], where x < S, then

$$W = \frac{S-x}{b-x}S + \frac{1}{b-x}\int_S^b m\,dm = \frac{S^2 - 2xS + b^2}{2(b-x)} > S$$

(iv) If (m | R) ~ U[x, b], where x ≥S, then

$$W = \frac{x+b}{2} > S$$

It is easy to verify that W, the payoff of investing in the risky project, is increasing in managerial ability.[54] Even the lowest ability manager, m = a, can achieve the first two states given in (B1) and (B2). Thus, the lowest ability manager strictly prefers to invest in the risky project and therefore all managers prefer to invest in the risky project. ■

Part Two:

The class of equilibria in which all managers invest in the risky project is unique. There does not exist a pooling equilibrium in which all managers invest in the safe project, nor do separating equilibria exist.

Consider a candidate equilibrium in which all managers invest in the safe project. If a manager invests in the safe project, the market cannot update its beliefs on managerial ability. That is, m ~ U[a, b]. Thus, the value of a manager investing in the safe project is determined by (ii) in (B2) where x = b.

What off-equilibrium beliefs should be used to support the candidate equilibrium? If a manager deviates from the candidate equilibrium by investing in the risky project, Bayes' Rule *does* provide a guide on how to rationally update beliefs. The return on the risky

[54] See Lemma 1 for a similar proof.

project provides another observation on managerial ability. Accordingly, the market should use Bayes' Rule and the return on the risky project to update its prior belief.

Consider the situation of a manager with the highest ability, m = b. If the manager invests in the risky project, she can only achieve the last two states in (B1) and (B2). Thus, investing in the risky project would lead to a greater payoff than investing in the safe project. Since the highest ability manager wishes to deviate from the candidate equilibrium, there does not exist a pooling equilibrium in which all managers invest in the safe project.

Next, consider a candidate equilibrium in which managers with ability, $m \geq m'$, invest in the risky project and managers with ability, $m < m'$, invest in the safe project. In such a separating equilibrium, the market belief on managerial ability based solely on project choice would be $(m \mid \text{safe}) \sim U[a, m']$ and $(m \mid \text{risky}) \sim U[m', b]$. Consider a manager with ability, $m = m' - \varepsilon$, where ε is an arbitrarily small positive number. In the candidate equilibrium, such a manager would invest in the safe project and no further updating would take place on beliefs. The belief on managerial ability would be $(m \mid \text{safe}) \sim U[a, m']$.

What would occur if the manager deviated from the proposed equilibrium and invested in the risky project? With probability $1 - \varepsilon$, the deviation would not be detected and the manager would earn a payoff consistent with $(m \mid \text{risky}, R) \sim U[m', z]$, where $z \leq b$. Such a payoff is clearly greater than the payoff associated with $m \sim U[a, m']$. With probability ε, the deviation would be detected. In this case, it is assumed that the market updates its beliefs solely on project returns and not on project choice (as its belief on project choice has been proven incorrect). The manager would earn a payoff consistent with U[a, m` – y], where y $\leq \varepsilon$. As ε is an arbitrarily small positive number, the manager is better off investing in the risky project. Since the manager wishes to deviate from the candidate equilibrium, there does not exist a separating equilibrium in which managers above a certain cutoff level of ability invest in the risky project and managers with ability below the cutoff invest in the safe project.

The argument above generalizes to rule out *any* equilibrium that specifies all managers above a certain cutoff level of ability invest in the risky project. The intuitive criterion of Cho and Kreps (1987) can be used to rule out *any* equilibrium in which the highest ability manager invests in the safe project. These two arguments combine to rule out all possible equilibria except for the class of equilibria in which all managers invest in the risky project. ■

Proof of Lemma 1

Differentiating V, given in (5), with respect to m yields:

$$\frac{\partial V}{\partial m} = \frac{1}{2c}\left[\frac{b-S}{2} - F(T(m+c)^{*} - T(m-c)^{*})\right] \qquad \text{(B3)}$$

If $T(R)^{*}$ is monotonically nonincreasing in R, then the expression $T(m+c)^{*} - T(m-c)^{*} \leq 0$ and V must be increasing in m. It should be noted that $T(R)^{*}$ being monotonically nonincreasing in R is sufficient, not necessary, for V to be increasing in m. ■

Proof of Proposition 2

Suppose that firms use a monotonically nonincreasing termination schedule that satisfies the following condition:

$$\text{P} = \int_{S-c}^{S+c} T(R)^{*}\, dR = \frac{-(S^2/4) + S((b/2)-c) + bc - (b^2/4)}{F} \qquad \text{(B4)}$$

If the market believes that firms employ an optimal termination schedule, then the value of investing in the risky project relative to the safe project, V, is given in (5). For a manager with ability, m = S, V is given by:

$$V = \frac{1}{2c}[S(\frac{b-S}{2}) + c(b+S) + \frac{S^2}{4} - \frac{b^2}{4} - F\int_{S-c}^{S+c} T(R)^* dR] - S \quad \text{(B5)}$$

Substituting (B4) into (B5), it is easy to verify that if firms choose a termination schedule that satisfies the condition given in (B4), then a manager with ability, m = S, would be indifferent between investing in the safe project and investing in the risky project. From Lemma 1, it follows that all managers with ability, m ≥S, will invest in the risky project and all managers with ability, m < S, will invest in the safe project. That is, $m^* = S$.

Note that the maximum possible value of $T(R)^*$ for any value of R is one. Thus, termination can only be effective when $\frac{-(S^2/4) + S((b/2)-c) + bc - (b^2/4)}{F}$ (from B4) is less than 2c. Rearranging terms, termination can only be effective when F $> \frac{b-S}{2} - \frac{(b-S)^2}{8c}$. ■

Proof of Proposition 3

The claims made in this proposition can easily be verified by differentiating the right hand side in (B4) with respect to F, c, b and S:

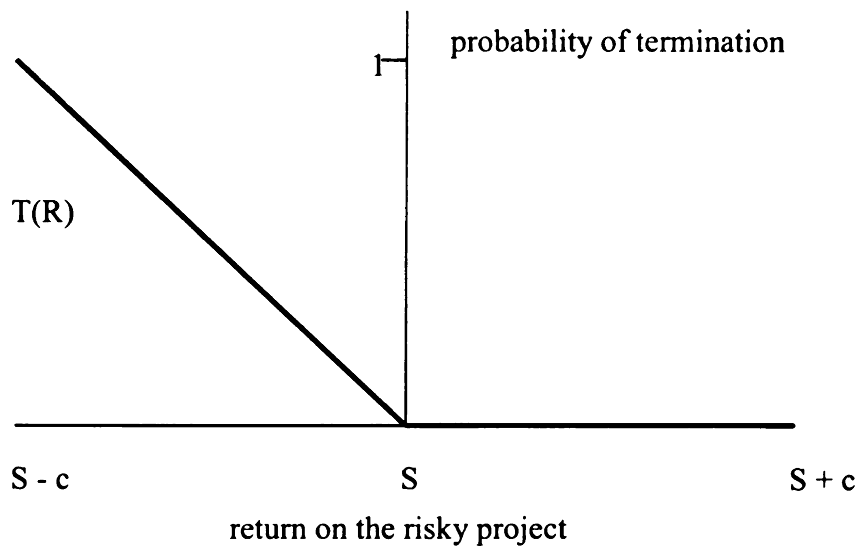
(i) $\frac{\partial P}{\partial F} = -(\frac{-(S^2/4) + S((b/2)-c) + bc - (b^2/4)}{F^2}) < 0$

(ii) $\frac{\partial P}{\partial c} = \frac{b-S}{F} > 0$

(iii) $\dfrac{\partial P}{\partial b} = \dfrac{2c-(b-s)}{2F} > 0$

(iv) $\dfrac{\partial P}{\partial S} = \dfrac{(b-S)-2c}{2F} < 0 \quad \blacksquare$

Figure 1



An Optimal Nonlinear Termination Schedule

APPENDIX C

Proof of Lemma 1

Lemma 1 claims that $\partial V/\partial\eta > 0$ when $\eta < W - c - d$ and $\partial V/\partial\eta < 0$ when $\eta \geq W - c - d$. V can be determined by subtracting (3) and (4) from (6) and (7). Doing so yields:

$$V = \frac{1}{d}\left[\int_{\eta+c}^{\eta^*+c} \frac{a+q_1-c}{2}dq_1 + \int_{\eta^*+c}^{b+c} \frac{\eta^*+q_1-c}{2}dq_1 + \int_{b+c}^{\eta+c+d} \frac{\eta^*+b}{2}dq_1\right] - \frac{a+\eta^*}{2} - z$$

$$\text{when } \eta < \eta^* \qquad \text{(C1)}$$

$$V = \frac{1}{d}\left[\int_{\eta+c}^{b+c} \frac{\eta^*+q_1-c}{2}dq_1 + \int_{b+c}^{\eta^*+c+d} \frac{\eta^*+b}{2}dq_1 + \int_{\eta^*+c+d}^{\eta+c+d} \frac{q_1-c-d+b}{2}dq_1\right]$$

$$- \frac{a+\eta^*}{2} - z \qquad \text{when } \eta^* \leq \eta < W - c - d \qquad \text{(C2)}$$

$$V = \frac{1}{d}\left[\int_{\eta+c}^{b+c} \frac{\eta^*+q_1-c}{2}dq_1 + \int_{b+c}^{\eta^*+c+d} \frac{\eta^*+b}{2}dq_1 + \int_{\eta^*+c+d}^{\eta+c+d} \frac{q_1-c-d+b}{2}dq_1\right]$$

$$- \frac{1}{d}\left[\int_{\eta+c}^{W} \frac{a+\eta^*}{2}dq_1 + \int_{W}^{\eta+c+d} \frac{q_1-c-d+b}{2}dq_1\right] - z$$

$$\text{when } W - c - d \leq \eta \qquad \text{(C3)}$$

From (C1) – (C3), it is easy to verify that V is continuous in η. Differentiating V with respect to η yields:

$$\partial V/\partial\eta = \begin{cases} \dfrac{(b-a)+(\eta^*-\eta)}{2d} \text{ when } \eta < \eta^* \\ \dfrac{b-\eta^*}{2d} \text{ when } \eta^* \le \eta < W-c-d \\ \dfrac{a-\eta}{2d} \text{ when } \eta \ge \text{W - c - d} \end{cases}$$

Given the assumptions on parameters, it is clear that $\partial V/\partial\eta > 0$ when $\eta < W - c - d$ and $\partial V/\partial\eta < 0$ when $\eta \ge W - c - d$. ■

Proof of Proposition 1

In a separating equilibrium it must be the case that $V \ge 0$ for firms with ability, $\eta \ge \eta^*$, and $V < 0$ for firms with ability, $\eta < \eta^*$. From Lemma 1 it follows that if a separating equilibrium is to exist, V must equal zero when $\eta = \eta^*$. Carrying out the integration in (C2), letting $\eta = \eta^*$, and setting V = 0 yields:

$$\frac{1}{d}\left[\frac{\eta^* b + d(b-a)}{2} - \frac{\eta^{*2} + b^2}{4}\right] - z = 0 \qquad \text{(C4)}$$

Solving (C4) for η^* and bearing in mind that $\eta^* \in [a, W - c - d)$ yields:

$$\eta^* = b - \sqrt{2d(b-a-2z)} \qquad \text{(C5)}$$

Thus, if η^* satisfies (C5), V = 0 when $\eta = \eta^*$. If (C5) is satisfied, it follows from Lemma 1 that $V < 0$ for firms with ability, $\eta < \eta^*$, and $V > 0$ for firms with ability, $\eta^* \le \eta < W - c - d$. When (C5) is satisfied it is not guaranteed that $V > 0$ for $\eta \ge W - c - d$ since $\partial V/\partial\eta < 0$ when $\eta \ge W - c - d$. In fact, if η^* satisfies (C5), as $\eta^* \to W - c - d$, $V < 0$ for firms with the highest ability level, $\eta = b$. A condition is now to found to ensure that when (C5) is satisfied, $V \ge 0$ for firms with ability, $\eta = b$.

Carrying out the integration in (C3) and letting $\eta = b$ yields:

$$V = \frac{1}{d}[(W-c)(\frac{b-\eta^{*}-a}{2}) + \frac{(W-c-d)^2+\eta^{*2}}{4} + \frac{d\eta^{*}+ab-b^2}{2}] - z$$

when $\eta = b$ (C6)

In a separating equilibrium, η^* must satisfy (C5). Solving (C5) for z and substituting into (C6) yields the expected net return to advertising for a firm with ability, $\eta = b$, when $V = 0$ for firms with ability, $\eta = \eta^*$.

$$V = \frac{1}{d}[(\frac{W-c-d}{2})(b+\frac{W-c-d}{2}-a-\eta^{*}) - (b)(\frac{\eta^{*}-a}{2}) + \frac{\eta^{*2}}{2} - \frac{b^2}{4}] \quad \text{(C7)}$$

In (C7), $\partial V/\partial\eta^* < 0$ since $\eta^* < W - c - d < b$. That is, the expression for V in (C7) is monotonically decreasing in η^*. Setting $V = 0$ in (C7) and solving for η^* yields the expression for η_{max} given in the proposition. Thus, when $V = 0$ for firms with ability, $\eta = \eta^*$, η_{max} provides the largest possible value of η^* that guarantees V is nonnegative for firms with ability, $\eta = b$.

In summary, if η^* satisfies the expression in (C5) and $\eta^* \in [a, \eta_{max}]$, then $V = 0$ for a firm with ability, $\eta = \eta^*$ and $V \geq 0$ for a firm with ability, $\eta = b$. It follows from Lemma 1 that a separating equilibrium exists under these conditions. From (C5), if

$z \in [\frac{2d(b-a)-(b-a)^2}{4d}, \frac{2d(b-a)-(b-\eta_{max})^2}{4d}]$, then $\eta \in [a, \eta_{max}]$. ■

Proof of Lemma 2

Let the prior on η be given by (12). For the technology given in (1), the probability distribution of q_1, $g(q_1)$, is given by:

$$
g(q_1) \sim \begin{cases}
\dfrac{x-(a+c)}{d(L_1-a+b-H_1)} \text{ if } x \in [a+c, L_1+c) \\
\dfrac{L_1-a}{d(L_1-a+b-H_1)} \text{ if } x \in [L_1+c, H_1+c] \\
\dfrac{x-(H_1+c)+(L_1-a)}{d(L_1-a+b-H_1)} \text{ if } x \in (H_1+c, b+c) \\
\dfrac{1}{d} \text{ if } x \in [b+c, a+c+d] \\
\dfrac{1}{d} - \dfrac{x-(a+c+d)}{d(L_1-a+b-H_1)} \text{ if } x \in (a+c+d, L_1+c+d) \\
\dfrac{b-H_1}{d(L_1-a+b-H_1)} \text{ if } x \in [L_1+c+d, H_1+c+d] \\
\dfrac{(b-H_1)-(x-(H_1+c+d))}{d(L_1-a+b-H_1)} \text{ if } x \in (H_1+c+d, b+c+d]
\end{cases}
$$

Assume for a moment that the exact value of q_1 is always observable even if a firm refrains from advertising. Bayes' Rule can be used to derive the following posteriors on a firm's ability level.

$$
f(\eta \mid \text{no advertise}, q_1) = \begin{cases}
\dfrac{1}{(q_1-c)-a} \text{ if } q_1 \in [a+c, L_1+c) \\
\dfrac{1}{L_1-a} \text{ if } q_1 \in [L_1+c, H_1+c] \\
\dfrac{1}{(L_1-a)+(q_1-c-H_1)} \text{ if } q_1 \in (H_1+c, b+c) \\
\dfrac{1}{(L_1-a)+(b-H_1)} \text{ if } q_1 \in [b+c, a+c+d] \\
\dfrac{1}{(L_1-(q_1-c-d))+(b-H_1)} \text{ if } q_1 \in (a+c+d, L_1+c+d) \\
\dfrac{1}{b-H_1} \text{ if } q_1 \in [L_1+c+d, H_1+c+d] \\
\dfrac{1}{b-(q_1-c-d)} \text{ if } q_1 \in (H_1+c+d, b+c+d]
\end{cases}
$$

The first two and last two posteriors given in $f(\eta \mid$ no advertise, $q_1)$ represent uniform distributions on the intervals $[a, q_1 - c]$, $[a, L_1]$, $[b, H_1]$ and $[q_1 - c - d, b]$, respectively. The third posterior distribution given represents a proper uniform distribution on the interval [a, q_1 - c] with support missing on the interval [L_1, H_1]. The fourth posterior distribution given is the prior on η given in (12). The fifth posterior distribution given represents a proper uniform distribution on the interval [q_1 – c – d, b] with support missing on the interval [L_1, H_1]. The means of the third, fourth, and fifth posterior distributions given above are $\frac{L_1 - a^2 + (q_1 - c)^2 - H_1^2}{2(L_1 - a + (q_1 - c) - H_1)}$, $\frac{L_1 - a^2 + b^2 - H_1^2}{2(L_1 - a + b - H_1)}$ and $\frac{L_1 - (q_1 - c - d)^2 + b^2 - H_1^2}{2(L_1 - (q_1 - c - d) + b - H_1)}$, respectively.

Of course, if a firm does not advertise and does not obtain positive word-of-mouth, q_1 is not exactly observed by consumers. All consumers learn is that $q_1 < W$. Profits from this event occurring can be determined by:

$$R = \frac{\int_{a+c}^{W} g(q_1)\left(\int (x) f(\eta \mid \text{no advertise}, q_1)dx\right)dq_1}{G(q_1 < W)}$$

Using g(q_1) and f(η| no advertise, q_1) from above, the numerator of R is given by:

$$\text{Numerator} = \int_{a+c}^{L_1+c} \left(\frac{x-(a+c)}{d(L_1 - a + b - H_1)}\right)\left(\frac{a+x-c}{2}\right)dx +$$

$$\int_{L_1+c}^{H_1+c} \left(\frac{L_1 - a}{d(L_1 - a + b - H_1)}\right)\left(\frac{a+L_1}{2}\right)dx$$

$$+ \int_{H_1+c}^{b+c} \left(\frac{x-(H_1+c)+(L_1-a)}{d(L_1 - a + b - H_1)}\right)\left(\frac{L_1^2 - a^2 + (x-c)^2 - H_1^2}{2(L_1 - a + (x-c) - H_1)}\right)dx$$

$$+ \int_{b+c}^{a+c+d} \left(\frac{1}{d}\right)\left(\frac{L_1^2 - a^2 + b^2 - H_1^2}{2(L_1 - a + b - H_1)}\right)dx \qquad +$$

$$\int_{a+c+d}^{L_1+c+d} \left(\frac{1}{d} - \frac{x-(a+c+d)}{d(L_1-a+b-H_1)}\right)\left(\frac{L_1^2-(x-c-d)^2+b^2-H_1^2}{2(L_1-(q_1-c-d)+b-H_1)}\right)dx$$

$$+ \int_{L_1+c+d}^{W} \left(\frac{b-H_1}{d(L_1-a+b-H_1)}\right)\left(\frac{b+H_1}{2}\right)dx$$

$$= \frac{3(b^2-H_1^2)(W-c-d)+3d(L_1^2-a^2+b^2-H_1^2)-2(b^3-H_1^3)}{6d(L_1-a+b-H_1)}$$

Using $g(q_1)$ above, $G\ (q_1 < W)$ is given by:

$$G\ (q_1 < W) = \int_{a+c}^{L_1+c} \frac{x-(a+c)}{d(L_1-a+b-H_1)}dx + \int_{L_1+c}^{H_1+c} \frac{L_1-a}{d(L_1-a+b-H_1)}dx$$

$$+ \int_{H_1+c}^{b+c} \frac{x-(H_1+c)+(L_1-a)}{d(L_1-a+b-H_1)}dx + \int_{b+c}^{a+c+d} \frac{1}{d}dx$$

$$+ \int_{a+c+d}^{L_1+c+d} \left(\frac{1}{d} - \frac{x-(a+c+d)}{d(L_1-a+b-H_1)}\right)dx +$$

$$\int_{L_1+c+d}^{W} \frac{b-H_1}{d(L_1-a+b-H_1)}dx$$

$$= \frac{2(b-H_1)(W-c-d)+2d(L_1-a+b-H_1)-(b^2-H_1^2)}{2d(L_1-a+b-H_1)}$$

Substituting the expressions for the numerator and denominator into R and simplifying yields the value of R given in the lemma. ■

Proof of Lemma 3

Lemma 3 claims that $\partial V/\partial\eta > 0$ when $\eta < W - c - d$ and $\partial V/\partial\eta < 0$ when $\eta \geq W - c - d$. V can be determined by subtracting (14)-(16) from (9)-(11). Doing so yields:

$$V = \frac{1}{d}\left[\int_{\eta+c}^{L_1+c} \frac{a+q_1-c}{2}dq_1 + \int_{L_1+c}^{H_1+c} \frac{L_1+q_1-c}{2}dq_1 + \int_{H_1+c}^{\eta+c+d} \frac{L_1+H_1}{2}dq_1\right] - z - R$$

when $\eta < L_1$ (C8)

$$V = \frac{1}{d}\left[\int_{\eta+c}^{H_1+c} \frac{L_1+q_1-c}{2}dq_1 + \int_{H_1+c}^{L_1+c+d} \frac{L_1+H_1}{2}dq_1 + \int_{L_1+c+d}^{\eta+c+d} \frac{q_1-c-d+H_1}{2}dq_1\right] - z$$

$$- R$$

when $L_1 \le \eta < W - c - d$ (C9)

$$V = \frac{1}{d}\left[\int_{\eta+c}^{H_1+c} \frac{L_1+q_1-c}{2}dq_1 + \int_{H_1+c}^{L_1+c+d} \frac{L_1+H_1}{2}dq_1 + \int_{L_1+c+d}^{\eta+c+d} \frac{q_1-c-d+H_1}{2}dq_1\right] - z$$

$$- \frac{1}{d}\left[\int_{\eta+c}^{W} R\,dq_1 + \int_{W}^{\eta+c+d} \frac{H_1+b}{2}dq_1\right]$$

when $W - c - d \le \eta \le H_1$ (C10)

$$V =$$

$$\frac{1}{d}\left[\int_{\eta+c}^{L_1+c+d} \frac{L_1+H_1}{2}dq_1 + \int_{L_1+c+d}^{H_1+c+d} \frac{q_1-c-d+H_1}{2}dq_1 + \int_{H_1+c+d}^{\eta+c+d} \frac{q_1-c-d+b}{2}dq_1\right] - z$$

$$- \frac{1}{d}\left[\int_{\eta+c}^{W} R\,dq_1 + \int_{W}^{H_1+c+d} \frac{H_1+b}{2}dq_1 + \int_{H_1+c+d}^{\eta+c+d} \frac{q_1-c-d+b}{2}dq_1\right]$$

when $H_1 < \eta$ (C11)

From (C8) – (C11), it is easy to verify that V is continuous in η. Differentiating V with respect to η yields:

$$\partial V/\partial \eta = \begin{cases} \dfrac{(L_1-\eta)+(H_1-a)}{2d} \text{ when } \eta < L_1 \\ \dfrac{H_1-L_1}{2d} \text{ when } L_1 \le \eta < W-c-d \\ \dfrac{R}{d} - \dfrac{b+L_1}{2d} \text{ when } W-c-d \le \eta \le H_1 \\ \dfrac{R}{d} - \dfrac{L_1+H_1}{2d} \text{ when } H_1 < \eta \end{cases}$$

Given the assumptions on parameters, it is clear that $\partial V/\partial \eta > 0$ when $\eta < W - c - d$. If $R < (L_1 + H_1)/2$, $\partial V/\partial \eta < 0$ when $\eta > H_1$. Since $b > H_1$, if $R < (L_1 + H_1)/2$, $\partial V/\partial \eta$ will be negative when $W - c - d \le \eta \le H_1$ as well. ■

Proof of Lemma 4

A countersignaling equilibrium exists if $V \ge 0$ for firms with ability, $L_1 \le \eta \le H_1$, and if $V < 0$ for firms with ability $\eta < L_1$ and for firms with ability, $\eta > H_1$. If the first condition given in the lemma holds, the result derived in Lemma 3 is valid. For existence of a countersignaling equilibrium, it follows from Lemma 3 that V must equal zero for firms with ability, $\eta = L_1$, and for firms with ability, $\eta = H_1$. Carrying out the integration in (C9), letting $\eta = L_1$, and setting V equal to zero provides condition (ii) in the lemma. Carrying out the integration in (C10), letting $\eta = H_1$, and setting V equal to zero provides condition (iii) in the lemma. ■

Proof of Lemma 5

After a fad occurs the posterior on a firm's ability is $\eta \sim U\,[q_1 - c - d, H_1]$. Let the market have the following beliefs: firms with ability, $\eta \ge S$, advertise; firms with ability, $\eta < S$, do not advertise. In a separating equilibrium, $V \ge 0$ for firms with ability, $\eta \ge S$, and $V < 0$ for firms with ability, $\eta < S$. If a firm advertises, the following reputational states can occur:

$$(\eta|\text{ advertise}, q_3) \sim \begin{cases} U[q_1 - c - d, q_3 - c] \text{ if } q_3 \in [q_1 + d, S + c) \\ U[S, q_3 - c] \text{ if } q_3 \in [S + c, H_1 + c) \\ U[S, H_1] \text{ if } q_3 \in [H_1 + c, S + c + d] \\ U[q_3 - c - d, H_1] \text{ if } q_3 \in (S + c + d, H_1 + c + d] \end{cases} \quad \text{(C12)}$$

The first reputational state given in (C12) is off the candidate equilibrium path. In this case, signaling is ignored and beliefs are simply updated using Bayes' Rule and the observation on q_3.

Given market beliefs, if a firm does not advertise and obtains positive word-of-mouth, the following reputational states can occur:

$$(\eta|\text{ no advertise}, q_3) \sim \begin{cases} U[q_1 - c - d, S] \text{ if } q_3 \in [W, q_1] \\ U[q_3 - c - d, S] \text{ if } q_3 \in (q_1, S + c + d] \\ U[q_3 - c - d, H_1] \text{ if } q_3 \in (S + c + d, H_1 + c + d] \end{cases} \quad \text{(C13)}$$

Note that the last reputational state given in (C13) is off the candidate equilibrium path. Once again, in this case beliefs are updated using Bayes' Rule and the observation on q_3.

Consider a firm with ability, $\eta \geq S$. For the firm, the expected return to advertising is given by:

$E(\pi|\text{ advertise}, q_3) =$

$$\frac{1}{d}\left[\int_{\eta+c}^{H_1+c} \frac{S + q_3 - c}{2} dq_3 + \int_{H_1+c}^{S+c+d} \frac{S + H_1}{2} dq_3 + \int_{S+c+d}^{\eta+c+d} \frac{q_3 - c - d + H_1}{2} dq_3\right] - z$$

(C14)

Given market beliefs, let L denote the payoff a firm receives when it does not advertise and does not obtain positive word-of-mouth. For a firm with ability, η ≥S, the expected return to not advertising is given by:

E(π| no advertise) =

$$\frac{1}{d}\Big[\int_{\eta+c}^{W} L\,dq_3 + \int_{W}^{q_1} \frac{q_1 - c - d + S}{2}dq_3 + \int_{q_1}^{S+c+d} \frac{q_3 - c - d + S}{2}dq_3 + \int_{S+c+d}^{\eta+c+d} \frac{q_3 - c - d + H_1}{2}dq_3\Big] \tag{C15}$$

Subtracting (C15) from (C14) would yield, V, the expected net return to advertising for a firm with ability, η ≥S. Differentiating V with respect to η yields:

$$\partial V/\partial \eta = \frac{1}{d}[L - \frac{S+\eta}{2}] < 0 \tag{C16}$$

Given parameter assumptions, (S+η)/2 >(q_1-c-d+S)/2. The result in (C16) holds because (q_1-c-d+S)/2 > L. For the given prior on η, let g(q_3) represent the probability distribution of q_3. Let G denote the cumulative density function of g(q_3). Let f(η| no advertise, q_3) represent the posterior on η if q_3 were always exactly observed when the firm refrained from advertising. It follows that L is given by:

$$L = \frac{\int_{q_1-d}^{W} g(q_3)(\int (x) f(\eta \mid \text{no advertise}, q_3)dx)dq_3}{G(q_3 < W)}$$

Due to the martingale property discussed in Holmstrom (1999), it is the case that

$$\frac{q_1 - c - d + S}{2} = \frac{\int_{q_1-d}^{S+c+d} g(q_3)(\int (x) f(\eta \mid \text{no advertise}, q_3)dx)dq_3}{G(q_3 \le S + c + d)}$$

Since S + c + d > W, it follows that $\frac{q_1 - c - d + S}{2} > L$. Thus, explaining the result in (C16).

Given the result in (C16), there does not exist a separating equilibrium. In a separating equilibrium, V ≥ 0 for firms with ability, $\eta \geq S$, and V < 0 for firms with ability, $\eta < S$. V is continuous in ability, so in a separating equilibrium, it must be the case that V = 0 for a firm with ability, $\eta = S$. Since, $\partial V/\partial\eta < 0$ when $\eta \geq S$, there does not exist a separating equilibrium. ■

Proof of Lemma 6

After a fad occurs the posterior on a firm's ability is $\eta \sim U\,[q_1 - c - d, H_1]$. Let the market have the following beliefs: Firms with ability, $\eta > H_2$, and firms with ability, $\eta < L_2$, do not advertise; firms with ability, $L_2 \leq \eta \leq H_2$, advertise. In a countersignaling equilibrium, V ≥ 0 for firms with ability, $L_2 \leq \eta \leq H_2$; V < 0 for firms with ability, $\eta < L_2$, and for firms with ability, $\eta > H_2$.

Consider a firm with ability, $\eta \geq L_2$. If the firm advertises, the following reputational states can occur:

$$(\eta|\text{advertise}, q_3) \sim \begin{cases} U[L_2, q_3 - c] \text{ if } q_3 \in [L_2 + c, H_2 + c) \\ U[L_2, H_2] \text{ if } q_3 \in [H_2 + c, L_2 + c + d] \\ U[q_3 - c - d, H_2] \text{ if } q_3 \in (L_2 + c + d, H_2 + c + d] \\ U[q_3 - c - d, H_1] \text{ if } q_3 \in (H_2 + c + d, H_1 + c + d] \end{cases} \quad \text{(C17)}$$

The first three states given in (C17) can be obtained by firms with ability, $L_2 \leq \eta \leq H_2$. The last three states given in (C17) can be obtained by firms with ability, $\eta > H_1$. The last state given in (C17) is off the candidate equilibrium path. In this case, countersignaling is ignored and beliefs are simply updated using Bayes' Rule and the observation on q_3.

Thus, the expected return to advertising is given by:

E(π| advertise, q_3) =

$$\frac{1}{d}\Big[\int_{\eta+c}^{H_2+c}\frac{L_2+q_3-c}{2}dq_3+\int_{H_2+c}^{L_2+c+d}\frac{L_2+H_2}{2}dq_3+\int_{L_2+c+d}^{\eta+c+d}\frac{q_3-c-d+H_2}{2}dq_3\Big]-z$$

when $L_2 \leq \eta \leq H_2$ (C18)

$$E(\pi|\text{ advertise}, q_3) = \frac{1}{d}\Big[\int_{\eta+c}^{L_2+c+d}\frac{L_2+H_2}{2}dq_3+\int_{L_2+c+d}^{H_2+c+d}\frac{q_3-c-d+H_2}{2}dq_3 + \int_{H_2+c+d}^{\eta+c+d}\frac{q_3-c-d+H_1}{2}dq_3\Big]-z$$

when $\eta > H_2$ (C19)

If a firm with ability, $\eta \geq L_2$, does not advertise and obtains positive word-of-mouth, the following posteriors on η can arise:

$$f(\eta|\text{ no advertise}, q_3) = \begin{cases} \dfrac{1}{(L_2-(q_1-c-d)+(H_1-H_2))} \text{ if } q_3 \in [W, q_1] \\ \dfrac{1}{(L_2-(q_3-c-d)+(H_1-H_2))} \text{ if } q_3 \in [q_1, L_2+c+d] \\ \dfrac{1}{H_1-H_2} \text{ if } q_3 \in (L_2+c+d, H_2+c+d] \\ \dfrac{1}{H_1-(q_3-c-d)} \text{ if } q_3 \in (H_2+c+d, H_1+c+d] \end{cases} \quad \text{(C20)}$$

Note that the last two posterior distributions given in (C20) are uniform distributions on the intervals [H_2, H_1] and [$q_3 - c - d$, H_1], respectively. The first posterior distribution in (C20) is a proper uniform distribution on the interval [$q_1 - c - d$, H_1] with support missing on [L_2, H_2]. The second posterior distribution given is a proper uniform dis

Th

ac

ac

distribution on the interval [q_3 – c –d, H_1] with support missing on the interval [L_2, H_2].

The means of the four posteriors given in (C20) are $\dfrac{L_2^2-(q_1-c-d)^2+H_1^2-H_2^2}{2(L_2-(q_1-c-d)+H_1-H_2)}$,

$\dfrac{L_2^2-(q_3-c-d)^2+H_1^2-H_2^2}{2(L_2-(q_3-c-d)+H_1-H_2)}$, $\dfrac{H_1+H_2}{2}$ and $\dfrac{q_3-c-d+H_1}{2}$, respectively.

Given market beliefs, let P denote the payoff a firm receives when it does not advertise and does not obtain positive word-of-mouth. Thus, the expected return to not advertising is given by:

E(π| no advertise) =

$$\frac{1}{d}\Big[\int_{\eta+c}^{W}Pdq_3+\int_{W}^{q_1}\frac{L_2^2-(q_1-c-d)^2+H_1^2-H_2^2}{2(L_2-(q_1-c-d)+H_1-H_2)}dq_3$$
$$+\int_{q_1}^{L_2+c+d}\frac{L_2^2-(q_3-c-d)^2+H_1^2-H_2^2}{2(L_2-(q_3-c-d)+H_1-H_2)}dq_3+\int_{L_2+c+d}^{\eta+c+d}\frac{H_1+H_2}{2}dq_3\Big]$$

when $L_2 \le \eta \le H_2$ (C21)

E(π| no advertise) =

$$\frac{1}{d}\Big[\int_{\eta+c}^{W}Pdq_3+\int_{W}^{q_1}\frac{L_2^2-(q_1-c-d)^2+H_1^2-H_2^2}{2(L_2-(q_1-c-d)+H_1-H_2)}dq_3$$
$$+\int_{q_1}^{L_2+c+d}\frac{L_2^2-(q_3-c-d)^2+H_1^2-H_2^2}{2(L_2-(q_3-c-d)+H_1-H_2)}dq_3$$
$$+\int_{L_2+c+d}^{H_2+c+d}\frac{H_1+H_2}{2}dq_3+\int_{H_2+c+d}^{\eta+c+d}\frac{q_3-c-d+H_1}{2}dq_3\Big]$$

when $\eta > H_2$ (C22)

Subtracting (C21) from (C18) would yield V for a firm with ability, $L_2 \leq \eta \leq H_2$. Subtracting (C22) from (C19) would yield V for a firm with ability, $\eta > H_2$. Differentiating V with respect to η yields:

$$\partial V/\partial \eta = \begin{cases} \frac{1}{d}[P - \frac{H_1 + L_2}{2}] \text{ when } L_2 \leq \eta \leq H_2 \\ \frac{1}{d}[P - \frac{L_2 + H_2}{2}] \text{ when } \eta > H_1 \end{cases} \quad \text{(C23)}$$

Note that $\partial V/\partial \eta$ is constant in ability when $L_2 \leq \eta \leq H_2$. Since V is continuous in ability, for existence of a countersignaling equilibrium, V must equal zero when $\eta = L_2$ and when $\eta = H_2$. From (C23), the only way V can equal zero at both $\eta = L_2$ and $\eta = H_2$ is if $P = (H_1 + L_2)/2$. However, if $P = (H_1 + L_2)/2$, then $\partial V/\partial \eta > 0$ when $\eta > H_1$, since $H_1 > H_2$. If this is the case, then $V > 0$ when $\eta > H_1$. Thus, there does not exist a countersignaling equilibrium. ■

Proof of Lemma 7

After a fad occurs the posterior on a firm's ability is $\eta \sim U[q_1 - c - d, H_1]$. Market beliefs are specified such that all firms advertise. If a firm does not advertise, beliefs are such that consumers believe the firm is the lowest ability firm that could produce q_3. If a firm advertises, the following reputational states can occur:

$$(\eta | \text{ advertise}, q_3) \sim \begin{cases} U[q_1 - c - d, q_3 - c] \text{if } q_3 \in [q_1 - d, H_1 + c) \\ U[q_1 - c - d, H_1] \text{if } q_3 \in [H_1 + c, q_1] \\ U[q_3 - c - d, H_1] \text{if } q_3 \in (q_1, H_1 + c + d] \end{cases}$$

The expected payoff from advertising is given by:

$E(\pi | \text{ advertising}, q_3) =$

$$\frac{1}{d}\Big[\int_{\eta+c}^{H_1+c}\frac{q_1-2c-d+q_3}{2}dq_3+\int_{H_1+c}^{q_1}\frac{q_1-c-d+H_1}{2}dq_3 + \int_{q_1}^{H_1+c+d}\frac{q_3-c-d+H_1}{2}dq_3\Big]-z \tag{C24}$$

Given off-equilibrium beliefs, the following reputational states can occur if a firm does not advertise:

$$(\eta|\text{ no advertise}) \sim \begin{cases}\eta = q_1 - c - d \text{ if } q_3 < W \\ \eta = q_3 - c - d \text{ if } q_3 \geq W\end{cases}$$

The expected payoff from not advertising is given by:

$$\text{E}(\pi|\text{ no advertise}) = \frac{1}{d}\Big[\int_{\eta+c}^{W}(q_1-c-d)dq_3+\int_{W}^{\eta+c+d}(q_3-c-d)dq_3\Big] \tag{C25}$$

Carrying out the integration in (C24) and (C25) and subtracting (C25) from (C24) yields V:

V =

$$\frac{1}{d}\Big[\eta\Big(\frac{H_1+q_1-c-d-\eta}{2}\Big)+d\Big(\frac{H_1+q_1-c-d}{2}\Big)-q_1(W-c)+\frac{(q_1-c-d)^2-H_1^2}{4} + \frac{d^2+W^2-c^2}{2}\Big]-z \tag{C26}$$

In (C26), V obtains the same minimum value at the endpoints, $\eta = q_1 - c - d$ and $\eta = H_1$. The value of z given in the lemma is sufficiently small enough so that V ≥ 0 at these endpoints. ■

Proof of Proposition 2

After a fad occurs the posterior on a firm's ability is $\eta \sim U\ [q_1 - c - d, H_1]$. Market beliefs are specified such that no firms advertise. If a firm advertises, beliefs are simply updated using Bayes' Rule and the observation on q_3. The expected payoff from advertising is the same as in (C24).

If firms do not advertise the market updates its beliefs based on whether or not word-of-mouth is obtained. Again, for the given prior on η, let $g(q_3)$ represent the probability distribution of q_3. It follows that:

$$g(q_3) \sim \begin{cases} \dfrac{x-(q_1-d)}{d(H_1+c+d-q_1)} \text{ if } x \in [q_1-d, H_1+c) \\ \dfrac{1}{d} \text{ if } x \in [H_1+c, q_1] \\ \dfrac{1}{d} - \dfrac{x-q_1}{d(H_1+c+d-q_1)} \text{ if } x \in (q_1, H_1+c+d] \end{cases}$$

Let $f(\eta|$ no advertise, $q_3)$ represent the posterior on η if q_3 were always exactly observed when the firm refrained from advertising. It follows that:

$$f(\eta|\text{ no advertise, } q_3) \sim \begin{cases} U[q_1-c-d, q_3-c] \text{ if } q_3 \in [q_1-d, H_1+c) \\ U[q_1-c-d, H_1] \text{ if } q_3 \in [H_1+c, q_1] \\ U[q_3-c-d, H_1] \text{ if } q_3 \in (q_1, H_1+c+d] \end{cases}$$

Let M denote the payoff to a firm when it does not advertise and does not receive word-of-mouth. Let G denote the cumulative density function of $g(q_3)$. It follows that M is given by:

$$M = \frac{\int_{q_1-d}^{W} g(q_3)\left(\int (x) f(\eta \mid \text{no advertise}, q_3)dx\right)dq_3}{G(q_3 < W)}$$

Similar to Lemma 2, carrying out the integration above yields the value of M given in the proposition.

The expected payoff from not advertising is given by:

E(π| no advertise) =

$$\frac{1}{d}\left[\int_{\eta+c}^{W} M dq_3 + \int_{W}^{q_1} \frac{q_1 - c - d + H_1}{2} dq_3 + \int_{q_1}^{\eta+c+d} \frac{q_3 - c - d + H_1}{2} dq_3 \right. \tag{C27}$$

Subtracting (C27) from (C24) and integrating yields V:

$$\mathrm{V} = \frac{1}{d}[\eta(M - \frac{q_1 - c - d}{2} - \frac{\eta}{4}) + (W - c)(\frac{H_1 + q_1 - c - d}{2} - M) - \frac{H_1^2}{4}] - z \tag{C28}$$

In (C28), V obtains its maximum value when $\eta = 2M - (q_1 - c - d)$. The value of z given in the proposition is sufficiently large to ensure that this firm will not want to advertise. ■

REFERENCES

Akerlof, G.A., 1976, "The Economics of Caste and of the Rat Race and Other Woeful Tales," *Quarterly Journal of Economics*, 90, 599 - 617.

Banerjee, A.V., 1992, "A Simple Model of Herd Behavior," *Quarterly Journal of Economics*, 107, 797-817.

Becker, G.S. and K. Murphy, 1993, "A Simple Theory of Advertising as a Good or Bad," *Quarterly Journal of Economics* 108, 941-964.

Becker, G.S., 1981, *A Treatise on the Family*, Harvard University Press, Cambridge.

Bikhchandani, S., D. Hirshleifer and I. Welch, 1992, "A Theory of Fads, Fashions, Custom, and Cultural Change as Informational Cascades," *Journal of Political Economy* 100, 992-1026.

Boot, A., 1992, "Why Hang on to Losers? Divestitures and Takeovers," *Journal of Finance*, 47, 1401-1424.

Brown, S.J., W.N. Goetzmann and J. Park, 2001, "Careers and Survival: Competition and Risk in the Hedge Fund and CTA Industry," *Journal of Finance*, 56, 1869-1886.

Burdett, K., 1978, "Theory of Employee Search: Quit Rates," *American Economic Review* 68, 212-220.

Carhart, M., J. Carpenter, A. Lynch and D. Musto, 2002, "Mutual Fund Survivorship," *Review of Financial Studies*, 15, 1439-1463.

Caves, R.E. and D.P. Greene, 1996, "Brands' Quality Levels, Prices, and Advertising Outlays: Empirical Evidence on Signals and Information Costs," *International Journal of Industrial Organization* 14, 29-52.

Chen, H., N. Jegadeesh and R. Wermers, 2000, "The Value of Active Fund Management: An Examination of the Stockholdings and Trades of Fund Managers," *Journal of Financial and Quantitative Analysis*, 35, 343-368.

Chevalier, J. and G. Ellison, 1999a, "Career Concerns of Mutual Fund Managers," *Quarterly Journal of Economics*, 114, 389-482.

Chevalier, J. and G. Ellison, 1999b, "Are Some Mutual Fund Managers Better Than Others? Cross-Sectional Patterns in Behavior and Performance," *Journal of Finance*, 54, 875-899.

Cho, I. and D.M. Kreps, 1987, "Signaling Games and Stable Equilibria," *Quarterly Journal of Economics*, 102, 179-222.

DeGroot, M., 1970, *Optimal Statistical Decisions*, McGraw-Hill, New York.

Demougin, D. and A. Siow, 1994, "Careers in Ongoing Hierarchies," *American Economic Review*, 84, 1261-1277.

Fama, E.F., 1980, "Agency Problems and the Theory of the Firm," *Journal of Political Economy*, 88, 288-307.

Fee, C.E. and C.J. Hadlock, 2003, "Raids, Rewards, and Reputations in the Market for Managerial Talent," *Review of Financial Studies*, 16, 1315-1357.

Fee, C.E. and C.J. Hadlock, 2004, "Management Turnover Across the Corporate Hierarchy," *Journal of Accounting and Economics*, 37, 3-38.

Feltovich, N., R. Harbaugh and T. To, 2002, "Too Cool for School? Signalling and countersignalling," *RAND Journal of Economics* 33, 630-649.

Gibbon, R. and K.J. Murphy, 1992, "Optimal Incentive Contracts in the Presence Of Career Concerns: Theory and Evidence," *Journal of Political Economy*, 100, 468-505.

Grossman, G. and G. Maggi, 2000, "Diversity and Trade," *American Economic Review*, 90, 155-1275.

Gruber, M. J., 1996, "Another Puzzle: The Growth in Actively Managed Mutual Funds," *Journal of Finance*, 51, 783-810.

Harbaugh, R. and T. To, 2004, "False Modesty: When Disclosing Good News Looks Bad," Working paper.

Hendricks, D., J. Patel and R. Zeckhauser, 1993, "Hot Hands in Mutual Funds: Short-run Persistence of Performance," *Journal of Finance*, 48, 93-130.

Hertzendorf, M.N., 1993, "I'm Not a High-Quality Firm – But I Play One on TV," *RAND Journal of Economics* 24, 236-247.

Holmstrom, B., 1982, "Moral Hazard in Teams," *Bell Journal of Economics* 13, 324-340.

Holmstrom, B., 1999, "Managerial Incentive Problems: A Dynamic Perspective," *Review of Economic Studies* 66, 169-182.

Holmstrom, B. and J. Ricart i Costra, 1986, "Managerial Incentives and Capital Management," *Quarterly Journal of Economics* 101, 835-860.

Hong, H., J.D. Kubik and A. Solomon, 2000, "Security Analysts' Career Concerns and Herding of Earnings Forecasts," *RAND Journal of Economics*, 31, 101-120.

Horstmann, I. and G. MacDonald, 1994, "When is Advertising a Signal of Product Quality?" *Journal of Economics and Management Strategy* 3, 561-584.

Horstmann, I. and G. MacDonald, 2003, "Is Advertising a Signal of Product Quality? Evidence from the Compact Disc Player Market, 1983-1992," *International Journal of Industrial Organization* 21, 317-345.

Johnson, W., 1978, "A Theory of Job Shopping," *Quarterly Journal of Economics*, 92, 261-277.

Jovanovic, B., 1979, "Job Matching and the Theory of Turnover," *Journal of Political Economy* 87, 972-990.

Kihlstrom, R. and M. Riordan, 1984, "Advertising as a Signal," *Journal of Political Economy* 92, 427-450.

Lazear, E. P. and S. Rosen, 1981, "Rank Order Tournaments as Optimum Labor Contracts," *Journal of Political Economy* 89, 841-864.

Linnemer, L., 2002, "Price and Advertising Signals of Quality when Some Consumers are Informed," *International Journal of Industrial Organization* 20, 931-947.

Meyer, M. A., 1994, "The Dynamics of Learning with Team Production: Implications for Task Assignment," *Quarterly Journal of Economics* 109, 1157-1184.

Milbourn, T., R. Shockley, and A. Thakor, 2001, "Managerial Career Concerns and Investments in Information," *RAND Journal of Economics*, 32, 334-351.

Milgrom, P. and J. Roberts, 1986, "Price and Advertising Signals of Product Quality," *Journal of Political Economy* 94, 796-821.

Nalebuff, B. and J.E. Stiglitz, 1984, "Prizes and Incentives: Toward a General Theory of Compensation and Competition," *Bell Journal of Economics*, 15, 21-43.

Nelson, P., 1974, "Advertising as Information," *Journal of Political Economy* 82, 729-754.

O'Flaherty, B. and A. Siow, 1992, "On the Job Screening, Up or Out Rules, and Firm Growth," *Canadian Journal of Economics*, 25, 346-368.

Orzach, R., P.B. Overgaard and Y. Tauman, 2002, "Modest Advertising Signals Strength," *RAND Journal of Economics* 33, 340-358.

Prendergast, C. and L. Stole, 1996, "Impetuous Youngsters and Jaded Old-timers: Acquiring a Reputation for Learning," *Journal of Political Economy* 104, 1105-1134.

Riley, J. G., 2001, "Silver Signals: Twenty-Five Years of Screening and Signaling," *Journal of Economic Literature* 39, 432-478.

Rosen, S., 1982, "Authority, Control and the Distribution of Earnings," *Bell Journal of Economics*, 1982, 13, 311-323.

Rosen, S., 1986, "Prizes and Incentives in Elimination Tournaments," *American Economic Review* 76, 701-715.

Scharfstein, D. and J. Stein, 1990, "Herd Behavior and Investment," *American Economic Review* 80, 465-479.

Spence, M., 1973, "Job Market Signaling," *Quarterly Journal of Economics* 87, 355-374.

Thomas, L., S. Shane and K. Weigelt, 1998, "An Empirical Examination of Advertising as a Signal of Product Quality," *Journal of Economic Behavior & Organization* 37, 415-430.

Wermers, R., 2000, "Mutual Fund Performance: An Empirical Decomposition into Stock-Picking Talent, Style, Transaction Costs and Expenses," *Journal of Finance*, 55, 1655-1695.

Zwiebel, J., 1995, "Corporate Conservatism and Relative Compensation," *Journal of Political Economy*, 103, 1-25.

MICHIGAN STATE UNIVERSITY LIBRARIES
3 1293 02736 4094