Spreading Information and Media Coverage: Theory and Evidence from Drug Approvals

Asaf Manela*

JOB MARKET PAPER

January 15, 2011

Abstract

This paper puts forth and estimates a dynamic asset pricing model with asymmetric information to study the relation between the media and the stock market. In the model, the speed by which information spreads through a population of rational agents determines both the amount of media coverage and the equilibrium fraction of informed agents. Information is more valuable when the intertemporal growth in the precision of the uninformed about the random payoff is large. The model predicts that faster-spreading positive news induces higher pre-announcement returns, lower post-announcement returns, and higher announcement trade volume. I investigate the performance of the model using a panel of stock returns, volume, and media coverage around FDA drug approvals. Reduced-form tests reveal that, consistent with the model, drug approvals that receive more media exposure on the approval day and the next exhibit higher turnover on those days and lower returns the following week. Using structural estimation, I estimate that the amount of non-informational trading (noise) necessary to keep prices from fully revealing all information is small. Model-generated effects are quantitatively similar to those in the data. In equilibrium, fast-spreading news is purchased at a high rate, whereas slow-spreading news is not pursued at all.

*The author is at the Booth School of Business, University of Chicago, 5807 S Woodlawn Avenue, Chicago, IL 60637, USA. Email: amanela@chicagobooth.edu. I thank my committee chair, Pietro Veronesi, and committee members Doug Diamond, John Heaton and Lubos Pastor for their guidance. This paper has benefited from helpful comments and suggestions from Tal Barak-Manela, Zahi Ben-David, Andrea Frazzini, Matthew Gentzkow, Michael Gofman, Roni Kisin, Alan Moreira, Toby Moskowitz, Alexi Savov, David Solomon, Paul Tetlock, Luigi Zingales, and participants of the Economic Dynamics Working Group, the Finance Lunch Workshop in Chicago, the Econometric Society meeting in Rio de Janeiro, and the LBS Trans-Atlantic Doctoral Conference. Remaining errors are my own.

1 Introduction

The covariation of media coverage with asset prices has been documented by several recent empirical studies. This literature documents two pervasive features of the data. The first is that more extensive media coverage of news pertaining to a publicly traded firm is followed by less price drift in its stock price (Chan, 2003; Peress, 2008; Fang and Peress, 2009). Second, more media coverage is associated with more trading activity (Peress, 2008; Barber and Odean, 2008; Engelberg and Parsons, 2010). These facts are hard to reconcile with the canonical asset pricing model in which prices fully reveal all publicly available information. The empirical literature thus far has focused on comparative statics of the one-period Merton (1987) model as the leading explanation of this phenomena. This paper explores the possibility that these stylized facts can be generated by a dynamic private information model in which the amount of media coverage relates to the speed by which information spreads through a population of rational agents.

I model media coverage as increasing in the level of public interest in a story. A simple observation motivates this approach: more interesting news spread faster. When people meet and only some of them have information, the probability that information is transmitted, the *transmission rate of information*, is higher for more interesting information. Media coverage of a particular story also decreases in the availability of other newsworthy material. Consider the problem of a representative daily newspaper editor. The amount of space the editor can allocate to a given story is limited by many factors, such as the attention span of its readers, the cost of print and distribution, and journalists' time. The editor caters to his readers' interests by giving the most exposure to the most interesting stories of the day. But these are the same stories that are likely to spread faster by word-of-mouth, even in the absence of media coverage. This insight can explain the covariation between stock returns and coverage by the popular press, even in a financial market where only professional traders matter for asset prices. Empirical evidence surveyed below shows that money managers do in fact spread and exploit information over their social networks. Media coverage can therefore mirror public interest in a story without directly affecting it.

To understand the economic mechanism behind this intriguing empirical phenomena, I construct a fully specified asset pricing model with the aim of generating both comparative statics and quantitative restrictions on the data. The model is a four-period noisy rational expectations model with asymmetric information about the payoff of a risky asset. Information gradually spreads through a large population of risk-averse agents. The transmission rate of information parametrizes variation in the speed of information diffusion. I derive the optimal demand and equilibrium prices in closed-form expressions, which eases the analysis of the model and proves instrumental for structural estimation purposes by reducing the computational demand.

I derive the following reduced-form predictions of the model regarding price reaction to news. Consider the introduction of positive news into the market. The model predicts that before an official announcement, for example, by a press release, returns are increasing in the transmission rate of information, when at least some agents are informed early, and post-announcement returns are decreasing in the transmission rate. Interestingly, I show that variation in the transmission rate has an ambiguous effect on the official announcement-day return. This result implies that using announcement returns to measure empirically how positive news are, might be confounded by the strategic behavior of investors who possess foresight and knowledge of the process by which information spreads. Negative news has symmetric effects. That is, more interesting negative news exhibits lower pre-announcement returns and higher post-announcement returns.

Studying the value of information, I show that information is more valuable when the intertemporal growth in the precision of the uninformed about the random payoff is large. This feature results in a hump-shaped demand for information as a function of the transmission rate of information. The intuition for this result is as follows. Early informed agents choose to acquire their private information and take into account the future spread of information. Information that spreads faster promises informed agents a quicker gain from trading on information and one that is less subject to price swings. However, as this potential gain increases, informed agents trade more aggressively, which makes prices more informative and reduces the value of information. These two contrasting forces determine the value of information and the equilibrium informed fraction of the population.

In information market equilibrium, uninteresting news that propagates slowly is not pursued by anyone before the official announcement, because the fixed cost of information is prohibitively high. Faster-spreading information is purchased at a higher rate, while the fastest spreading news is somewhat less valuable. This unique feature of the model accentuates the covariation between transmission rates and the demand for the risky asset by influencing the extensive margin of information acquisition by the population as a whole. I find that this feature is essential for generating the large covariation between media coverage and stock returns and volume.

I investigate the performance of the model using a panel of stock returns, volume, and media coverage around new drug approvals by the FDA. These events provide a particularly clean laboratory for examining stock market reaction to news. Importantly, the selection of relevant articles is straightforward for the approval of drugs that have unique names. I measure the *media exposure* given to an approval story as the sum of all articles that report the approval on the official approval day and the next, weighted by the price of their adjacent advertising space. A first look at the data reveals that the patterns found by previous literature are borne out in my sample. Specifically, drug approvals that receive more media exposure on the approval day and the next exhibit higher turnover on those days, and then lower subsequent returns in the following week. These effects are statistically strong even after controlling for previously suggested proxies for information dissemination such as firm size, analyst coverage, analyst estimates dispersion, and other characteristics of the approval.

I then use the number of pages in the *Wall Street Journal* as a measure of the availability of other newsworthy material on approval day. Using this instrument for media exposure, I find that the newspaper editor's decision to grant more media exposure to an approval story causes a greater price increase at the approval. However, I do not find that this exogenous variation in media coverage leads to variation in post-approval price drift or in turnover. This evidence is consistent with my model in which both asset demand and media coverage respond to an exogenous transmission rate of information. What matters is not necessarily the media exposure of the story, but rather whether the story captures the imagination of investors and induces them to propagate it faster.

Using structural estimation, I show that the gradual information spread model quantitatively matches fairly well many empirical moments around drug approvals. Quantitative assessments of asymmetric information asset pricing models are for the most part missing in the empirical literature thus far.[1] I attempt to bridge this gap. I apply the Indirect Inference estimation method of Gourieroux et al. (1993) to the cross section of drug approvals and their short time series. I estimate that an uninformed agent must pay 50% of his position in the stock to become informed. I interpret this as the expected cost of trading on private information with a chance of adverse legal consequences. On average, only one in a million agents is informed about a drug approval before the official FDA approval. The signal-to-noise ratio for the average drug approval is about 10%, which means that the reduction in uncertainty regarding the value of the drug developing firm associated with an approval is small. This finding is consistent with DiMasi (2001) which studies the drug approval process and estimates that the probability that a drug will be approved conditional on surviving to the marketing application stage is about 75%. I estimate that non-informational supply shocks have a standard deviation of 1% of the total supply of the stock. Thus the model requires only a small amount of non-informational trading to prevent prices from fully revealing the news. Overall, the results from the structural estimation support the model. The model-generated effects are not substantially different from those in the data. The empirical success of the model in explaining the stylized facts about media coverage and stocks suggests that the process by which information spreads can be important for asset pricing.

One implication of my findings concerns the informational efficiency of capital markets. Variation in transmission rates of information means that a finer distinction can be made between the informational content of prices of different projects and firms. Specifically, in the context of drug approvals, I find that priority-review drugs exhibit a higher price increase upon approval and a lower post-approval return. The public is particularly interested in these drugs, which promise a significant benefit over existing treatment. The FDA is likely to prioritize a new drug for Cancer or HIV. Orphan drugs on the other hand, which treat only a small fraction of the population, exhibit higher price drift. The informational frictions seem less important

[1] A quantitative analysis of this type of models is Campbell et al. (1993), which considers time-varying risk aversion instead of time-varying information sets as I do. The authors report that implausibly large transitory shifts in risk aversion are required to match data on aggregate daily volume and serial correlation of returns. Changes in information sets I consider here can potentially resolve that puzzle.

for an economic activity society as a whole finds interesting. In equilibrium, the allocation of capital to such activities will be larger. Thus what may seem an arbitrary influence of fads and fashions on prices, can actually be an efficiency-enhancing mechanism for the allocation of capital.

A second policy implication for firms that aim to raise capital in financial markets concerns their expenditure on investor relations. In an economy with informational frictions, whether the result of limited attention or communication costs, a firm with a new project can benefit from increasing the transmission rate of information about its prospects to lower its cost of capital. Previous literature on the role of the media in finance has concluded that investor-relations firms can raise share prices by publicizing the news in the popular press (Fang and Peress, 2009) or giving news a positive spin (Solomon, 2009). But if the media is in fact just a mirror of public interest, the role of investor-relations agencies is more subtle. The goal of a successful publicity campaign is then to create an appealing narrative around news, one that captures the imagination of investors and makes positive information propagate faster through social networks. Viral investor relations, much like viral marketing, can be more successful than traditional attempts to increase firm visibility.

This paper contributes to several literatures. Recent work that studies the relationship between media coverage and stocks includes Fang and Peress (2009) which studies how cross-sectional variation in media coverage relates to variation in expected monthly returns. Chan (2003) documents differences in monthly returns for firms with and without headline news. Other studies try to better control for the content of news using earnings announcement surprises and focus on shorter return horizons.[2] These include Peress (2008) which documents lower post-earnings announcement drift for announcements covered in the *Wall Street Journal*, as well as, stronger trading volume upon announcement. Tetlock (2010) documents return predictability and trading volume patterns following news that are consistent with an asymmetric information model's predictions. News versus no-news correspond to comparative statics of his model about the variance of a private signal. Endogeneity of media coverage is a concern in all these studies. Engelberg and Parsons (2010) identify a causal effect of media coverage on the trading behavior of individual investors using local weather as an instrument. They find that investors exposed to media coverage of an earnings announcement in their local newspaper trade more. Using reporter connections and geographical links to newspapers, Solomon (2009) finds that IR firms are causally affecting both media coverage and returns. I provide a new model to study these results that features the transmission rate of information as the key variable of interest that drives both media coverage and stock behavior. I also suggest a finer-grained measure of media exposure using advertising rates, a new instrument for media coverage using the number of pages in the *Wall Street Journal*, and FDA drug approvals as a different and in many ways superior laboratory for studying the spread of information.

[2]A related but orthogonal strand of literature fixes media exposure and examines how variation in the content of news relates to stock behavior. Recent such studies include Tetlock (2007), Dougal et al. (2010), and Neuhierl et al. (2010).

The process by which information spreads in my model resembles the spread of an infectious disease. The notion that information spreads in this fashion has both theoretical and empirical footing. Information, unlike tangible goods, can be transmitted from one individual to the next without loss to the transmitting party. Since the cost of communication over a social network is rather low, any positive utility gained from passing along information can result in free and direct information transmission. Romer (1990) argues that ideas or technological improvements are inherently nonrival and only partially excludable, allowing their accumulation without bound and for informational spillovers. Dawkins (1989) considers the similarity between a piece of information and other replicating units, such as genes or living organisms. Successful biological traits replicate rapidly and spread through a naturally selected population. Similarly, more appealing ideas are retained and communicated at higher rates and prevail in social networks. More recently, Stein (2009) provides microfoundations for truthful exchanges of information between competitors. Empirical work by Shiller and Pound (1989) provides survey evidence that direct interpersonal communication is important in investment decisions, and that investor interest in specific stocks spreads like an epidemic. Hong et al. (2005) provide further evidence that mutual fund managers spread information directly, through word-of-mouth communication. Furthermore, Cohen et al. (2008) find that portfolio managers gain an informational advantage through education networks, and that their returns from this channel are concentrated around corporate news announcements. Gray (2010) finds that skilled investors share their profitable ideas with their competition.

My model builds on foundations laid by previous asset pricing theories with sequential information arrival. Hirshleifer et al. (1994) randomly assigns the informed population into early and late informed groups. In addition to liquidity traders, meant to keep prices from fully revealing all information, they introduce a risk-neutral competitive fringe of market-makers. This feature, which simplifies the solution, implies that prices follow a martingale and are semi-strong informationally efficient in this setup. In a similar setting, Holden and Subrahmanyam (2002) allow agents to purchase information in any of two periods of their model and investigate the serial correlation of stock returns and trade volume. I improve on their numerical work by characterizing the value of information in an intuitive closed-form expression. I further consider the problem of uninformed investors that randomly become informed in the future. Hong and Stein (1999) provides a behavioral model with gradual information diffusion across a population of differentially and symmetrically informed "newswatchers." Information in their model rotates between groups of investors, thus diffusing in a linear manner until the circle completes and all groups are fully informed. These newswatchers are boundedly rational in that they do not condition their trades on current or past prices, leading to underreaction to news. Furthermore, they interact with "momentum traders" who follow univariate strategies, resulting in overreaction at long horizons. The rational part of their model is closely related to my model. I extend their work with optimal learning from prices, information acquisition, and information asymmetry in the sense that different agents have different precision of information.

Information acquisition is also central in Veldkamp (2006). It studies media frenzies in emerging markets and uses the aggregate number of articles that reference an emerging market as a proxy for the cost of information in different environments. By contrast, the focus in my paper is on specific news announcements and the market reactions they induce. Hong et al. (2011) study a dynamic model of opinions, prices and volume based on word-of-mouth communication. They allow for intricate information diffusion dynamics at the cost of a simplified asset pricing framework in which agents agree to disagree and behave myopically as does much of the above theoretical work. I avoid this assumption and characterize the resulting hedging demand of forward-looking agents.

The paper proceeds as follows. Section 2 describes the media coverage and asset pricing model. Section 3 provides a first look at the data and establishes the main empirical facts to be explained. Section 4 then proceeds with structural estimation to determine whether the model can quantitatively match the data. Section 5 concludes.

2 Theory

The role of the media in financial markets as it pertains to the dissemination of information can be both active and passive. One can think of the media as providing an essential service to investors by informing them in a timely manner about valuation-relevant news. On the other hand, the popular press can perform a more passive role by picking up and exposing stories investors find interesting, but that would propagate through the relevant investor population even without any media coverage. It is the latter approach which I take here. I begin by describing a passive model of media coverage. I then construct a simple asset pricing model with the necessary features for the discussion and empirical analysis, and derive its main predictions. The link between the two is that both media coverage and asset demands respond to the level of public interest in the news, which determines the transmission rate of information.

2.1 A Passive Model of Media Coverage

Consider the problem of a representative daily newspaper editor. The editor has limited space he can allocate to any given story. This real estate is limited by the attention span of readers and by the cost of print and distribution. Readers often concentrate on the front page of the paper and skim the rest. The danger for the editor is that his reader will not feel up-to-date about current events when he encounters his peers and as a result cancel his subscription and move to a competing news source. Therefore, the editor gives more exposure to more interesting stories. But these are the same stories that are likely to spread faster through word-of-mouth, even in the absence of media coverage. Thus media coverage can depend on the level of public interest in a story without directly affecting it.

A news cycle at a modern newspaper like the *Wall Street Journal* begins when the production department allocates a certain number of columns for the editors to fill with news in the next day's edition. This space in the book plan between the ads is called the *newshole*. The amount of advertising sold and the availability of news determine its size. Newspapers have a target average advertising-to-news ratio they aim for on a weekly, or quarterly basis. Within this pre-specified newshole, the editors must decide what stories are newsworthy enough to be published. More interesting news features more prominently, say, on the front page, whereas other news is relegated to inner pages. At times of major news events such as election cycles or natural disasters, editors request additional space for news. When they do so, they compensate by reducing the amount of news on other days to maintain their target.[3]

Furthermore, since the editor cannot publish a marginal article, for example, just a word or two about the story, some stories will not make it into the news at all. With this additional constraint, media coverage is actually either positive if the story is interesting enough to pass the threshold, or is exactly zero. The editor must pick the day's most interesting stories and discard or postpone publication of the rest. Thus each story's chances of being published in a prominent section of the newspaper depends critically on the availability of other competing stories.

Consequently, I model the media coverage of a story as an increasing function of how interesting it is, which determines its transmission rate in the population, and a decreasing function of the availability of other newsworthy material. Let m^* denote the media coverage a story would receive if there was no fixed cost associated with publishing it:

$$m^* = \boldsymbol{x}^T \boldsymbol{\beta}_m + \gamma z + logit\left(\delta\right), \tag{1}$$

where δ is the transmission rate. It is defined as the probability of information transmission when asymmetrically informed agents happen to meet. $\boldsymbol{x}$ denotes a vector of story-specific characteristics and z measures the availability of other newsworthy material. With the fixed cost, the outcome of the editor's choice is a media coverage censored at zero:

$$m = \begin{cases} m^* & m^* > 0 \\ 0 & otherwise. \end{cases} \tag{2}$$

I try to measure this variable m in the empirical part below. The logit function provides a one-to-one mapping from $\delta \in (0,1)$ to the entire real line, as is common in probability models. This functional form assumption means media coverage is more sensitive to variation in transmission rates around the boundaries of its support.

2.2 Gradual Information Spread Asset Pricing Model

I extend the Grossman and Stiglitz (1980) framework to a three-trading period economy with a fraction of informed investors that evolves deterministically over time and is known by all

[3] Special thanks to Jim Pensiero and Robin Haynes of the *Wall Street Journal* for taking the time to explain this process.

market participants. The model has four periods labeled $t = 0, 1, 2, 3$, where the first three involve trade, and all uncertainty is revealed in period 3.

2.2.1 Preferences, Endowments and Tradable Assets

Agents can trade an elastically supplied risk-free asset with constant gross return R per period that is exogenously given. The second asset offers a risky payoff u in period 3 equal to:

$$u = \theta + \epsilon, \tag{3}$$

where $\theta \sim \mathcal{N}(\mu_0, \frac{1}{\tau_0})$ and $\epsilon \sim \mathcal{N}(0, \frac{1}{\eta})$ are independent shocks and the precisions τ_0 and η are positive. Agents enter period 0 with identical preferences and endowments. Agents of type i maximize their constant absolute risk aversion (CARA) utility over terminal wealth

$$V_0^i(W_0^i) = \max_{\{q_t^i\}_{t=0}^2} E\left[-e^{-\frac{1}{\phi}W_3^i}|\mathcal{F}_0^i\right], \tag{4}$$

by choosing the amount of shares q_t, while satisfying the law of motion of their stochastic wealth

$$W_{t+1}^i \quad = \quad RW_t^i + q_t^i Q_{t+1} \tag{5}$$

where $Q_{t+1} \equiv P_{t+1} - RP_t$ is the return to a zero-investment portfolio long one share of the risky asset. The initial endowments, W_0, are given and $\mathcal{F}_t^i$ is the information set of agent i described next.

2.2.2 Information Structure

At the initial period 0, all agents are identical and choose not to trade at the market-clearing price. A signal θ is revealed in period 1 to a fraction $I_1 \in [0, 1]$ of the population, which choose to pay a fixed cost c for this information. This fraction is determined in equilibrium. Information spreads deterministically between periods 1 and 2, before the payoff u is revealed in period 3:

$$I_2 - I_1 = \Gamma(I_1; \delta)(1 - I_1). \tag{6}$$

The incidence probability $\Gamma(I; \delta)$ is the probability that an uninformed agent at time 1 will become informed at time 2. It potentially depends on the fraction informed to allow for word-of-mouth transmission of information. $\frac{\partial \Gamma}{\partial \delta} > 0$ so that, all else equal, the uninformed are more likely to receive more interesting news, for which the transmission rate is higher. All agents know I_t and its law of motion.[4]

[4]That I_t is public knowledge is important to preserve the tractability of the model. Since I_1 is an equilibrium outcome, it is reasonable that all agents can infer I_2 from their knowledge of the process by which information spreads. An alternative specification could assume a stochastic version of (6). Such an extension could prove to be an important source of risk in a small market. However, as the size of the population grows large, the variance of the change in the fraction of informed tends to zero. Therefore, assuming a deterministic process is

Figure 1 illustrates the time line of information spread. Agents of type II observe the signal. They remain informed for the rest of time. Uninformed agents at time 1 become informed next period with probability $\Gamma(I_1; \delta)$. This important addition to the canonical model adds a new, third type of agent that will affect the dynamics of asset prices and the volume of trade. These agents of type UI are uninformed agents that become informed via direct information transmission. The rest, agents of type UU, remain uninformed until time 3. The three return periods of the model are meant to capture the periods before, on, and after an official release of news such as a drug approval as illustrated at the top.

All agents know the structure of the world and its parameters, and observe current prices, as well as their entire history. Thus an informed agent's information set at time t is $\mathcal{F}_t^I = \{P^t, I^t, \theta\}$, whereas that of an uninformed agent at time t is $\mathcal{F}_t^U = \{P^t, I^t\}$, where the superscript t denotes the entire history of the variable.

2.2.3 Noise

To prevent prices from fully revealing all information, I allow for noise in the supply of the risky asset. This noise traditionally has several interpretations, such as liquidity shocks (Wang, 1994), or allocational price changes (Grossman, 1995).

Let X denote the mean supply of the risky asset, which is augmented in every trading period by i.i.d noise $x_t \sim \mathcal{N}(0, \frac{1}{\xi})$ for $t = 1, 2$ with strictly positive precision ξ. Thus the risky asset market-clearing condition is

$$I_t q_t^I + (1 - I_t) q_t^U = X + x_t \qquad t = 0, 1, 2, \tag{7}$$

where $x_0 = 0$, q_t^I is the risky asset holdings of an informed agent leaving period t and q_t^U is that of an uninformed agent. In general, type UI demand would enter separately in (7), but since wealth does not enter the optimal demand functions with these preferences, agents of type UI will behave just like the initially informed of type II going forward.

2.2.4 A Noisy Linear Rational Expectations Equilibrium

Definition. A rational expectations equilibrium (REE) is an allocation $\{q_t^i\}_{t=0}^3$ for each agent type i and prices $\{P_t\}_{t=0}^3$ such that:

1. Taking prices as given, agents form correct Bayesian beliefs and maximize expected utility (4) subject to the budget constraint (5) every period.

2. Market clearing prices satisfy (7).

After the payoff is revealed, should trading open for a claim to the payoff, the price must be $P_3 = u$. At that time, agents are again identical in the absence of wealth effects and I assume

reasonable as long as the population exhibits sufficient mixing.

for volume calculation purposes that $q_3^i = X$ for all i. The assumed structure allows for a noisy rational expectations equilibrium with prices that are linear in the state as made formal by the following proposition:

Proposition 1 (Existence of Equilibrium)**.** *There exists a REE with prices that are linear in the state. Specifically, prices take the form:*

$$\begin{aligned} P_t &= a_t + b_t \Delta_t + c_t x_t \qquad t = 1, 2 \\ P_0 &= a_0, \end{aligned} \tag{8}$$

where $\Delta_t \equiv \theta - \mu_t$ is the prediction error of the uninformed and the coefficients a_t, b_t, and c_t are measurable with respect to the uninformed information set at time $t - 1$.

The coefficient a_t is the uninformed agent's expectation of the price next period. It is composed of the discounted present value of the expected payoff minus a risk premium. The coefficient b_t measures the price sensitivity to the uninformed agent's error at time t. It is non-negative and grows larger the more informed agents there are and the less uncertainty there is in the market about the return going into the next period. Finally, $c_t < 0$ is the price sensitivity to supply shocks. The ratio between b_t and c_t determines the informativeness of the equilibrium price.

The proof of Proposition 1 proceeds in several steps and constitutes the remainder of this section. First, I conjecture that (8) holds and derive agents' beliefs given this conjecture. I then solve for agents' optimal demand functions given their beliefs starting with the final period and proceeding by backward induction. In each period, I impose market clearing and verify the price functions satisfy the linear prices conjecture.

2.2.5 Learning from Equilibrium Prices

Uninformed agents are Bayesian and form rational expectations based on all of their available information, which in this case is restricted to the history of market clearing prices. The linearity of prices, together with the normality of the noise and the prior on the signal θ, imply a normal posterior distribution with a mean that is linear in prices and a precision that increases deterministically over time as summarized by the following:

Lemma 1 (Learning from Prices)**.** *Uninformed agent's beliefs evolve:*

$$E_{t+1}^U[\theta] \equiv \mu_{t+1} = \mu_t + k_{t+1}(P_{t+1} - a_{t+1}) \tag{9}$$

$$\Psi_{t+1}^U[\theta] \equiv \tau_{t+1} = \tau_t + \xi \beta_{t+1}^2 \tag{10}$$

for $t = 0, 1$ with (μ_0, τ_0) given and where $k_{t+1} \equiv \frac{\xi b_{t+1}}{\tau_t c_{t+1}^2}$ and $\beta_{t+1} \equiv \frac{b_{t+1}}{c_{t+1}}$.

For notational brevity, $E_t^i[\cdot]$ denotes the expectation operator conditional on agent i's information set at time t. Similarly, $\Psi_t^i[\cdot]$ denotes the conditional precision operator, defined as the inverse of the corresponding variance $Var_t^i[\cdot]$. Given their information of the signal, informed agents can perfectly infer the supply shock. Their expectation of the payoff u is just θ, which they observe, and its precision conditional on the signal, η, is constant. Note that nothing can be learned from period 0 prices so no updating occurs.

2.2.6 Optimal Asset Holdings and Prices

Closed-form expressions can be attained for all variables in the model as the appendix details. Period 2 of this model is identical to the classic Grossman and Stiglitz (1980) problem. Optimal holdings of the risky asset are:

$$q_2^i \quad = \quad \phi \Psi_2^i[Q_3] E_2^i[Q_3], \tag{11}$$

where the optimal asset holdings of type UI are the same as those of type II even though the changes in their holdings, their demand, in period 2 can be different. This time-varying heterogeneity is important for volume, but does not matter for prices in the absence of wealth effects. Agents trade off the expected excess return per share with the variance of the payoff that is scaled by ϕ, their risk tolerance. Substituting optimal demands (11) into the market-clearing condition at time 2 yields a linear price function $P_2 = P_2(\Delta_2, x_2; P_1)$. Matching coefficients yields the three P_2 coefficients that appear in the appendix. Importantly, the ratio of the news coefficient b_2 to the supply shocks coefficient c_2 has a simple expression:

$$\beta_2 \equiv \frac{b_2}{c_2} = -\phi I_2 \Psi_2^I[Q_3]. \tag{12}$$

The larger the magnitude of β_2 in absolute value, the more informative the price about the private information of the informed. This result shows that P_2 is more informative when agents are more tolerant of risk, which comes from holding the asset well into the future. The level of risk depends on the residual uncertainty about the payoff. In a multiple-asset model, this risk would come from the residual uncertainty about the systematic component of the portfolio. Importantly, when I_2 is high and more agents are informed at time 2 prices are more informative.

The choice of risky asset holdings at time 1 is somewhat more interesting yet harder to solve. Informed agents know they will remain informed in the future and therefore maximize:

$$V_1^I\left(W_1^I; \Delta_1, x_1\right) \quad = \quad \max_{q_1^I} E_1^I\left[V_2^I\left(W_2^I; \Delta_2, x_2\right)\right]. \tag{13}$$

The problem of uninformed agents at time 1 involves additional uncertainty about their informational type in period 2. Although the informed know they will remain informed next period, the uninformed at time 1 may get the signal for free at the beginning of time 2. The uninformed agent's problem at time 1 can be summarized by his value function:

$$V_1^U\left(W_1^U;\mu_1\right) = \max_{q_1^U} E_1^U\left[\Gamma\left(I_1;\delta\right)V_2^I\left(W_2^U;\Delta_2,x_2\right)+\left[1-\Gamma\left(I_1;\delta\right)\right]V_2^U\left(W_2^U;\mu_2\right)\right]. \quad (14)$$

With probability $\Gamma(I_1;\delta)$, the uninformed agent will receive the signal by period 2 and will then use it to make his informed portfolio decision. The gain he expects from this future informational advantage over less fortunate investors depends on the fraction informed both directly and indirectly. The direct effect is simply due to the increase in the probability of observing the signal. The indirect effect is from the reduction in uncertainty the signal provides by making prices more informative.

The exponential utility assumption allows us to express the expected informational advantage of the informed over the uninformed in period 2 as follows:

$$E_2^U\left[V_2^I\right] = \sqrt{\frac{\tau_2}{\eta+\tau_2}}V_2^U. \quad (15)$$

Due to the choice of a negative utility function, a larger proportionality term means a smaller informational advantage. Intuitively, when the informed have more trading periods to use their informational advantage, then its expected gain in utility is larger. The proportionality result in (15), together with the law of iterated expectations, allow us to rewrite the uninformed agent's problem in (14) as follows:

$$V_1^U\left(W_1^U;\mu_1\right) = \max_{q_1^U}\left[\Gamma\left(I_1;\delta\right)\sqrt{\frac{\tau_2}{\tau_{2+\eta}}}+\left(1-\Gamma\left(I_1;\delta\right)\right)\right]E_1^U\left[V_2^U\left(W_2^U;\mu_2\right)\right]. \quad (16)$$

Although this formulation greatly simplifies the problem, this special feature of the CARA-normal framework also eliminates $\Gamma(I_1;\delta)$ from the agent's first-order condition for q_1^U. Therefore, the change in I only enters the uninformed portfolio problem indirectly, through its equilibrium effects on means and variances.

An important feature of demand depends on the relationship between the period 2 excess return and the period 2 expected excess return in period 3. This relationship takes the following linear form:

$$E_2^i[Q_3] = E_1^i[Q_3]+\rho_1^i\left(Q_2-E_1^i[Q_2]\right), \quad (17)$$

where the coefficients ρ_1^i for both types of agents are negative. Therefore Q_2 is perfectly negatively correlated with $E_2^i[Q_3]$. The intuition is easy to see for the informed agent whose expected payoff at the final period does not change over time. From his perspective, any unexpected return in time 2 is due to the transitory supply shocks and he therefore expects it to reverse in the future. The same force operates on the uninformed agent's expectation, but his learning from prices attenuates this effect. The reason is that the uninformed agent cannot fully distinguish between the transitory supply shocks and the permanent news shocks that affect prices. Nonetheless, I show in the appendix that ρ_1^U is negative as well.

Faced with this opportunity set and full knowledge of the evolution of I over time, agents in period 1 choose their optimal holdings of the risky asset. Optimal holdings of the risky asset by an agent of type i going into period 2 are

$$q_1^i = \frac{\phi}{R}\left(\Omega_1^i \Psi_1^i[Q_2] E_1^i[Q_2] - \rho_1^i \Psi_2^i[Q_3] E_1^i[Q_3]\right), \tag{18}$$

where $\Omega_1^i \equiv 1 + \frac{\Psi_2^i[Q_3](\rho_1^i)^2}{\Psi_1^i[Q_2]}$. The first term in parentheses can be interpreted as a myopic demand component. Since Ω_1^i is larger than one, agents demand more of the asset than if no further uncertainty remained in period 2. Specifically, demand increases with the growth in precision, or equivalently, when agents expect a sharper decrease in uncertainty from period 1 to period 2. Agents also consider their continuation utility, which gives rise to the hedging demand component. Since ρ_1^i is negative, a higher expected return in period 3 induces a higher demand for the asset in period 1. Holding the asset to maturity provides a natural hedge since unexpectedly low returns today are expected to be offset in the future. Substituting (18) into the market-clearing condition at time 1 yields a linear price function $P_1 = P_1(\Delta_1, x_1)$, which is explicitly derived in the appendix.

Before examining the informativeness of prices, focusing on the expected return of an informed agent in excess of the risk premium demanded by the uninformed (their alpha) is useful. In the CARA-Normal framework, this conditional expected return can be expressed as a constant multiplying the news:

$$E_1^I[Q_2] - E_1^U[Q_2] = G_{12}\Delta_1,$$

where the sensitivity of their informational expected return to news is $G_{12} = b_2\frac{\tau_1}{\tau_2}$. The greater it is, the greater the return following positive fundamental value shocks and conversely for negative news.

The ratio of the news coefficient b_1 to the supply shocks coefficient c_1, which takes a more elaborate form than before, determines the informativeness of P_1. Specifically,

$$\beta_1 \equiv \frac{b_1}{c_1} = -\phi I_1\left(\Psi_2^I[Q_3] + \frac{1}{R}\Psi_1^I[Q_2]G_{12}\right). \tag{19}$$

The first term in parenthesis is just as before but with the current fraction informed I_1. As in β_2, the greater the risk tolerance or the greater the fraction informed at time 1, the more informative P_1 is about θ. Uncertainty about the payoff at time 3 again diminishes the informativeness of the price. Additionally, the second component, which depends on the intermediate period 2, increases the informativeness of P_1. Prices are more informative when the opportunity cost of trading on information, R, is small or when the informed are more certain about the intermediate return Q_2. Finally, the more sensitive the informed expected return $E_1^I[Q_2]$ is to the news, the more informative the price P_1.

In fact, as opposed to the first component of β_1, the second is recursively stated since both

$\Psi_1^I[Q_2]$ and G_{12} depend on both β_1 and β_2. In the appendix, I show that β_1 is the root of a cubic polynomial. As such, it admits a single real solution as long as its discriminant is negative. At least one real solution must exist by the intermediate value theorem, establishing the existence of equilibrium. The condition on the discriminant implies the following condition for uniqueness of the equilibrium in the linear class:

Proposition 2 (Uniqueness of Equilibrium)**.** *Suppose*

$$4\left(\gamma^3+\xi^2\zeta^4\omega\right)>\xi\zeta^2\left(\gamma^2+18\gamma\omega-27\omega^2\right),$$

where $\zeta=\phi I_1\left(\eta+\eta^2 I_2\xi\phi^2\right)$, $\gamma=\eta^2 I_2\xi\phi^2+\eta+\tau_0$ *and* $\omega=\eta^2 I_2^2\xi\phi^2+\eta+\tau_0$.
Then (19) *admits a single real solution and therefore the equilibrium is unique in the linear class.*

Having characterized the existence and uniqueness of equilibrium when the information structure is given, I next turn to analysis of the market for information and endogenize this structure.

2.2.7 The Value of Information

Agents at time 1 must choose whether to purchase the signal θ at the exogenously given cost c. The value of information v is the amount of wealth an uninformed agent is willing to pay to become informed, such that he is indifferent between purchasing the signal or not. Therefore, v is such that

$$E_1^U\left[V_1^U\left(W_1;\mu_1\right)\right] \quad = \quad E_1^U\left[V_1^I\left(W_1-v;\Delta_1,x_1\right)\right] \tag{20}$$

holds. Solving this equation for the value of information as a function of the informed fraction yields the following result:

Proposition 3 (Value of Information)**.** *The value of information v given the initial fraction informed I_1 is*

$$v\left(I_1;\delta\right) \quad = \quad \frac{\phi}{R^2}\left\{\frac{1}{2}\log\Omega_v+\frac{1}{2}\log\frac{\Omega_1^I}{\Omega_1^U}+\log\left[\Gamma\left(I_1;\delta\right)\sqrt{\frac{\tau_2}{\tau_{2+\eta}}}+\left(1-\Gamma\left(I_1;\delta\right)\right)\right]\right\}, \tag{21}$$

where $\Omega_v=\frac{\tau_2+\eta}{\tau_1}$ *and* $\Omega_1^i=1+\frac{\Psi_2^i[Q_3](\rho_1^i)^2}{\Psi_1^i[Q_2]}$ *for* $i\in\{U,I\}$.

The first component of the value of information is positive and increasing in the intertemporal growth rate of uninformed agents' precision about returns. The steeper this increase is, the more valuable the information. The intuition is that informational gains that are realized earlier are better than those that will only be realized in the future and are subject to additional shocks. However, as this potential gain increases, informed agents trade more aggressively, which makes P_1 more informative. Such an increase in τ_1 reduces the intertemporal precision growth and lowers the value of information.

The second component is positive as well and has to do with the intertemporal increase in precision of the informed relative to this increase by the uninformed. The third component is negative and represents the extent of information spillover to uninformed agents who do not pay for the signal. Intuitively, an increase in the probability that an uninformed agent becomes informed next period, $\Gamma(I_1;\delta)$, decreases the value of information to those who purchase the signal because of the spillover effect. Although this effect works in the same direction as strategic information substitutability, a range of I_1 close to zero exists such that strategic complementarity arises. That is, an increase in the cost of information will result in increased demand for information. The relative contribution of each of these components depends on the parameters of the model, but mainly on the precisions ξ, τ_0 and η. For example, the third component will be more negative when η is high relative to τ_2. This means informational spillovers are more important when the variance of the signal θ is large compared with the variance of ϵ. In Section 4, I estimate these parameters and the relative magnitudes of the three components.

In equilibrium, the fraction of informed agents, I_1^*, is such that agents are indifferent between purchasing the signal or not. Equivalently, the equilibrium value of information $v(I_1^*;\delta)$ equals the cost of information c.

2.2.8 Ex-ante Prices

Ex-ante, at time 0, all agents are identical and symmetrically uninformed. Furthermore, there is no noisy supply at time 0. The Pareto optimal allocation is one without trade, in which prices are such that agents are happy holding on to their endowments. Assuming a competitive equilibrium in the market for information, an agent at time 0 is indifferent between being informed or uninformed in period 1. Therefore, he maximizes the expected value of $V_1^U(W_1;\mu_1)$, which depends on future expected returns. As before, it can be shown that the following linear relationship holds between Q_1 and time 1 expectations of future returns:

$$\begin{aligned} E_1^U[Q_2] &= E_0[Q_2] + \rho_0\left(Q_1 - E_0[Q_1]\right) \\ E_1^U[Q_3] &= E_0[Q_3], \end{aligned} \tag{22}$$

where the loading on the innovation ρ_0 is negative. Thus the period 1 return is perfectly negatively correlated with the uninformed agent's expected return from time 1 to time 2. Furthermore, the expected return of the uninformed at time 1 about period 3 returns is independent of the period 1 return. In both periods 0 and 1, the uninformed expect the same final period return.

Ex-ante optimal holdings of the risky asset are

$$q_0 = \frac{\phi}{R^2}\left\{\Omega_0\Psi_0[Q_1]E_0[Q_1] - \rho_0\Omega_1^U\Psi_1^U[Q_2]E_0[Q_2] + \rho_0\rho_1^U\Psi_2^U[Q_3]E_0[Q_3]\right\}, \tag{23}$$

where $\Omega_0 = 1 + \frac{\rho_0^2\Omega_1^U\Psi_1^U[Q_2]}{\Psi_0[Q_1]}$. The myopic demand appears first, followed by hedging demand induced by future expected returns. Since the coefficients ρ_0 and ρ_1^U are both negative, all hedging demand components are positive. The expectations only depend on P_0, and $q_0(P_0)$ is linear in P_0. As a result, the equilibrium price P_0 is just a constant given the parameters of the model. The ex-ante price P_0 turns out to not depend on I_1 or the process assumed for information spread. Since the intermediate process by which information spreads changes nothing about the expected payoff from holding the asset to maturity at time 3, the equilibrium price is the same as the no-information economy price:

$$P_0 = \frac{\mu_0}{R^3} - \frac{X}{R^3\phi\Psi_0[u]}. \tag{24}$$

This result implies that if we were to examine low-frequency returns and the spread of information were to complete within a period of observation, then variation in transmission rates would not be observed.

This result concludes the proof of Proposition 1 establishing the existence of a noisy linear REE in this model. With the asymmetric information asset pricing framework in place, we can examine how information spread dynamics affect how market prices respond to news.

2.3 Price and Volume Reaction to News

The events I study below pertain to positive news. I next investigate how returns and trade volume are expected to respond to positive news ($\Delta_0 > 0$), while noting that symmetric effects arise for negative news. Because I study idiosyncratic news that pertains to individual firms, I concentrate on expected returns in excess of the unconditional risk premium. Excess expected returns conditional on news Δ_0 can be expressed as a constant multiplying the news:

$$E_0[Q_t|\Delta_0] - E_0[Q_t] = G_t\Delta_0,$$

where G_t is the sensitivity of expected returns to cash flow news. When news is introduced to the market, some of it will manifest immediately and the rest gradually until it is fully incorporated into prices or, in the case of the model, no later than period 3. In fact, it can be shown that the following relationship must hold between the sensitivities:

$$R^2G_1 + RG_2 + G_3 = 1. \tag{25}$$

If we consider a short amount of time so that $R \approx 1$ then the Gs must sum up to one. Therefore any change in the process by which information spreads that increases one period's return must be offset by a change at a different period. The following proposition establishes the direction of a marginal change in δ:

Proposition 4 (Price Reaction to News)**.** *Holding the fraction informed at time 1 and the precision τ_1 constant and positive, a marginal increase in the transmission rate δ:*

1. *Increases the period 1 expected return conditional on positive news.*
2. *Has an ambiguous effect on period 2 returns.*
3. *Decreases the period 3 expected return conditional on positive news;*

i.e., $\frac{\partial G_1}{\partial \delta} > 0$, $\frac{\partial G_2}{\partial \delta} \lesseqgtr 0$ *and* $\frac{\partial G_3}{\partial \delta} < 0$.

The intuition for this result is straightforward given (25). The channel through which an increase in δ operates is through an increase in I_2. The first-period return depends on P_1, whose sensitivity to payoff news shown in (19) is increasing in the fraction informed at time 2. The intuition is that informed agents trade more aggressively when they expect the information gap between informed and uninformed agents to close faster. When they do so, prices convey more information about the signal θ and less about the supply shocks. Period 3 expected returns conditional on positive news depend on the expected payoff and decrease with P_2. More informed agents at time 2 implies a more informative P_2, as can be clearly seen in (12). This brings prices closer to their frictionless benchmark and results in a lower expected Q_3. The intermediate-period expected return can increase or decrease because it is wedged between the two. An increase in P_1 informativeness lowers G_2, whereas an increase in P_2 informativeness increases it. Figure 2 plots excess returns for each of the three periods and confirms that while the dependency of G_1 and G_3 on transmission rates is monotone, the relationship for G_2 is not.

Let T_t denote turnover in period t defined as the number of shares traded over shares outstanding. The relationship between turnover and the transmission rate is neither linear nor monotone, as can be seen in Figure 2. It is the case however that announcement turnover is first increasing in the transmission rate and then remains around the same level. We can gain some intuition for this relationship by examining the expression for turnover in the intermediate period when $I_1 \downarrow 0$:

$$T_2 = \frac{1}{2}\left\{\Gamma\left(I_1,\delta\right)\left|q_2^I - X\right| + \left(1-\Gamma\left(I_1,\delta\right)\right)\left|q_2^U - X\right| + \left|x_2\right|\right\}. \tag{26}$$

The first term in absolute value is the demand of uninformed agents that become informed late in period 2. For positive news, we would expect it to be positive. The second term in absolute value is demand by agents who remain uninformed, we would expect it to be negative. The last term is the supply shocks which contribute $\sqrt{\frac{\pi}{2\xi}}$ to the level of turnover (ignoring covariances). When the incidence probability $\Gamma\left(I_1,\delta\right)$ is small, demand by the small number of informed agents dominates since they are able to hide their information. As a result, an increase in transmission rates results in higher turnover. However, for large $\Gamma\left(I_1,\delta\right)$, prices are highly revealing so both demands are similar in magnitude. Therefore, we would expect period 2 turnover to be higher for fast-spreading ($\delta = 1$) news than for uninteresting or slow-spreading

($\delta = 0$) news. The intuition is that a higher transmission rate means that a larger part of the previously uninformed population becomes informed in period 2 and increases its holdings of the risky asset. This intuition can break down when so many agents become informed that prices are close to fully revealing, in which case turnover is insensitive to δ.

The predictions of Proposition 4 and the numerical analysis giving rise to the turnover prediction are about changes to the transmission rate holding all else equal. However, when agents acquire information strategically, they take into account how quickly this information will spread. Therefore, the reduced-form analysis in the next section tests these predictions in a ceteris-paribus experiment. By contrast, the structural estimation exercise in Section 4 fully accounts for strategic information choice and its influence on observable outcomes such as returns and trade volume.

3 A First Look at the Data

The model described above yields predictions about observables such as returns and trade volume, and then ties them to media coverage and to the process by which information spreads. What remains is finding a suitable empirical environment to bring these predictions to data. I study a cross section of new drug approvals by the U.S. Food and Drug Administration. I examine a short time series of prices, volume, and media coverage around each approval. Times 0, 1, and 2 in the model correspond to the beginning of the pre-approval, approval, and post-approval event windows, respectively. Thus the informed fraction at time 1 trades on information that has not yet been publicly disclosed. The informed at time 2 include investors that receive the news directly from centralized news outlets or alternatively by word-of-mouth from the informed at time 1. In any case, public interest in the drug approval story determines both the transmission rate of information and the amount of media coverage the story receives upon the approval. I therefore use media coverage as a proxy for the transmission rate of information.

Drug approvals provide a particularly clean laboratory for examining stock market reaction to news that varies in its prominence for several reasons. First, information on approved drugs is readily available from the FDA. Second, the event's timing is exogenous to the firm developing the drug (its sponsor). This exogeneity is important since firms could otherwise time this event to maximize their share value, for example, by releasing the information on a certain day of the week (DellaVigna and Pollet, 2009). Third, many public pharmaceutical companies apply for drug marketing approvals, which allows for large sample studies. Fourth, a marketing approval is always a positive shock to the sponsor's future cash flow since it basically provides the sponsor with a real option on the drug's production. Thus the direction of the effect is predictable, and, to some extent, the impact on future cash flow can be estimated ex-ante. Finally and importantly, the unique drug names and active ingredients allow for a free-text article search that is likely to produce only articles that discuss the approval story. This is not the case for other well-studied events, such as earnings announcements, whose covering articles are harder to classify. In addition, earnings announcements often include soft information pertaining to future profitability that can be confused with earnings surprise.

3.1 Measuring Media Exposure

Previous literature on the effects of media coverage has focused on the extensive margin alone and compares news covered by the popular press with news not covered at all.[5] I suggest a new measure of *media exposure* that takes into account both the extensive and the intensive margin of media coverage using the price of advertising adjacent to a newspaper article.

3.1.1 Advertising Rates

To measure media coverage, I need a way to compare the prominence of an article published on the front page of the *Wall Street Journal* to an article placed on an inner page of a small-town newspaper. I approximate the relative emphasis a particular news item receives from the media, its media exposure, by weighting each news item by the price of its adjacent advertising space. This method allows me to construct a uniform dollar-value measure of media coverage and aggregate it over a period of time and across different types of publications. Although this paper focuses on daily newspapers and magazines, once could construct a similar measure for other mediums, such as television and radio.

I assume the market for advertising space is competitive and that an advertiser maximizes the media coverage of the good it is promoting. Thus, the price of ad space should reflect its marginal value to the advertiser. Determinants of this value include the ad's prominence, the medium's readership, and, to some extent, its editorial reputation. The size of the ad is obviously important as it serves to capture the reader's attention. Advertisers pay a premium for color ads as well as for special-position ads placed on the first few pages of a newspaper section. Circulation is a major determinant of ad rates, and the purchasing power of the paper's audience is important as well. Ferguson (1983) shows that daily newspaper ad rates are increasing not only in circulation but also in the local income per city household. Advertising space in higher-longevity publications such as magazines is more expensive than in daily newspapers, which have a high turnover. Finally, the publication's reputation for editorial scrutiny can play a role. For example, an advertiser must pay a premium for the scrutiny of the *New York Times*, which is important for the credibility of its content. Therefore, my measure of media exposure also captures the publication's credibility. In terms of media exposure, the space an ad occupies is a resource that is practically identical to the space a news item occupies. Therefore, I use advertising rates to quantify the news item's media exposure.

[5] See, for example, Chan (2003); Eisensee and Stromberg (2007); Peress (2008); Fang and Peress (2009)

3.1.2 Methodology

To determine the price of ad space, I use the open (non-contract) display advertising rate per column-inch quoted for an ad on the same page as the article. I focus on print publications that have a fairly standardized market for advertising. Rate cards are published yearly by each publication and collected by several agencies. My sample includes a 1998 world wide cross section of 726 newspapers, weeklies, and monthly magazines covered by the news database. All publications have an open column-inch rate and circulation. Daily newspapers usually have a different rate for Sunday, in which case they also specify Sunday circulation. Magazines quote a rate for a full black-and-white page, which I convert to a column-inch rate using the magazine's layout specification. Circulation and rates are hand collected from Editor and Publisher (1998), Gale Research (1998), Oxbridge Communications (1998), Hollis Directories (1998) and Stamm (1998). U.S. daily newspapers provide most of the media coverage of drug approvals in my sample. Figure 3 plots a regression of their advertising rates on circulation. We can see that circulation can explain much of the variation ($R^2 = 80\%$). The rest of the variation is likely due to the purchasing power of the readers and to some extent the prestige of the publication.[6]

For each article of interest I can match to an ad rate, I calculate a media exposure grade equal to the regular weekday price per column-inch. If the article was published on Sunday and the publication has a special Sunday rate then I use that rate instead. Sunday rates are 20% higher on average and circulation is 38% higher on average than on weekdays. Newspapers also charge a special premium for guaranteed positions. If the article was featured on the front page of the paper then I multiply its grade by 5. Pages 2 and 3 get a 30% premium and pages 4 and 5 a 20% premium. Front-page advertising is a relatively recent phenomena many journalists consider taboo (Shaw, 2007). Thus, although the premia I assign for pages 2-5 are based on a small sample survey of newspapers' actual premiums, front-page advertising rates are practically impossible to get and the premium is based on media experts' estimates and a few small newspapers that quote such a rate. In unreported tests, I multiply each article's grade by its word count to proxy for the size of the article, which can be important for grabbing the attention of readers. This modification adds no further explanatory power.

I am interested in an effect on markets that are usually closed when newspapers are printed late at night. Therefore, I match daily price and volume data of traded securities from CRSP to an aggregated measure of daily media exposure by summing the grades of all articles published between the previous day's market close and the return day market close. This way, each closing price and daily volume is matched with the new media exposure they should reflect. The news database contains duplicate articles, mainly when it subscribes to an agency that provides it with full-content articles as well as a second agency that provides abstracted articles listed under a different source code. Therefore, I omit duplicate articles from the same source if they are published on the same day and their headlines' first three words are the same as those of a

[6] For a detailed discussion of newspaper advertising rate structure and terminology, see Ferguson (1963) and Ferguson (1983).

previously aggregated article.[7]

3.2 New Drug Approvals Sample

I obtain from the Drugs@FDA database all Original New Drug Approvals from January 1990 to June 2007. The FDA does not disclose information about applications it does not approve or that were withdrawn. Many marketing approvals refer to the same drug or active ingredient but for different dosage forms. Since my identification strategy relies on the accuracy of the article search results, I only keep drugs based on an active ingredient that has never before been marketed in the United States in any form. The FDA marks these as Chemical Type 1 or New Molecular Entity applications. In addition, if the same drug is administered in more than one form, I keep only the first approval.

I match each drug approval with its original sponsor's daily share information from CRSP. Of the 320 approvals I can match to a publicly traded security, 65 are American Depository Receipts (ADRs) for which the CRSP record of outstanding shares has a different meaning. I attain market value for ADRs used in the firm size controls from Datastream. Since the FDA's working calendar coincides with that of the U.S. financial market, event day zero is also a trading day, although the FDA can issue the approval letter after market close. The FDA's policy as described in U.S. Food and Drug Administration (1998) is to convey this information to the applicant within one business day, at which point at least some market participants know with certainty the drug is approved. Even though the sponsoring firm is not obliged to make this news public, the vast majority issue press releases so that newswires and the popular press report the story within a day. In any case, FDA policy is to make the approval letter publicly available on its web-site and through a fax-on-demand system as soon as possible and no more than three working days past approval. Nonetheless, the choice of event window involves a tradeoff between, on the one hand, clear identification of articles that discuss only news of the approval as opposed to media coverage of the stock market's reaction and, on the other hand, capturing all of the media exposure of the drug approval. As Figure 4 shows, although many news articles are written on the second day, newswires begin to report on day zero and the largest price change is on day 1. Therefore, I calculate CAR for days 0 to 1 to capture the immediate market reaction and days 2 to 6 to capture the delayed response. I use the pre-approval window on days -5 to -1 to allow for abnormal market reaction prior to the official approval that better-informed traders possibly induce.

I systemize the collection of event-relevant articles with a predetermined template for a text search specification using the drug's name, active ingredient, and approval date. For example, the drug Lamisil, based on the active ingredient Terbinafine Hydrochloride and approved on

[7] When I run the same tests but use only daily newspaper data (no weeklies or monthly magazines), the regression coefficients on media exposure become slightly more statistically significant. This can be because weeklies and monthly magazines relate to prices in different ways, because my advertising rates data on them are flawed or due to chance. Future work can probably use daily newspaper rates alone to measure media exposure, making the data-collection process less demanding.

March 9, 1999, has the following search specification:

Free text	((LAMISIL) or (TERBINAFINE HYDROCHLORIDE)) and ((”food and drug administration”) or (FDA) or (F.D.A))
Date range	03/09/1999 to 06/07/1999
Search in	Full text articles

This procedure yields 46,338 unique articles about the approvals beginning one year before each approval and ending three months later. Of these, I can match 14,203 to advertising rates. I calculate *Media Exposure* for each approval event as the sum of all news articles that report the approval on the official approval day and the next, weighted by the price of their adjacent advertising space. *Preceding Media Exposure* and *Subsequent Media Exposure* are similarly calculated over the pre-approval and post-approval time frames.

Table 1 provides summary statistics for the 320 drug approvals. Sixty-five percent of the approvals in the sample received no immediate measurable media exposure. The average media exposure is $1,500, which is approximately equivalent to three regular *Washington Post* articles. The print media further covered most approvals over the subsequent business week. Firm size can play an important role in explaining abnormal returns since the impact of a new drug approval on a firm's earnings is closely related to its pre-approval market value. For a giant like Pfizer, an additional drug probably makes less of a difference than it would to a drug-developing start-up whose entire value stems from the prospects of a single drug. The sampled firms exhibit a wide variation in size measured as market capitalization one year before the approval in 1990 dollars. With an average firm value of $28 billion, the findings of this study cannot be dismissed as a small or illiquid stocks phenomena. Analyst coverage is also a potential source of information dissemination to a wide audience. On average, 19 analysts cover the sampled securities.

A variety of drugs were approved in the sampled period. 45% received priority-drug-review classification to speed up the process. Orphan drugs for rare diseases constitute 19% of the sample. Cancer and HIV/AIDS drugs comprise 16% and 6% respectively, of the drugs for which I have indication data. The average approval granted its sponsor with 16 years remaining of intellectual property rights from patents and 5 years of exclusivity rights by the FDA, resulting in a real option for non-trivial monopoly rents.

Using the sample averages in Table 2, we can discern several interesting features of drug approvals. The average drug approval generated a 1.33% abnormal return in the pre-approval period of trading days -5 to -1. Upon approval, it returned an additional 1.35% and then declined 0.53% over the subsequent five trading days. Since the standard errors of the pre-approval and approval means are small, we can reject that they are zero at usual significance levels. About half of the price appreciation occurs before the drug is officially approved which is indicative of superiorly informed agents participating in the market. The post-approval abnormal return of the average drug is statistically no different from zero. When we do not condition on any information other than the approval itself, the market's reaction is consistent with semi-strong market efficiency.

However, if we condition on a certain level of media exposure, the picture is different. In Figure 5, I split the sample into high, low, and zero media exposure sub-samples. Recall that media exposure is measured on the approval day and the next. The top panel plots the average cumulative abnormal return for each sub-sample. All three sub-samples feature a price increase in the days before the approval. The pre-approval return seems higher for drugs that would later appear in the news. This suggests insider-trading activity is increasing in *future* media exposure. At approval time, drugs covered by the media exhibit a higher price increase than the rest. Post-approval, the stock price of drug sponsors that received no initial media exposure continue to appreciate while low-media-exposure firms maintain their valuation. Interestingly, high-media-exposure approvals exhibit a negative drift following the approval, which continues even at a longer horizon than the one I test below. The bottom panel plots abnormal turnover for each of the same sub-samples. Turnover on the approval day is higher for better exposed approvals. In fact, high-media-exposure approvals exhibit higher turnover throughout the examined 90-day period surrounding the approval.

Although these results suggest variation in media exposure is related to the path of price adjustment to news, the small number of observations in each sub-sample and the lack of obvious controls, such as firm size and calendar effects, yield large confidence intervals around the means that make the three lines statistically indistinguishable. Also note that the no-media-exposure sub-sample is larger than the two others, which considerably reduces the volatility of its mean estimates. I next turn to linear regressions that allow me to control for observed heterogeneity along these dimensions.

3.3 A Linear Model of Media Coverage

Before moving on to a fully specified model, first considering a simple linear regression specification of media coverage and stock returns may be instructive. Denote m_j as the media exposure of drug approval j over days 0 and 1. Denote R_{jt}, $t = 1, 2, 3$ as the Pre-Approval, Approval, and Post-Approval cumulative abnormal returns of j. We can think of the joint behavior of media coverage and returns in a reduced form of the model as follow:

$$\begin{aligned} m_j &= \boldsymbol{x}_j^{\mathbf{T}} \boldsymbol{\beta}_m + \gamma z_j + \delta_j + \epsilon_{mj} \\ R_{tj} &= \boldsymbol{x}_j^{\mathbf{T}} \boldsymbol{\beta}_t + \pi_t \delta_j + \epsilon_{tj} \qquad t = 1, 2, 3, \end{aligned} \tag{27}$$

where δ_j is the transmission rate of information specific to this approval, $\boldsymbol{x}_j$ is a vector of controls, and z_j measures the availability of other newsworthy material.

The availability of other newsworthy material should crowd out the exposure an approval gets. To capture the editor's outside option, I use WSJ News Pressure defined as the 40-day moving average of the number of pages in section A of the *Wall Street Journal* published on day

1 when drug-approval newspaper coverage usually appears.[8] A plot of WSJ News Pressure over the sample period appears in Figure 6. We can discern both a cyclical time trend and a seasonal component. The time trend seems to peak around the end of the 20th century, when the print newspaper industry begins its decline in favor of online alternatives. The seasonal component peaks around May and November and bottoms out in February and August. I include month and year fixed effects in all regressions to control for these effects that are likely due to changes in the demand for advertising.

Since transmission rates are not directly observable, they would appear in the error term of the regressions in (27). Therefore, I employ a two-stage estimation procedure by first estimating a latent $\hat{\delta}_j \equiv \delta_j + \epsilon_{mj} = m_j - \hat{m}_j$ from the first regression. I test the sign of the coefficients π_t from

$$\begin{aligned} m_j &= \boldsymbol{x}_j^{\mathbf{T}} \boldsymbol{\beta}_m + \gamma z_j + \hat{\delta}_j \\ R_{tj} &= \boldsymbol{x}_j^{\mathbf{T}} \boldsymbol{\beta}_t + \pi_t \hat{\delta}_j + \epsilon_{tj} \qquad t = 1, 2, 3 \end{aligned} \tag{28}$$

under the assumption that $E[\epsilon_{mj}\epsilon_{tj}] = 0$. To account for estimation error in $\hat{\delta}_j$, I estimate this four-equation system simultaneously using GMM and cluster standard errors by the approval day.

The first prediction of the model from Proposition 4 is that pre-approval returns are increasing in δ, so we would expect $\pi_1 > 0$. The second prediction is that post-approval returns are decreasing in δ, so we would expect $\pi_3 < 0$. Furthermore, time 2 returns are non-monotonic in δ and therefore, without better knowledge of the parameters of the model, the model does not predict a sign for π_2.

The results in Table 3 support the predictions of the model. Column (1) reports the first-stage regression estimates. We can see that media exposure of drug approvals is lower when WSJ News Pressure is high. The availability of other newsworthy material in the business press seems to crowd out media coverage of drug approvals. Column (2) shows that pre-approval returns are increasing with transmission rates, though the effect is not statistically significant. One potential explanation for the large confidence interval around the effect is perhaps that the pre-approval window I use, days -5 to -1, is misspecified. Although the official announcement is relatively straight-forward to map to the model, knowing when private information about the approval begins to spread in the market is difficult.[9] Column (3) shows that approval

[8] This measure of news pressure is related to the television news pressure instrument constructed in Eisensee and Stromberg (2007) that proxies for the availability of newsworthy material using a 40-day moving average of daily TV news pressure defined as the median (across broadcasts in a day) number of minutes a news broadcast devotes to the top three news segments in a day. In addition, the authors use Olympic games incidence as a second source of variation. In unreported results, I find that TV news pressure does not crowd out drug approval news, perhaps because major TV networks cater to a different clientele than that of the business press I try to capture with my instrument. Olympic games do crowd out drug-approval news, but only coincide with 8 drug approvals in my sample.

[9] See Ellison and Mullin (2007) for a discussion of this issue and an estimation procedure for the day information begins to leak.

returns are uncorrelated with transmission rates. Column (4) shows that post-approval returns are decreasing in δ. This effect is statistically and economically significant. A one standard deviation increase in $\hat{\delta}$ implies a 0.6 percentage points higher pre-approval return and a 0.9% lower post-approval return. For the average drug developing firm in this sample, this effect means a $161 million higher pre-approval market value and a $265 million lower value post-approval.

3.3.1 Does the Media Affect Asset Prices?

Huberman and Regev (2001) describe in detail a thrilling case study in which extreme media exposure that contained no news in the traditional sense coincided with a spike in a stock price. The authors conclude with a provocative hypothetical question, asking what the stock price of the studied firm would have been if the editor of the *New York Times* had chosen to kill the story. In an attempt to answer this question, the last three specifications in Table 3 entertain the possibility that the media are the ultimate drivers of stock returns around approvals. WSJ News Pressure is plausibly exogenous to variation in media coverage of drug approvals, which are usually not front-page news and appear in the business section if at all. I employ a two-stage least squares instrumental variable regression to identify such an effect.

The first stage regression is the same as the one reported in column (1). The statistically significant negative coefficient on the instrument supports its validity. The F-test statistic for γ is 6.5 with p-value 0.01. The second-stage IV regressions in columns (5) to (7) regress returns on the predicted value $\hat{m}$ and the controls. Neither the pre-approval nor the post-approval returns seem to be sensitive to this exogenous variation in media coverage. Their loadings on $\hat{m}$ are statistically no different from zero and even go the opposite way. On the other hand, the estimates in column (6) show the approval return is higher for drugs with more media exposure. The coefficient on $\hat{m}$ is statistically different from zero at the 10% significance level. This evidence suggests that an exogenous increase in the media exposure of a drug approval results in a higher stock price of the developing firm. However, this price increase does not seem to revert in the business week following the approval as would be predicted by a model of overreaction to prominent news (e.g., Hong and Stein, 1999; DeMarzo et al., 2003).

3.3.2 Identification Concerns and Alternative Proxies for Transmission Rates

As mentioned above, the sample includes only approved drugs. This selection should be kept in mind when interpreting the results, and the question of whether they extend empirically to negative news remains unanswered here. Theoretically, the same forces that operate for positive shocks would operate on negative ones in a symmetric way. If anything, real-world frictions, such as short-sale constraints, might make underreaction to negative news more pronounced, suggesting the estimates above bound these effects from below. In the structural estimation below, I address this issue directly by modeling the selection.

A second concern is that one of the identifying assumptions in the linear system (27), specifically $E[\epsilon_{mj}\epsilon_{1j}] = 0$, might be violated because media coverage may be higher due to abnormally high pre-approval returns drawing the attention of reporters. In case that this covariation is positive, the regression coefficient might be upwards biased. The usual error-in-variables that biases the coefficient toward zero would offset this bias. Thus, a priori, the bias can go either way.[10]

To try to alleviate concerns about the endogeneity of media coverage, I exploit other features of the institutional setting that can proxy for a drug approval's transmission rate. When the FDA admits a drug for review, the drug is designated a review status. Priority review status is granted to drugs that promise a significant advance over existing treatments. Innovative cancer or HIV/AIDS treatments will often be granted priority review. The public's interest in these drugs is exactly why the regulator speeds up their review; therefore, news of their approval will likely propagate faster than others. Orphan drug status is granted if the treatment is for a disease affecting fewer than 200,000 Americans. The FDA grants an exclusivity period to the drug's sponsor to encourage its development. The general public is likely less interested in approval of orphan drugs, which is precisely why pharmaceutical companies require additional incentives to invest in their development. According to the model, high δ priority drug approvals should exhibit lower post-approval returns than regular drugs. On the other hand, low δ orphan drug approvals should exhibit higher post-approval returns than regular drugs.

The results in Table 4 agree with the model's prediction about covariation between the transmission rate of information and post-approval returns. Priority drug approvals exhibit 1.9% lower post-announcement returns than regular review drugs. Orphan drug approvals exhibit 3.1% higher post-approval returns. Both effects are statistically significant. The inclusion of media exposure in the regressions slightly shrinks these estimates toward zero but all coefficients remain statistically significant. This change should be expected if they all proxy for the same transmission rate. Priority drug and orphan drug dummies are respectively 0.24 and 0.08 correlated with media exposure. Examining the approval returns in columns (3) and (4), we can see that priority drugs exhibit a larger price increase than regular review drugs, though this effect is statistically weak. Finally, the coefficient estimates for pre-approval returns are all statistically no different from zero. Thus, the results of this experiment with alternative proxies for δ are quite similar to the results in Table 3 that use media exposure.

3.3.3 Turnover Effects

A model with time-varying heterogeneity generates trade volume by varying the optimal holdings of various agents in the economy. In the gradual information spread model, an increase in the transmission rate of information results in the transition of a larger fraction of initially

[10] To see this, note that $plim(\hat{\pi}_1) = \pi_1 - \pi_1 \frac{Var[\epsilon_{mj}]}{Var[\delta_j]+Var[\epsilon_{mj}]} + \frac{E[\epsilon_{1j}\epsilon_{mj}]}{Var[\delta_j]+Var[\epsilon_{mj}]}$. The second term on the right is the error-in-variables which biases π_1 toward zero. The last term is an upward bias in the regression from correlation in the errors across equations.

uninformed agents to the informed group. All else equal, faster information spread would result in larger trade volume in period 2, when information spills over according to the transmission rate δ.

Table 5 reports results from an experiment similar to the one in (28), but replacing returns with cumulative abnormal turnover over the same three periods around drug approvals. The sample size is reduced because I have no turnover data for ADRs. The first three columns show regression results on $\hat{\delta}$. Standard errors are again calculated using GMM. In all three return periods, the approval of more interesting drugs exhibits significantly higher turnover. The largest effect is at approval time as reported in column (2). A one standard deviation increase in $\hat{\delta}$ implies a 0.9 percentage points higher approval turnover. This effect is about half of the average approval turnover. Columns (4) through (6) are second stage IV regressions that use WSJ News Pressure as an instrument for media exposure. The first-stage regression is the same as that reported in Table 3. None of the coefficients on the instrumented $\hat{m}$ are statistically different from zero. These results suggest the media is not the cause of variation in turnover in this setting.

Overall, the results on turnover provide additional support for the gradual information spread model. The IV regression results suggest that thinking about the media more as a passive reporter of interesting news and less as an active shaper of public knowledge seems to conform better with the data.

4 Structural Estimation

Having established that the model can explain the negative correlation between media coverage and post-announcement returns and the positive correlation with trade volume, the question remains whether the effects of the gradual spread of information on returns and volume the model implies are quantitatively important and in line with those in the data. In this section, I apply the Indirect Inference estimation method as described in Gourieroux et al. (1993) to the cross-section of drug approvals and their short time series around approvals. Indirect Inference is a generalization of the Simulated Method of Moments of Duffie and Singleton (1993) that allows for exogenous observables in the data-generating-process. Although the predictions from Proposition 4 are about partial derivatives, the numerical analysis of this section allows me to consider the total effects of variation in transmission rates, which changes the equilibrium in the information market. The goal of this exercise is to find a set of model parameters that generate moments close to those in the data. By defining a clear loss function that can be used to judge the fit of the model to the data, this exercise can guide future improvement of the theory (Hansen and Heckman, 1996).

I begin by parameterizing the incidence probability as follows:

$$\Gamma(I; \delta) = \delta I^n, \tag{29}$$

where $\delta \in [0, 1]$ is the transmission rate of the information specific to each drug approval. It parametrizes public interest in the news. The elasticity of $\Gamma(I)$ with respect to I is parametrized by $n \geq 0$. It has an intuitive interpretation as the number of informed agents required for effective information transmission to a randomly matched uninformed agent. One can think of more complicated news as requiring more than a one-on-one meeting for transmission. The information percolation literature studies the aggregation of information in a random matching environment and shows that agents' posterior beliefs evolve according to a similar process (Duffie and Manso, 2007).

4.1 Maximum Likelihood Estimation of Media Coverage Model

My identification strategy relies on observable variation in transmission rates as proxied by the media exposure of various drug approvals. I estimate the parameters of the model (2) using the drug approvals sample. The model of observed media exposure for drug j as constructed in (2) can be summarized as follows:

$$m_j = \max\left\{0, \boldsymbol{x}_j^T \boldsymbol{\beta}_m + \gamma z_j + logit\left(\delta_j\right)\right\}, \tag{30}$$

where $\delta_j \sim Beta(\alpha_\delta, \beta_\delta)$ is the transmission rate of the news story with positive shape and scale parameters α_δ and β_δ. This distributional assumption for the transmission rate is necessary in order to restrict it to the range $(0, 1)$. The censoring of media exposure, m_j, at zero is apparent in Table 1. It can be the result of a fixed cost in publishing a story or of measurement. The exogenous variation provided by the availability of other newsworthy material as measured by WSJ News Pressure (z_j) allows me to identify the parameters of the δ distribution by crowding out media exposure for some observations below the censoring threshold.

Maximum likelihood estimation is a straightforward exercise given this model. As before, I include in the control vector $\boldsymbol{x}_j$ a constant, firm size, and year and month fixed-effects. Table 6 reports MLE estimated parameters and their standard errors. As expected, WSJ News Pressure is negatively correlated with media exposure. The point estimate ($\gamma = -0.71$) is larger in absolute value than the one estimated using the linear media coverage model in section 3, which makes sense because, unlike the linear model, the censored model does not restrict observations where the crowding-out effect is large to have zero media exposure. The left panel of Figure 7 shows the actual mapping from latent residuals ($\hat{\epsilon}_j = m_j - \boldsymbol{x}_j^T \hat{\boldsymbol{\beta}}_m - \hat{\gamma} z_j$) to latent transmission rates via the logistic function as estimated by maximum likelihood. The logit functional form assumption is that media coverage is more sensitive to variation in transmission rates around the boundaries of its support. The resulting distribution of latent $\hat{\delta}$ is shown in the right panel.

Ideally, I would proceed by including conditional distributions of returns and volume given the parameters of the model and maximize the likelihood that the process that generated the sample of drug approvals is the gradual information spread model. Several issues prevent me from proceeding this way. First, expressions for returns and volume are in closed form up to the determination of the price coefficients ratio β_1 in (19), which can potentially depend on different real roots for each parameter set. Second, turnover is not normally distributed, which makes numerical simulation necessary to calculate its conditional means. Finally, MLE requires distributional assumptions about residuals on which I have no strong prior. For these reasons, I proceed by using the cross section of latent transmission rates $\hat{\delta}$ from MLE estimation of (1) as input to the indirect inference procedure. Observations where media exposure is positive use the residual after controlling for other sources of variation. Censored observations use the expected value of δ, which is a non-linear function of $\boldsymbol{x}$ and z. Specifically,

$$\hat{\delta}_j = \begin{cases} logistic\left(m_j - \boldsymbol{x}_j^T \boldsymbol{\beta}_m - \gamma z_j\right) & m_j > 0 \\ E\left[\delta | \boldsymbol{x}_j, z_j, m_j \leq 0\right] & m_j = 0, \end{cases}$$

where $E\left[\delta | \boldsymbol{x}, z, m \leq 0\right] = \frac{Beta\left[logistic\left(-\boldsymbol{x}^T \boldsymbol{\beta}_m - \gamma z\right), \alpha_\delta + 1, \beta_\delta\right]}{Beta\left[logistic\left(-\boldsymbol{x}^T \boldsymbol{\beta}_m - \gamma z\right), \alpha_\delta, \beta_\delta\right]}$.

4.2 Indirect Inference

Moments I choose to match include the mean cumulative abnormal returns as well as the mean cumulative abnormal turnover of the first two periods. The model counterpart to abnormal return is the realized return less the expected risk premium $R_t^e \equiv Q_t - E_0\left[Q_t\right]$. In addition, I include the mean product of returns with $\hat{\delta}$, to capture the dependence of these outcomes on the transmission rate of information. Turnover moments are especially informative about the supply shocks variance. The levels of abnormal returns change considerably when the precision of the signal τ_0 changes. Covariance of returns with transmission rates is informative about the fraction that is early informed and by the competitive information market assumption on the cost of information acquisition.

The sample of drug approvals is selected because no denied approval requests are in my sample. To account for this selection, I use only the right tail of the signal distribution. Specifically,

$$\theta_j = \mu_0 + \tau_0^{-\frac{1}{2}} \upsilon_j, \tag{31}$$

where υ_j is a truncated Gaussian restricted to be positive. Thus all the news simulated is positive news, just as in the sample. Risk tolerance (ϕ) is difficult to identify in this model together with the cost of information (c). To see why, note that in equation (21), multiplying both v and ϕ by a constant leaves the relationship intact. Therefore, I estimate the model for five different values of risk tolerance. I fix the mean supply of the risky asset to 1 and the risk-free rate to 8.7 basis points, which is the average weekly risk free rate over the sampled period. One unfortunate feature of the CARA-Normal framework is that the definition of returns its agents care about is the per-share excess return (Q_t). Following Campbell et al. (1993), I constrain the ex-ante mean dividend (μ_0) so that, on average, $P_0 = 1$. Under this constraint, share returns closely approximate dollar returns, as are the sampled cumulative daily returns in excess of the market return. Furthermore, the specific values of η and τ_0 are not as important as their ratio, which determines the signal-to-noise ratio. Therefore, I normalize the standard deviation of ϵ to 0.1.

Denote R^e_{tj} the excess return of observation $j = 1..N$ in period $t = 1, 2, 3$ whose sample counterpart is the cumulative return in excess of the market return. Let T_{tj} denote the corresponding turnover whose sample counterpart is cumulative turnover in excess of market turnover. The four-period gradual information spread model in section 2, for each observation j, maps the exogenous observable δ_j, unobservable shocks $u_j = [\theta_j, x_{1j}, x_{2j}, \epsilon_j]$, and model parameters $\alpha = [\xi, \tau_0, v, n]$ into a vector of endogenous observables $y_j = [R^e_{1j}, R^e_{2j}, R^e_{3j}, T_{1j}, T_{2j}, T_{3j}] = r(\delta_j, u_j, \alpha)$. The shocks are just a linear function $u_j = \varphi(\upsilon_j, \alpha)$ of a vector of i.i.d Gaussian noise υ_j.

Let $k_j = k(y_j, \delta_j)$ denote the multidimensional function of the data with associated empirical moments $\bar{k}_N = \frac{1}{N}\sum_{j=1}^{N} k(y_j, \delta_j)$. Specifically, my choice of moments is

$$\bar{k}_N = \frac{1}{N}\sum_{j=1}^{N}\left[\begin{array}{cccccccc} R^e_{1j}, & R^e_{2j}, & R^e_{3j}, & T_{1j}, & T_{2j}, & R^e_{1j}\delta_j, & R^e_{2j}\delta_j, & R^e_{3j}\delta_j \end{array}\right]^{\mathrm{T}}.$$

Simulated moments are similarly calculated conditional on the same sample δs, but with $\tilde{y}^h_j(\alpha) = r\left(\delta_j, \tilde{u}^h_j(\alpha), \alpha\right)$, where $h = 1 \ldots H$ indexes simulations. That is, $\tilde{y}^h_j$ depends on the random draw of υ^h_j and on the parameters α. Thus the numerical optimization procedure attempts to match 8 moments with 4 free parameters. The additional moments provide overidentifying restrictions. The indirect estimator $\tilde{\alpha}^H_N$ is obtained by minimizing

$$\min_{\alpha}\left[\bar{k}_N - \frac{1}{NH}\sum_{j=1}^{N}\sum_{h=1}^{H} k\left(\tilde{y}^h_j(\alpha), \delta_j\right)\right]^{\mathrm{T}}\left[\bar{k}_N - \frac{1}{NH}\sum_{j=1}^{N}\sum_{h=1}^{H} k\left(\tilde{y}^h_j(\alpha), \delta_j\right)\right].$$

For each δ_j, I simulate $H = 300$ different random draws of shocks. Since I have no reason to favor one moment over another, I use the identity matrix as a weighting matrix.

4.3 Results

Table 7 presents indirect estimation results for five different values of risk tolerance. The mean payoff μ_0 is constrained by the other parameters so the ex-ante price $P_0 = 1$. More risk-averse agents require a higher expected payoff to value the risky asset at the same price. Both the precision of the supply shocks (ξ) and that of the signal (τ_0) are decreasing in risk aversion. By examining the values of the minimized objective, we can see the model with risk aversion $\phi^{-1} = 5$ performs better than the rest, and I therefore focus on these results for the remainder of the analysis.

With a constant absolute risk aversion of 5, the cost of inside information is about 52% of an uninformed agent's position in the risky asset at time 1. This cost can be interpreted as the expected cost of trading on private information with a chance of adverse legal consequences.

On average, agents informed about a drug approval before the official announcement control only a tiny fraction of invested wealth; it is on the order of 10^{-6}. The signal-to-noise ratio for the average drug approval $\frac{Var[\theta]}{Var[\theta]+Var[\epsilon]}$ is about 10%, which means the reduction in uncertainty about the value of the drug developing firm that is associated with an approval is small relative to the remaining uncertainty. This ratio is consistent with DiMasi (2001), which studies the drug approval process and estimates the probability that a drug will be approved conditional on surviving to the marketing application stage is about 75%. The elasticity n is estimated at 0.014.

Non-informational supply shocks have a standard deviation of 1% of the total supply of the stock. Compared with the 6.5% standard deviation of approval turnover in the sample, this number is small. Thus the model requires only a small amount of non-informational trading to prevent prices from fully revealing the news. Since there are no borrowing constraints in the model, informed agents can hold a highly levered position at time 1 to exploit their informational advantage. The more these agents can borrow, the more informative are prices. Extending the model to incorporate borrowing constraints will decrease the required amount of noise. Therefore, in a sense, this estimate of the standard deviation of the supply shocks is an upper bound.

Table 8 provides a comparison between sample moments and their simulated counterparts using the estimated parameters. The goal of the numerical procedure is to minimize the discrepancy between the two, which I report in the last column. The covariances are those implied by the 8 matched moments and the average δ of 0.54. The pre-approval average return, as well as its covariance with the transmission rate, match well. The approval return and turnover do not match as well, but are still reasonable. Turnover moments implied by the model actually match the data fairly well. The turnover level is fairly easy to match with this model using the parameter ξ, which controls the precision of the supply shocks. The negative post-approval return in the sample cannot be generated by this model which features only rational agents. Behavioral extensions can help there, but the standard deviation around the mean cannot statistically reject that it is zero and worth explaining. Pre-approval returns are increasing in δ in both the sample and the model. The approval return generates the opposite sign on the covariation, but again, the point estimate is quite noisy. The main stylized fact, namely, the negative covariation between post-approval returns and δ is present in both the sample and the model.

4.4 Simulated Regression Analysis

The structural estimates provide us with a relevant set of parameters that can be used to answer the question: what should we expect from a linear regression when the data generating process is that modeled in this paper? To answer this question, I simulate a large cross-section 10 times as large as the sample, using the estimated parameters. I draw the random shocks as before, and in addition, draw δ from a Beta distribution with parameters estimated using MLE in section 4.1. These draws approximate the theoretical distribution implied by the data without conditioning on δ.

First, I consider the dependence of price drift as measured by the post-approval return on the transmission rate, previous returns, and concurrent turnover. Matching these moments was not the objective of the optimization procedure above, which makes it less likely that they match those in the data. Thus they provide an additional source of validation for the model. Table 9 compares the regression results.

The results show that the model can certainly generate the negative covariation between transmission rates and post-approval returns. If anything, the effect predicted is too large. Regression coefficients of post-approval returns on their lagged values match fairly well. Little covariation seems to exist between post-approval turnover with returns in the sample or in the simulated data. Finally, note that the level of R_3^e is somewhat higher in the model than it is in the data. The reason is that the reduction in uncertainty in the model is absolute, whereas a considerable amount of uncertainty remains in the real world. An extension of the model to more periods can fix this issue.

Table 10 reports regression results from a similar exercise, but with post-approval turnover as the dependent variable and its lagged values as well as δ as explanatory variables. The level matches well. Covariance with the transmission rate is statistically insignificant in the sample but strongly significant in the simulated data. The model-predicted effect, however, is quite small. One standard deviation in δ (0.39) results in a 0.03 percentage points lower turnover. The regression coefficient on approval turnover is positive, significant, and well matched. The coefficient on pre-approval turnover, which is statistically significant in both, is positive in the sample but negative in the model. Finally, so much of the variation in turnover is persistent that the R-squared in the sample is 0.59. Model-generated data produce an R^2 of 0.86, which is high as well.

4.5 The Implied Market for Information

Proposition 3 decomposes the value of information into three components. We can use the parameter estimates to get a sense of their relative magnitude. Consider an average transmission rate and an average informed fraction as calculated in the simulated sample. The first component is 0.62, whereas the second and third components are on the order of 10^{-5}. Thus at least in the neighborhood of the estimated parameters, the value of information stems entirely from the intertemporal growth rate of uninformed agents' precision Ω_v. This growth is not exogenous but rather an equilibrium outcome. Because it is the dominant component of the value of information, given the constant cost of information c, either there are no informed in equilibrium, or

$$\Omega_v = \frac{\tau_2 + \eta}{\tau_1} \approx e^{\frac{2}{\phi} R^2 c}, \tag{32}$$

which implies the ratio $\frac{\tau_2}{\tau_1}$ is kept constant across the various transmission rates.

This nature of investor demand for information together with the fixed cost of acquiring it produce an interesting equilibrium outcome in the market for information plotted in Figure 8. The plots show that no information is acquired in equilibrium when δ is low, that is, for uninteresting news. Around $\delta = 0.25$, the fraction informed rises sharply and then slowly declines. Thus faster-spreading information is purchased at a high rate, while the fastest-spreading news is less valuable.

The parametrization of the incidence probability in (29) has a discontinuity at $n = 0$. When n is zero, a meeting with an informed agent is not necessary for information transmission. A high δ results in a high I_2 even if I_1 is close to zero. However, when n approaches zero from the right, a small I_1 implies a small I_2 regardless of δ. Unreported attempts to generate the covariance pattern in the data with $n = 0$ have been unsuccessful, because to match the magnitudes of covariation between media exposure and stock returns, the fixed cost of acquiring as well as the dependency of $\Gamma(I_1; \delta)$ on I_1 are essential. The fixed cost affects the extensive margin and creates two effective types of news, one with positive equilibrium information acquisition in which pre-approval returns are high and post-approval returns are low, and another type with no information being pursued until the post-approval period. Thus endogenous information acquisition accentuates the covariation between the transmission rate of information and returns.

The policy implications of this exercise are interesting. Consider the increase in the transmission rate of information that has resulted from the technological progress in communication technology. Holding fixed the cost of trading on inside information, we would expect more investment in the acquisition of private information to occur in the Internet age than before, because the present value of informational rents is larger and is subject to less noise. Policies like the one followed by government agencies such as the FDA to publicize drug approvals as soon as possible aims to reduce the asymmetry of information in the market. But this policy also results in the perhaps undesirable increased incentive to acquire private information before the announcement.

5 Conclusion

The suggested asset pricing model is useful for the study of the relation between the media and the stock market. I show that thinking about the media as responding to interesting news as opposed to shaping stock market reaction to news can explain the covariation between media coverage and the stock market within a rational framework with a plausible informational friction. Using the WSJ News Pressure instrument for media coverage, I find no evidence of the media causing the observed variation in post-approval returns. Reduced-form tests reveal that, consistent with the model, drug approvals that receive more media exposure on the approval day and the next, exhibit higher turnover on those days, and lower returns in the subsequent week. A third prediction of the model, that pre-approval returns are increasing in the transmission rate of information cannot be rejected by the data.

The structural estimation exercise reveals that effects the model generates are quantitatively similar to those in the data and that the parameter estimates are reasonable. Importantly, I find that the amount of non-informational trading (noise) required to keep prices from fully revealing all information is small. Analyzing the model around the estimated parameters, I find that the main source of value from becoming informed is the intertemporal growth in the precision of the uninformed about the random payoff. This feature results in a hump-shaped demand for information as a function of the transmission rate of information. In information market equilibrium, no one pursues uninteresting, slowly propagating news before a drug approval, because the fixed cost of information is prohibitively high. Faster-spreading information is purchased at a higher rate, whereas the fastest-spreading news is somewhat less valuable. This unique feature of the model accentuates the covariation between transmission rates and the demand for the risky asset, by influencing the extensive margin of information acquisition by the population as a whole. I find this feature is essential for generating the large covariation between media exposure and stock returns and volume.

The empirical success of the model in explaining the stylized facts about media coverage and stocks suggests that the process by which information spreads can be important for asset pricing. I show that the introduction of such a process to a simple yet relevant asset pricing model generates rich testable predictions about returns and volume around news releases. Although this study is silent about the importance of such informational frictions in other settings, I focus here on drug approvals because they allow for cleaner identification of the effects in question compared with other events previously documented in the literature. Nonetheless, the intuitive thought process of an investor who receives news and considers not only who else is informed today, but also how many others will be informed tomorrow, seems robust once we deviate from a fully revealing prices setup. That investors in asset markets expand considerable effort and resources for the attainment of information, even a mere split second before their peers, is difficult to explain within the canonical model, but is quite reasonable once informational gains are allowed.

One implication of my findings concerns the informational efficiency of capital markets. Variation in transmission rates of information means that a finer distinction can be made between the informational content of prices of different projects and firms. The informational frictions seem less important for an economic activity society as a whole finds interesting. Thus what may seem to be an arbitrary influence of fads and fashions on prices can actually be an efficiency-enhancing mechanism. A second practical implication for firms that aim to raise capital in financial markets concerns their expenditure on investor relations. If the media are in fact just a mirror of what investors find interesting, the goal of a successful publicity campaign is then to create an appealing narrative around news, one that captures the imagination of investors and makes positive information propagate faster through social networks. Viral investor relations, much like viral marketing, can therefore be more successful than traditional attempts at increasing firm visibility.

Finally, two empirical measures I constructed above have much broader applicability than the context of this study and can be used in future research. Advertising rates can be used for the measurement of media exposure and its relationship with the economy. For example, testing whether post-earnings announcement drift is affected by media exposure would be interesting. Identification of situations in which media coverage is responsible for the variation in the data as opposed to passively responding to the public's interest in the news will plague most such studies. For this reason, the proxy for business news pressure I construct using daily data on the number of pages in the *Wall Street Journal* can be used in such circumstances as a plausibly exogenous source of variation.

References

Barber, B., Odean, T., 2008. All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. Review of Financial Studies 21 (2), 785–818.

Birkhoff, G., Mac Lane, S., 1997. A Survey of Modern Algebra, 5th Edition. A. K. Peters.

Brunnermeier, M., 2001. Asset pricing under asymmetric information: Bubbles, crashes, technical analysis, and herding. Oxford University Press, USA.

Campbell, J. Y., Grossman, S. J., Wang, J., 1993. Trading volume and serial correlation in stock returns. The Quarterly Journal of Economics 108 (4), 905–939.

Chan, W. S., 2003. Stock price reaction to news and no-news: drift and reversal after headlines. Journal of Financial Economics 70 (2), 223–260.

Cohen, L., Frazzini, A., Malloy, C., 2008. The small world of investing: Board connections and mutual fund returns. Journal of Political Economy 116 (5), 951–979.

Dawkins, R., 1989. The Selfish Gene (New Edition). Oxford University Press, Oxford, New York.

DellaVigna, S., Pollet, J. M., 2009. Investor Inattention and Friday Earnings Announcements. Journal of Finance 64 (2), 709–749.

DeMarzo, P. M., Vayanos, D., Zwiebel, J., 2003. Persuasion bias, social influence, and unidimensional opinions. The Quarterly Journal of Economics 118 (3), pp. 909–968.

DiMasi, J. A., 2001. Risks in new drug development: approval success rates for investigational drugs. Clinical Pharmacology and Therapeutics 69 (5), 297–307.

Dougal, C., Engelberg, J., Garcia, D., Parsons, C., 2010. Journalists, Persuasion, and Asset Prices. University of North Carolina - Chapel Hill working paper.

Duffie, D., Manso, G., 2007. Information percolation in large markets. American Economic Review 97 (2), 203–209.

Duffie, D., Singleton, K. J., 1993. Simulated moments estimation of markov models of asset prices. Econometrica 61 (4), 929–952.

Editor and Publisher, 1998. Editor and Publisher International Yearbook. New York: Editor and Publisher Company.

Eisensee, T., Stromberg, D., 2007. News Floods, News Droughts, and US Disaster Relief. Quarterly Journal of Economics 122 (2), 693–728.

Ellison, S. F., Mullin, W. P., 2007. Gradual incorporation of information into stock prices: An empirical study of informed trading. Working Paper.

Engelberg, J., Parsons, C. A., 2010. The causal impact of media in financial markets. Forthcoming, Journal of Finance.

Fang, L., Peress, J., 2009. Media Coverage and the Cross-section of Stock Returns. The Journal of Finance 64 (5), 2023–2052.

Ferguson, J. M., 1963. The advertising rate structure in the daily newspaper industry. Ph.D. thesis, University of Chicago.

Ferguson, J. M., October 1983. Daily newspaper advertising rates, local media cross-ownership, newspaper chains, and media competition. Journal of Law and Economics 26 (3), 635–654.

Gale Research, 1998. Gale Directory of Publications and Broadcast Media. Detroit, Mich: Gale Research.

Gourieroux, C., Monfort, A., Renault, E., 1993. Indirect inference. Journal of Applied Econometrics 8, S85–S118.

Gray, W. R., 2010. Do Hedge Fund Managers Identify and Share Profitable Ideas? Working Paper.

Grossman, S. J., 1995. Dynamic asset allocation and the informational efficiency of markets. The Journal of Finance 50 (3), 773–787.

Grossman, S. J., Stiglitz, J. E., 1980. On the impossibility of informationally efficient markets. The American Economic Review 70 (3), 393–408.

Hansen, L. P., Heckman, J. J., 1996. The empirical foundations of calibration. The Journal of Economic Perspectives 10 (1), 87–104.

Hirshleifer, D., Subrahmanyam, A., Titman, S., 1994. Security analysis and trading patterns when some investors receive information before others. The Journal of Finance 49 (5), 1665–1698.

Holden, C. W., Subrahmanyam, A., 2002. News events, information acquisition, and serial correlation. The Journal of Business 75 (1), 1–32.

Hollis Directories, 1998. Willings Press Guide 1998. Teddington, Middlesex, UK, Hollis Directories Ltd.

Hong, D., Hong, H. G., Ungureanu, A., 2011. An Epidemiological Approach to Opinion and Price-Volume Dynamics. SSRN eLibrary.

Hong, H., Kubik, J. D., Stein, J. C., 2005. Thy neighbor's portfolio: Word-of-mouth effects in the holdings and trades of money managers. The Journal of Finance 60 (6), 2801–2824.

Hong, H., Stein, J. C., 1999. A unified theory of underreaction, momentum trading, and overreaction in asset markets. The Journal of Finance 54 (6), 2143–2184.

Huberman, G., Regev, T., February 2001. Contagious speculation and a cure for cancer. Journal of Finance.

Merton, R. C., July 1987. A simple model of capital market equilibrium with incomplete information. Journal of Finance 42 (3), 483–510.

Neuhierl, A., Scherbina, A. D., Schlusche, B., 2010. Market Reaction to Corporate News and the Influence of the Financial Crisis. SSRN eLibrary.

Oxbridge Communications, 1998. The Standard Periodical Directory. New York, N.Y., Oxbridge Communications, Inc.

Peress, J., 2008. Media coverage and investors' attention to earnings announcements. Working Paper.

Romer, P. M., 1990. Endogenous technological change. Journal of Political Economy 98 (s5), S71.

Shaw, D., 2007. A fading taboo: Paper by paper, advertising is making its way onto the nation's front pages and section fronts. American Journalism Review 29 (3), 34–39.

Shiller, R. J., Pound, J., 1989. Survey evidence on diffusion of interest and information among investors. Journal of Economic Behavior and Organization 12, 47–66.

Solomon, D., 2009. Selective publicity and stock prices. Working Paper.

Stamm, 1998. Stamm Leitfaden für Presse und Werbung. Essen: Stamm Verlag.

Stein, J. C., 2009. Conversations among competitors. American Economic Review (forthcoming).

Tetlock, P. C., 2007. Giving Content to Investor Sentiment: The Role of Media in the Stock Market. The Journal of Finance 62 (3), 1139–1168.

Tetlock, P. C., 2010. Does public financial news resolve asymmetric information? Review of Financial Studies 23, 3520–3557.

U.S. Food and Drug Administration, March 1998. CDER Manual of Policies and Procedures. Communicating Drug Approval Information. Online.

Veldkamp, L. L., 2006. Media frenzies in markets for financial information. The American Economic Review 96 (3), 577–601.

Wang, J., 1994. A model of competitive stock trading volume. The Journal of Political Economy 102 (1), 127–168.

A Solution Details

A.1 Proof of Proposition 1 (Existence of REE)

The solution method for the rational expectations equilibrium follows the guess-and-verify approach as usual. I first conjecture that prices are linear in the state, and solve for agents' optimal behavior given these prices in each period starting with the last one and working backward through time. Finally, I impose state-by-state market clearing and verify that equilibrium prices satisfy the conjectured form (8).

A.1.1 Proof of Lemma 1 (Learning from Prices)

The inference problem of the uninformed is to form a posterior belief about the fundamental value θ conditional on the current-period price. The prior belief of uninformed at time t is that $\theta|P_t \sim \mathcal{N}(\mu_t, \tau_t^{-1})$. The distribution of P_{t+1} conditional on θ can be attained by rearranging the price conjecture: $x_{t+1} = \frac{P_{t+1}-a_{t+1}-b_{t+1}\Delta_{t+1}}{c_{t+1}} \sim \mathcal{N}(0, \xi^{-1})$. Substituting these into Bayes' rule for the posterior distribution of θ and solving for μ_{t+1}, which appears also in Δ_{t+1}, yields the solution in (9).

A.1.2 Period 2 Optimal Behavior

The indirect utility of an agent of type $i \in \{U, I\}$ in period 2 is

$$V_2^i(W_2^i) \quad = \quad \max_{q_2^i} -e^{-\frac{1}{\phi}\left(RW_2^i + q_2^i E_2^i[Q_3] - \frac{1}{2\phi}\left(q_2^i\right)^2 \Psi_2^i[Q_3]^{-1}\right)},$$

where $Q_3 = u - RP_2$ is normally distributed. The expectation can be evaluated and the first-order condition on q_2^i provides the result in (11). Using the optimal demands the value function is

$$V_2^i(W_2^i) \quad = \quad -e^{-\frac{1}{\phi}RW_2^i - \frac{1}{2}\Psi_2^i[Q_3]E_2^i[Q_3]^2}. \tag{33}$$

A.1.3 Period 2 Market-Clearing Prices

By substituting the optimal holdings from (11) into the market-clearing condition, I verify the equilibrium P_2 indeed has the desirable form. Define $U_t \equiv 1 - I_t$ as the fraction uninformed at time t. Noting that $E_2^I[Q_3] = E_2^U[Q_3] + \Delta_2$, we get

$$P_2 = \frac{\mu_1 - k_2 a_2}{(R - k_2)} + \frac{I_2\Psi_2^I[Q_3]}{(R-k_2)\left(I_2\Psi_2^I[Q_3] + U_2\Psi_2^U[Q_3]\right)}\Delta_2 - \frac{X + x_2}{\phi(R-k_2)\left(I_2\Psi_2^I[Q_3] + U_2\Psi_2^U[Q_3]\right)}.$$

The first object that can be easily determined is the ratio: $\beta_2 = \frac{b_2}{c_2} = -\phi I_2 \Psi_2^I[Q_3] = -\phi I_2 \eta$. This result reveals that the relative signal strength in prices is always non-positive ($\beta_2 \leq 0$) and decreasing in I_2 ($\frac{\partial \beta_2}{\partial I_2} \leq 0$). Next we can determine the Kalman gain:

$$k_2 = R\left[\frac{\xi\phi^2 I_2\eta\left(I_2\Psi_2^I[Q_3] + U_2\Psi_2^U[Q_3]\right)}{\tau_1 + \xi\phi^2 I_2\eta\left(I_2\Psi_2^I[Q_3] + U_2\Psi_2^U[Q_3]\right)}\right] < R, \tag{34}$$

where the result (34) will turn out useful. The price coefficients are

$$\begin{aligned} a_2 &= \frac{\mu_1}{R} - \frac{X}{\phi R\left(I_2\Psi_2^I[Q_3] + U_2\Psi_2^U[Q_3]\right)} = \frac{\mu_1}{R} - \frac{E_0[Q_3]}{R} \\ b_2 &= c_2\beta_2 = \frac{I_2\Psi_2^I[Q_3]}{R\left(I_2\Psi_2^I[Q_3] + U_2\Psi_2^U[Q_3]\right)} + \frac{\xi\beta_2^2}{R\tau_1} \\ c_2 &= -\frac{1}{\phi(R-k_2)\left(I_2\Psi_2^I[Q_3] + U_2\Psi_2^U[Q_3]\right)} = -\frac{1}{\phi R\left(I_2\Psi_2^I[Q_3] + U_2\Psi_2^U[Q_3]\right)} + \frac{\xi\beta_2}{R\tau_1} \end{aligned} \tag{35}$$

Finally, note that the precisions of each agent about final period returns are $\Psi_2^I[Q_3] = \eta$ and $\Psi_2^U[Q_3] = \left[\eta^{-1} + \tau_2^{-1}\right]^{-1}$. We have solved for the equilibrium price P_2 taking μ_1 and τ_1 as given. Beliefs of both agents can be derived by substituting these coefficients into the precisions.

A.1.4 Period 2 Expected Returns

The expected returns of each agent at time 2 about time 3 returns can be expressed as a linear function of the unexpected return at time 2. Define the innovation to time 2 return for agent of type i $\tilde{Q}_2^i \equiv Q_2 - E_1^i[Q_2]$ and note that for an uninformed agent, $\tilde{Q}_2^U = b_2\Delta_2 + c_2 x_2$, and for an informed agent, $\tilde{Q}_2^I = b_2\left(\Delta_2 - E_1^I[\Delta_2]\right) + c_2 x_2$. The difference between their innovations is that whereas the uninformed expect their error in period 2 to be zero, the informed have superior information about how the uninformeds' beliefs will evolve. Specifically, $E_1^I[\Delta_2] = \frac{\tau_1}{\tau_2}\Delta_1$ so that the informed expect the error to decrease in the next period by the ratio of the current precision to the future precision. Agents' future expected returns can be expressed as in (17), where $E_1^U[Q_3] = \mu_1 - Ra_2$ and the loading on the innovation $\rho_1^U = k_2 - R$. From (34), we can see that $\rho_1^U < 0$. For informed agents' expected returns, $E_1^I[Q_3] = \theta - R\left(a_2 + b_2 E_1^I[\Delta_2]\right)$ and $\rho_1^I = -R$, which is negative. Thus the expected returns of both agents in period 2 are perfectly negatively correlated with time 2 returns.

A.1.5 Period 1 Informed Agents Behavior

Period 1 informed agents know they will remain informed next period and therefore have the following indirect utility:

$$V_1^I(W_1) = \max_{q_1^I} -e^{-\frac{1}{\phi}R^2 W_1} E_1^I\left[e^{d + b\tilde{Q}_2^I + A(\tilde{Q}_2^I)^2}\right], \tag{36}$$

where $d = -\frac{1}{\phi}Rq_1^I E_1^I[Q_2] - \frac{1}{2}\Psi_2^I[Q_3]E_1^I[Q_3]^2$, $b = -\Psi_2^I[Q_3]E_1^I[Q_3]\rho_1^I - \frac{1}{\phi}Rq_1^I$ and $A = -\frac{1}{2}\Psi_2^I[Q_3](\rho_1^I)^2$. Thus the expectation is of an exponential of a quadratic in $\tilde{Q}_2^I \sim \mathcal{N}(0, \Psi_1^I[Q_2]^{-1})$. Using the moment generating function of the Wishart distribution, it can be shown that for a multivariate normal random vector $z \sim \mathcal{N}(0, \Sigma)$, the following result holds (see Brunnermeier, 2001):

$$E\left[e^{z^T Az+b^T z+d}\right] = |Id-2\Sigma A|^{-\frac{1}{2}} e^{\frac{1}{2}b^T(Id-2\Sigma A)^{-1}\Sigma b+d}, \tag{37}$$

which means the value function for the informed investor is

$$V_1^I(W_1) = \max_{q_1^I} -e^{-\frac{1}{\phi}R^2 W_1}\left(\Omega_1^I\right)^{-\frac{1}{2}} e^{\frac{1}{2}\Psi_2^I[Q_3]^2 E_1^I[Q_3]^2\left(\rho_1^I\right)^2\Lambda_1^I+\frac{1}{\phi}Rq_1^I\Lambda_1^I\Psi_2^I[Q_3]E_1^I[Q_3]\rho_1^I+\frac{1}{2}\frac{R^2}{\phi^2}\left(q_1^I\right)^2\Lambda_1^I-\frac{1}{\phi}Rq_1^I E_1^I[Q_2]-\frac{1}{2}\Psi_2^I[Q_3]E_1^I[Q_3]^2},$$

where $\Omega_1^I \equiv 1+\frac{\Psi_2^I[Q_3](\rho_1^I)^2}{\Psi_1^I[Q_2]}$ and $\Lambda_1^I \equiv \left(\Omega_1^I\right)^{-1}\Psi_1^I[Q_2]^{-1}$. The first-order optimality condition for the informed agents at time 1 gives the optimal holding for informed agents entering period 2 (18). Evaluating the right-hand side of the value function at the optimum investment choice we get

$$V_1^I(W_1) = -\left(\Omega_1^I\right)^{-\frac{1}{2}} e^{-\frac{1}{\phi}R^2 W_1-\frac{1}{2}E_1^I[Q_2]^2\Psi_1^I[Q_2]-\frac{1}{2}\Psi_2^I[Q_3]\left(E_1^I[Q_3]-\rho_1^I E_1^I[Q_2]\right)^2}.$$

The variance informed agents is $\Psi_1^I[Q_2]^{-1} = b_2^2 Var_1^I[\Delta_2] + c_2^2 Var_1^I[x_2] + 2b_2c_2Cov_1^I[\Delta_2 x_2]$. The covariance arises because supply shocks are partially confused by the uninformed as payoff shocks: $Cov_1^I[\Delta_2 x_2] = \frac{-k_2c_2\frac{1}{\xi}}{1+k_2b_2} = \frac{-\beta_2}{\tau_2}$ and the variance of Δ_2 is $Var_1^I[\Delta_2] = \frac{k_2^2c_2^2Var_1^I[x_2]+2k_2^2b_2c_2Cov_1^I[\Delta_2x_2]}{1-k_2^2b_2^2}$. Therefore, we can express the variance of Q_2 simply as $\Psi_1^I[Q_2]^{-1} = \frac{c_2^2\tau_1^2}{\xi\tau_2^2}$.

A.1.6 Period 1 Uninformed Agents' Behavior

Period 1 uninformed agents face a more complicated problem. With probability $\Gamma(I_1)$, they become informed next period, otherwise they remain uninformed but learn more about the payoff through market price. Their problem is:

$$V_1^U(W_1) = \max_{q_1^U} E_1^U\left[\Gamma(I_1)E_2^U\left[V_2^I(W_2)\right]+(1-\Gamma(I_1))V_2^U(W_2)\right],$$

where I have used the law of iterated expectations. Next notice the only source of randomness in $E_2^U\left[V_2^I(W_2)\right]$ is $\Delta_2 \sim \mathcal{N}\left(0, Var_2^U[\Delta_2]\right)$, so

$$E_2^U\left[V_2^I(W_2)\right] = -e^{-\frac{1}{\phi}RW_2-\frac{1}{2}\Psi_2^I[Q_3]E_2^U[Q_3]^2}E_2^U\left[e^{-\Psi_2^I[Q_3]E_2^U[Q_3]\Delta_2-\frac{1}{2}\Psi_2^I[Q_3]\Delta_2^2}\right],$$

where once again, we use the result in (37) with $A=-\frac{1}{2}\Psi_2^I[Q_3]$, $b=-\Psi_2^I[Q_3]E_2^U[Q_3]$ and $d=0$ to get

$$E_2^U\left[e^{-\Psi_2^I[Q_3]E_2^U[Q_3]\Delta_2-\frac{1}{2}\Psi_2^I[Q_3]\Delta_2^2}\right] = \sqrt{\frac{\tau_2}{\tau_{2+\eta}}}e^{\frac{1}{2}\Psi_2^I[Q_3]^2E_2^U[Q_3]^2\frac{\tau_2}{\tau_{2+\eta}}\tau_2^{-1}}$$

so that

$$E_2^U\left[V_2^I(W_2)\right] = -\sqrt{\frac{\tau_2}{\tau_{2+\eta}}}e^{-\frac{1}{\phi}RW_2-\frac{1}{2}\Psi_2^U[Q_3]E_2^U[Q_3]^2} = \sqrt{\frac{\tau_2}{\tau_{2+\eta}}}V_2^U(W_2).$$

The value function of the uninformed at time 2 is proportional to the expected value function of the informed in time 2. This proportionality simplifies the uninformed's problem at time 1 to:

$$V_1^U(W_1) = \max_{q_1^U}\left[\Gamma(I_1)\sqrt{\frac{\tau_2}{\tau_{2+\eta}}}+(1-\Gamma(I_1))\right]E_1^U\left[V_2^U(W_2)\right],$$

where the constant term multiplying the expectation does not depend on the investment choice. Since $V_2^U(W_2)$ and the expectations of the uninformed take the same form as those of the informed agent, their first-order condition takes the same form. The uninformeds' value function at the optimum is

$$V_1^U(W_1) \quad = \quad -\left[\Gamma(I_1)\sqrt{\frac{\tau_2}{\tau_{2+\eta}}} + (1-\Gamma(I_1))\right]\left(\Omega_1^U\right)^{-\frac{1}{2}} e^{-\frac{1}{\phi}R^2 W_1 - \frac{1}{2}E_1^U[Q_2]^2\Psi_1^U[Q_2] - \frac{1}{2}\Psi_2^U[Q_3]\left(E_1^U[Q_3]-\rho_1^U E_1^U[Q_2]\right)^2}.$$

A.1.7 Period 1 Market-Clearing Prices

By substituting the optimal holdings into the market-clearing condition, I verify that equilibrium P_1 indeed has the desirable form. Noting that $E_1^I[Q_2] = E_1^U[Q_2] + b_2\frac{\tau_1}{\tau_2}\Delta_1$ and that $E_1^I[Q_3] = E_1^U[Q_3] + \left(1 - Rb_2\frac{\tau_1}{\tau_2}\right)\Delta_1$, the market clearing condition at time 2 gives

$$\begin{aligned} P_1 &= \frac{\mu_1}{R^2} - \frac{X}{\phi R^2\left(I_2\Psi_2^I[Q_3] + U_2\Psi_2^U[Q_3]\right)} - \frac{\frac{R}{\phi}X + \left(I_1\rho_1^I\Psi_2^I[Q_3] + U_1\rho_1^U\Psi_2^U[Q_3]\right)E_1^U[Q_3]}{R\left(I_1\Omega_1^I\Psi_1^I[Q_2] + U_1\Omega_1^U\Psi_1^U[Q_2]\right)} \\ &+ \frac{I_1\left(\Omega_1^I\Psi_1^I[Q_2]b_2\frac{\tau_1}{\tau_2} - \rho_1^I\Psi_2^I[Q_3]\left(1 - Rb_2\frac{\tau_1}{\tau_2}\right)\right)}{R\left(I_1\Omega_1^I\Psi_1^I[Q_2] + U_1\Omega_1^U\Psi_1^U[Q_2]\right)}\Delta_1 - \frac{x_1}{\phi\left(I_1\Omega_1^I\Psi_1^I[Q_2] + U_1\Omega_1^U\Psi_1^U[Q_2]\right)}. \end{aligned}$$

Determining the coefficients, beginning with $\beta_1 \equiv \frac{b_1}{c_1} = -\phi I_1\left(\Psi_2^I[Q_3] + \frac{1}{R}b_2\frac{\tau_1}{\tau_2}\Psi_1^I[Q_2]\right) = -\phi I_1\left(\frac{k_2}{R}\tau_2 + \eta\right)$. We can see that $\frac{\partial\beta_1}{\partial I_1} < 0$, which confirms the intuition that the more informed agents there are in period 1, the more informative prices are about this hidden signal. Furthermore, both $\frac{\partial k_2}{\partial I_2} > 0$ and $\frac{\partial\tau_2}{\partial I_2} > 0$, which implies $\frac{\partial\beta_1}{\partial I_2} < 0$ as well. That is, ceteris paribus, the informativeness of prices at time 1 is increasing not only with the fraction informed at time 1, but also with the fraction informed at time 2. Next, we can determine the Kalman gain:

$$k_1 = \frac{\xi b_1}{\tau_0 c_1^2} \quad = \quad R^2\left[\frac{\xi\phi^2 I_1\left(\frac{k_2\tau_2}{R} + \eta\right)\left(I_1\Omega_1^I\Psi_1^I[Q_2] + U_1\Omega_1^U\Psi_1^U[Q_2]\right)}{R^2\tau_0 + \xi\phi^2 I_1\left(\frac{k_2\tau_2}{R} + \eta\right)\left(I_1\Omega_1^I\Psi_1^I[Q_2] + U_1\Omega_1^U\Psi_1^U[Q_2]\right)}\right]. \tag{38}$$

Notice that since $k_2 > 0$, the term in brackets is between zero and one so $0 < k_1 < R^2$ and therefore $\left(1 - \frac{k_1}{R^2}\right) > 0$. The price coefficients are

$$\begin{aligned} a_1 &= \frac{\mu_0}{R^2} - \frac{E_0[Q_3]}{R^2} - \frac{E_0[Q_2]}{R} \\ b_1 &= c_1\beta_1 = -\frac{\beta_1}{\phi\left(I_1\Omega_1^I\Psi_1^I[Q_2] + U_1\Omega_1^U\Psi_1^U[Q_2]\right)} + \frac{\xi\beta_1^2}{R^2\tau_0} \\ c_1 &= -\frac{1}{\phi\left(I_1\Omega_1^I\Psi_1^I[Q_2] + U_1\Omega_1^U\Psi_1^U[Q_2]\right)} + \frac{\xi\beta_1}{R^2\tau_0}. \end{aligned} \tag{39}$$

A.1.8 Period 1 Expected Returns

The expected returns of each agent at time 1 about times 2 and 3 returns can be expressed as a linear function of the unexpected return at time 1. Define the innovation to time 1 return for agent of type i $\tilde{Q}_1^i \equiv Q_1 - E_0^i[Q_1]$. Since all agents are ex-ante identical and uninformed, we can discard the type superscripts and note that for both agents, $\tilde{Q}_1 = b_1\Delta_1 + c_1x_1$. Future expected returns can be expressed as in (22), where $E_0[Q_2] = \frac{\frac{R}{\phi}X + \left(I_1\rho_1^I\Psi_2^I[Q_3] + U_1\rho_1^U\Psi_2^U[Q_3]\right)E_0[Q_3]}{I_1\Omega_1^I\Psi_1^I[Q_2] + U_1\Omega_1^U\Psi_1^U[Q_2]}$ and the loading on the innovation $\rho_0 = \frac{k_1}{R} - R$. From (38) we can see $\rho_0 < 0$. Thus the period 1 return is perfectly negatively correlated with the uninformed agent's expected return from time 1 to time 2. Furthermore, the expected return of the uninformed at time 1 about period 3 returns is independent of the period 1 return. In both periods 0 and 1, the uninformed expect $E_0[Q_3] = \frac{X}{\phi\left(I_2\Psi_2^I[Q_3] + U_2\Psi_2^U[Q_3]\right)}$.

A.1.9 Period 0 Optimal Behavior

Ex-ante, all agents are identical and uninformed about θ. There is no noise, and nothing can be learned from equilibrium prices. The assumption about information market equilibrium that makes uninformed agents just as well off as informed agents once the latter pay for their information simplifies the problem of the agent at time 0. All he has to do is maximize the expected value of either $V_1^I(W_1)$ or $V_1^U(W_1)$. For simplicity, we maximize the uninformed expected utility. Thus the problem of the agent is to maximize:

$$
\begin{aligned}
V_0(W_0) &= \max_{q_0} - \left[\Gamma(I_1)\sqrt{\frac{\tau_2}{\tau_{2+\eta}}} + (1-\Gamma(I_1))\right] \left(\Omega_1^U\right)^{-\frac{1}{2}} e^{-\frac{1}{\phi}R^3 W_0 - \frac{1}{\phi}R^2 q_0 E_0[Q_1] - \frac{1}{2}\Psi_2^U[Q_3]E_0[Q_3]^2 + E_0[Q_2]\Psi_2^U[Q_3]E_0[Q_3]\rho_1^U - \frac{1}{2}E_0[Q_2]^2\Omega_1^U\Psi_1^U[Q_2]} \\
&\times \quad E_0\left[e^{\left[-\frac{1}{\phi}R^2 q_0 - E_0[Q_2]\rho_0\Omega_1^U\Psi_1^U[Q_2] + \rho_0\Psi_2^U[Q_3]E_0[Q_3]\rho_1^U\right]\tilde{Q}_1 - \frac{1}{2}\rho_0^2\Omega_1^U\Psi_1^U[Q_2]\tilde{Q}_1^2}\right].
\end{aligned}
$$

Evaluating the expectation using the result in (37) with $A = -\frac{1}{2}\rho_0^2\Omega_1^U\Psi_1^U[Q_2]$ and $b = \rho_0\Psi_2^U[Q_3]E_0[Q_3]\rho_1^U - \frac{1}{\phi}R^2 q_0 - E_0[Q_2]\rho_0\Omega_1^U\Psi_1^U[Q_2]$ and the expectation is of an exponential quadratic form in $\tilde{Q}_1 \sim \mathcal{N}(0, \Psi_0[Q_1]^{-1})$:

$$
E\left[e^{A\tilde{Q}_1^2 + b\tilde{Q}_1}\right] = \Omega_0^{-\frac{1}{2}} e^{\frac{1}{2}\Lambda_0\left(\rho_0\Psi_2^U[Q_3]E_0[Q_3]\rho_1^U - \frac{1}{\phi}R^2 q_0 - E_0[Q_2]\rho_0\Omega_1^U\Psi_1^U[Q_2]\right)^2},
$$

where $\Omega_0 \equiv 1 + \frac{\rho_0^2\Omega_1^U\Psi_1^U[Q_2]}{\Psi_0[Q_1]}$ and $\Lambda_0 \equiv \Omega_0^{-1}\Psi_0[Q_1]^{-1}$. The f.o.c for q_0 give the optimal demand (23). Equating $q_0 = X$ and simplifying yields the equilibrium price $P_0 = \frac{\mu_0}{R^3} - \frac{X}{R^3\phi\Psi_0[u]}$.

A.2 Proof of Proposition 2 (Uniqueness of Equilibrium)

By substituting the endogenous expressions for $\Psi_2^I[Q_3]$, $\Psi_1^I[Q_2]$ and G_{12} into (19), we get the following condition for the equilibrium β_1 as a function of the primitives of the model:

$$
\beta_1 = -\frac{I_1\phi\left(\eta + \eta^2 I_2\xi\phi^2\right)\left(\eta + \tau_0 + (\beta_1)^2\xi + \eta^2 I_2^2\xi\phi^2\right)}{\eta + \tau_0 + (\beta_1)^2\xi + \eta^2 I_2\xi\phi^2}. \tag{40}
$$

Therefore, β_1 is the root of a cubic polynomial:

$$
0 = \xi(\beta_1)^3 + \xi\zeta(\beta_1)^2 + \gamma\beta_1 + \omega\zeta,
$$

where $\zeta = \phi I_1\left(\eta + \eta^2 I_2\xi\phi^2\right)$, $\gamma = \eta^2 I_2\xi\phi^2 + \eta + \tau_0$ and $\omega = \eta^2 I_2^2\xi\phi^2 + \eta + \tau_0$. At least one real solution must exist by the Intermediate Value Theorem. The condition for uniqueness is that the discriminant is strictly negative, which implies

$$
4\left(\gamma^3 + \xi^2\zeta^4\omega\right) > \xi\zeta^2\left(\gamma^2 + 18\gamma\omega - 27\omega^2\right).
$$

See Birkhoff and Mac Lane (1997) for an efficient method for finding real roots of cubic polynomials.

A.3 Proof of Proposition 3 (Value of Information)

The indifference condition (20) implicitly determines I_1 as a function of the parameters of the model. By plugging in the value functions of both types of agents, recalling that $E_1^I[Q_2] = E_1^U[Q_2] + b_2\frac{\tau_1}{\tau_2}\Delta_1$ and that $E_1^I[Q_3] = E_1^U[Q_3] + \left(1 - Rb_2\frac{\tau_1}{\tau_2}\right)\Delta_1$, we get

$$
V_1^U(W_1) = -\left(\Omega_1^I\right)^{-\frac{1}{2}} e^{-\frac{1}{\phi}R^2(W_1 - v)} e^{-\frac{1}{2}E_1^U[Q_2]^2\Psi_1^I[Q_2] - \frac{1}{2}\Psi_2^I[Q_3]\left[E_1^U[Q_2]\rho_1^I - E_1^U[Q_3]\right]^2} E_1^U\left[e^{b\Delta_1 + A\Delta_1^2}\right],
$$

where $b = \Psi_2^I[Q_3]\left[E_1^U[Q_2]\rho_1^I - E_1^U[Q_3]\right] - E_1^U[Q_2]b_2\frac{\tau_1}{\tau_2}\Psi_1^I[Q_2]$ and $A = -\frac{1}{2}\left(b_2^2\frac{\tau_1^2}{\tau_2^2}\Psi_1^I[Q_2] + \Psi_2^I[Q_3]\right)$ and $\Delta_1 \sim \mathcal{N}\left(0, \tau_1^{-1}\right)$. Evaluating the expectation using the result in (37) we get

$$
E_1^U\left[e^{b\Delta_1 + A\Delta_1^2}\right] = \Omega_v^{-\frac{1}{2}} e^{\frac{1}{2}\left(\Psi_2^I[Q_3]\left[E_1^U[Q_2]\rho_1^I - E_1^U[Q_3]\right] - E_1^U[Q_2]b_2\frac{\tau_1}{\tau_2}\Psi_1^I[Q_2]\right)^2\Omega_v^{-1}\tau_1^{-1}},
$$

where $\Omega_v = \frac{\tau_2+\eta}{\tau_1}$. Plugging in the value functions and rearranging we get

$$
e^{\frac{1}{\phi}R^2 v} = \Omega_v^{\frac{1}{2}}\left[\Gamma(I_1)\sqrt{\frac{\tau_2}{\tau_{2+\eta}}} + (1-\Gamma(I_1))\right]\left(\frac{\Omega_1^U}{\Omega_1^I}\right)^{-\frac{1}{2}} e^{-\frac{1}{2}E_1^U[Q_2]^2 B_1 + E_1^U[Q_3]E_1^U[Q_2]B_2 - \frac{1}{2}E_1^U[Q_3]^2 B_3}, \tag{41}
$$

where B_1, B_2, and B_3 are all zero. Therefore, the exponential term on the r.h.s of (41) vanishes. Taking logs and rearranging we get the desired result (21).

A.4 Proof of Proposition 4 (Price Reaction to News)

Expected returns conditional on news Δ_0 can be expressed as a constant multiplying the news:

$$E_0\left[Q_t|\Delta_0\right] - E_0\left[Q_t\right] = G_t\Delta_0,$$

where $G_1 = b_1\frac{\tau_0}{\tau_1}$, $G_2 = \left[b_1\rho_0 + b_2\frac{\tau_1}{\tau_2}\right]\frac{\tau_0}{\tau_1}$ and $G_3 = \left[1 - Rb_2\frac{\tau_1}{\tau_2}\right]\frac{\tau_0}{\tau_1}$. Suppose we hold I_1 and τ_1 fixed and consider a marginal change in δ. From the GIS process, $\frac{\partial I_2}{\partial\delta} = \frac{\partial\Gamma(I_1)}{\partial\delta}U_1$; therefore, assuming that the incidence probability Γ is increasing in the transmission rate we get $\frac{\partial I_2}{\partial\delta} > 0$. The following results will be establishes in terms of marginal changes in I_2, but it should be kept in mind that these directly sign the derivatives with respect to δ. The price coefficient on Δ_1 is b_1. Differentiating (39) w.r.t I_2 and recalling that $\frac{\partial\beta_1}{\partial I_2} \leq 0$ and $\beta_1 \leq 0$ gives

$$\frac{\partial b_1}{\partial I_2} = \frac{\partial b_1}{\partial\beta_1}\frac{\partial\beta_1}{\partial I_2} = -\frac{\frac{\partial\beta_1}{\partial I_2}}{\phi\left(I_1\Omega_1^I\Psi_1^I[Q_2] + U_1\Omega_1^U\Psi_1^U[Q_2]\right)} + \frac{2\xi\beta_1\frac{\partial\beta_1}{\partial I_2}}{R^2\tau_0} \geq 0,$$

where the inequality is strict when $I_1 > 0$. A similar argument applies for b_2. Differentiating (35) w.r.t I_2 and recalling that $\frac{\partial\beta_2}{\partial I_2} \leq 0$ and $\beta_2 \leq 0$ gives

$$\frac{\partial b_2}{\partial I_2} = \frac{\partial b_2}{\partial\beta_2}\frac{\partial\beta_2}{\partial I_2} = \frac{1}{R}\left[\frac{\Psi_2^U[Q_3]\Psi_2^I[Q_3]}{\left(I_2\Psi_2^I[Q_3] + U_2\Psi_2^U[Q_3]\right)^2}\right] + \frac{2\xi\beta_2\frac{\partial\beta_2}{\partial I_2}}{R\tau_1} > 0.$$

With these results in mind, consider the cash-flow news sensitivities:

$$\begin{aligned}
\frac{\partial G_1}{\partial I_2} &= \frac{\partial}{\partial I_2}\left[b_1\frac{\tau_0}{\tau_1}\right] = \frac{\tau_0}{\tau_1}\frac{\partial b_1}{\partial I_2} \geq 0\\
\frac{\partial G_2}{\partial I_2} &= \frac{\partial}{\partial I_2}\left[\left(b_1\rho_0 + b_2\frac{\tau_1}{\tau_2}\right)\frac{\tau_0}{\tau_1}\right] = \frac{\tau_0}{\tau_1}\left(\frac{\partial b_1}{\partial I_2}\rho_0 + \frac{\partial b_2}{\partial I_2}\frac{\tau_1}{\tau_2}\right) \lesseqgtr 0\\
\frac{\partial G_3}{\partial I_2} &= \frac{\partial}{\partial I_2}\left[\left(1 - Rb_2\frac{\tau_1}{\tau_2}\right)\frac{\tau_0}{\tau_1}\right] = -\frac{\tau_0}{\tau_1}R\frac{\partial b_2}{\partial I_2}\frac{\tau_1}{\tau_2} \leq 0.
\end{aligned}$$

Since in information market equilibrium, $\frac{\tau_1}{\tau_2}$ is a constant that depends on the parameters of the model and in particular on the cost of information c, G_2 can either increase or decrease from a marginal change in I_2. The reason is that the positive derivatives $\frac{\partial b_1}{\partial I_2}$ and $\frac{\partial b_2}{\partial I_2}$ enter with opposite signs since $\rho_0 < 0$.

B Figures and Tables

Figure 1: Informational Types over Time

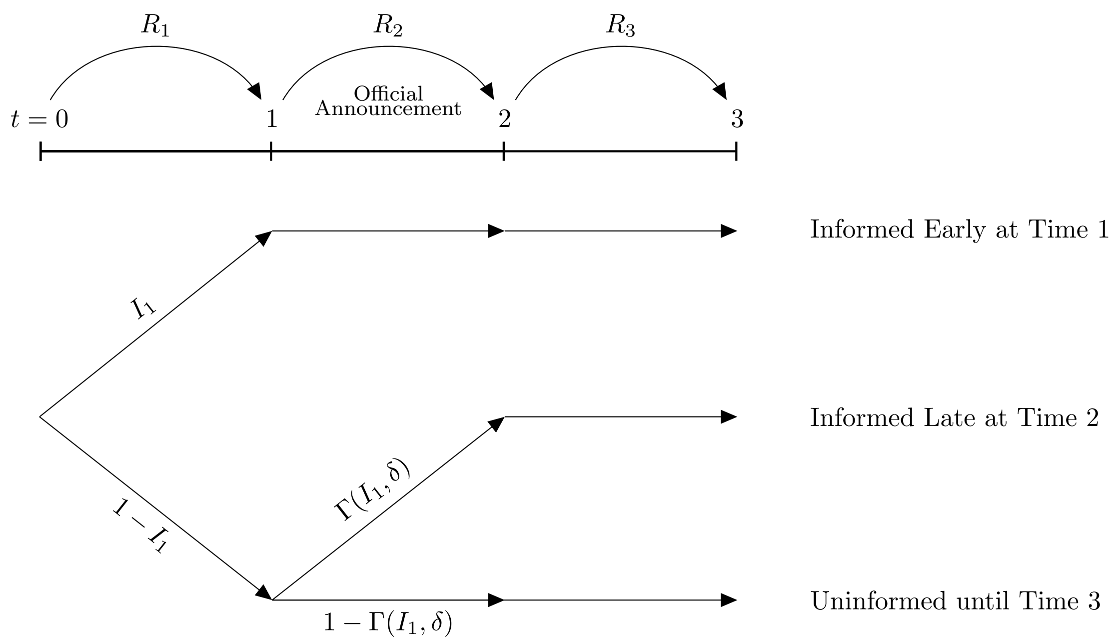
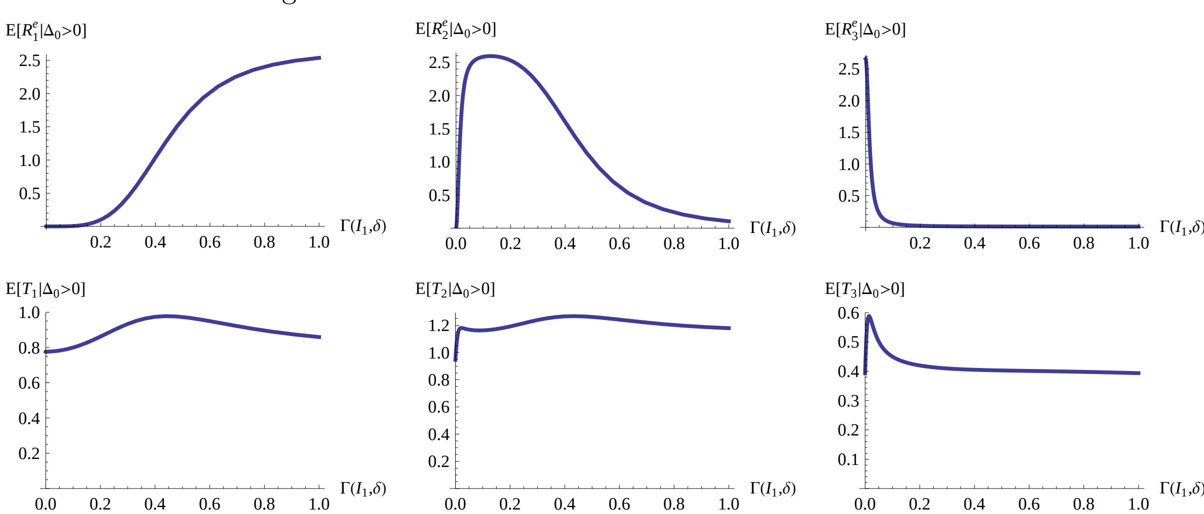
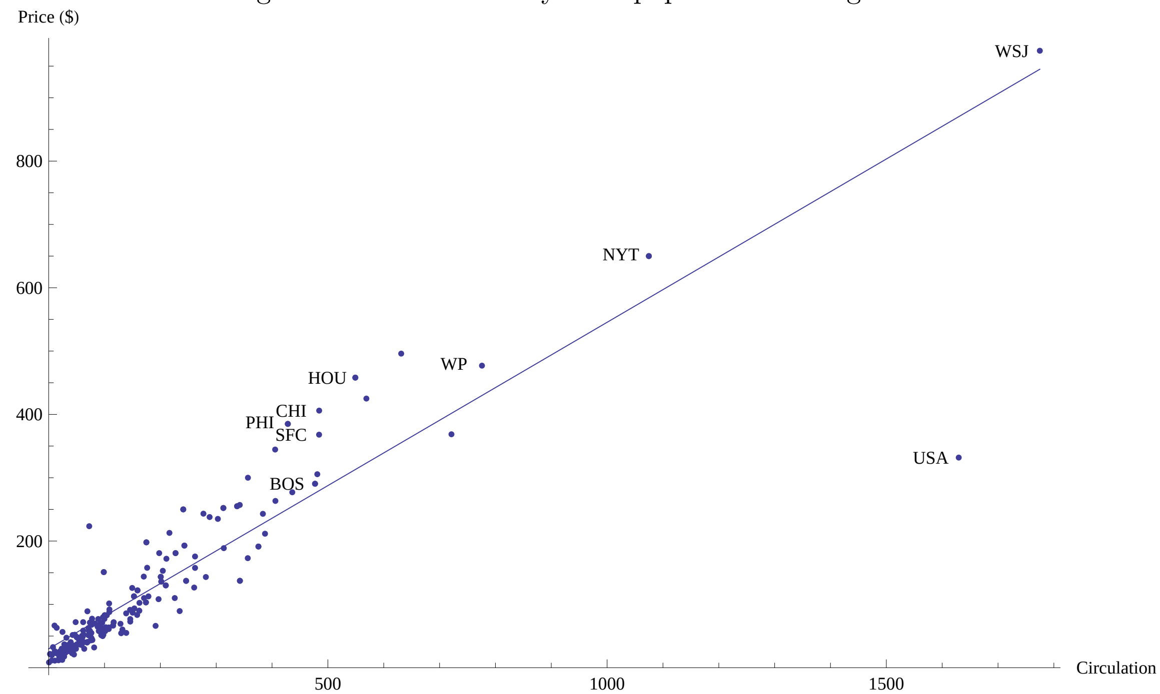


Figure 2: Returns and Turnover Around Positive News



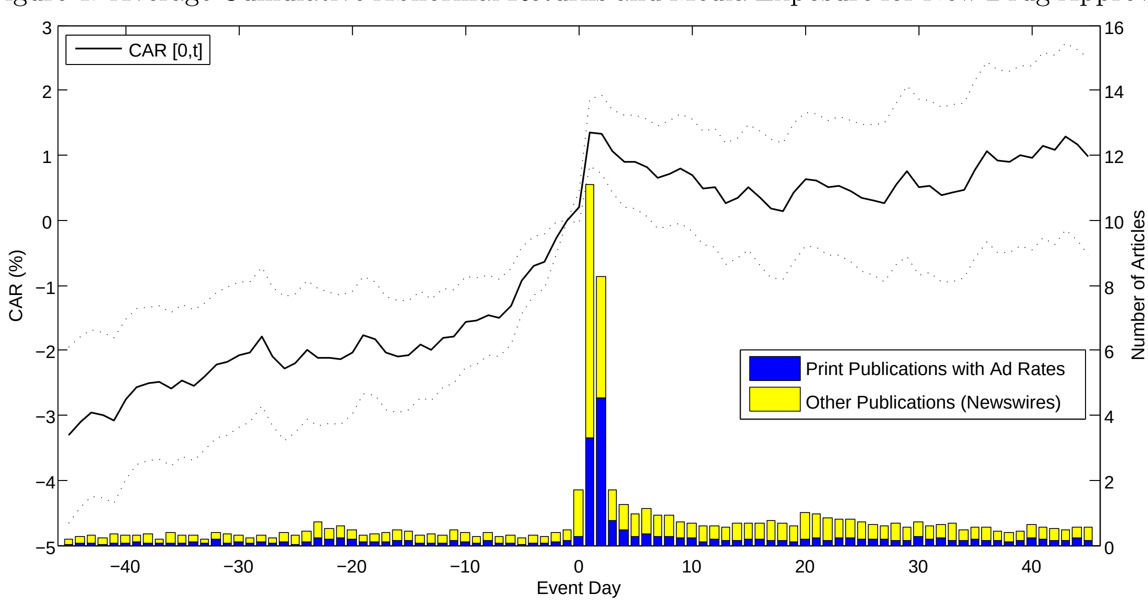
Plotted are percent expected returns in excess of the unconditional risk premium R_t^e and turnover T_t as a function of the incidence probability $\Gamma(I_1, \delta)$ at times 1, 2 and 3 of the model corresponding to the pre-announcement, announcement and post-announcement periods. Parameter values are those estimated in section 4. Expected values are calculated using 10,000 simulated draws of the shocks. θ shocks are drawn from a truncated normal distribution as in (31) to condition on positive news.

Figure 3: 1998 U.S. Daily Newspaper Advertising Rates



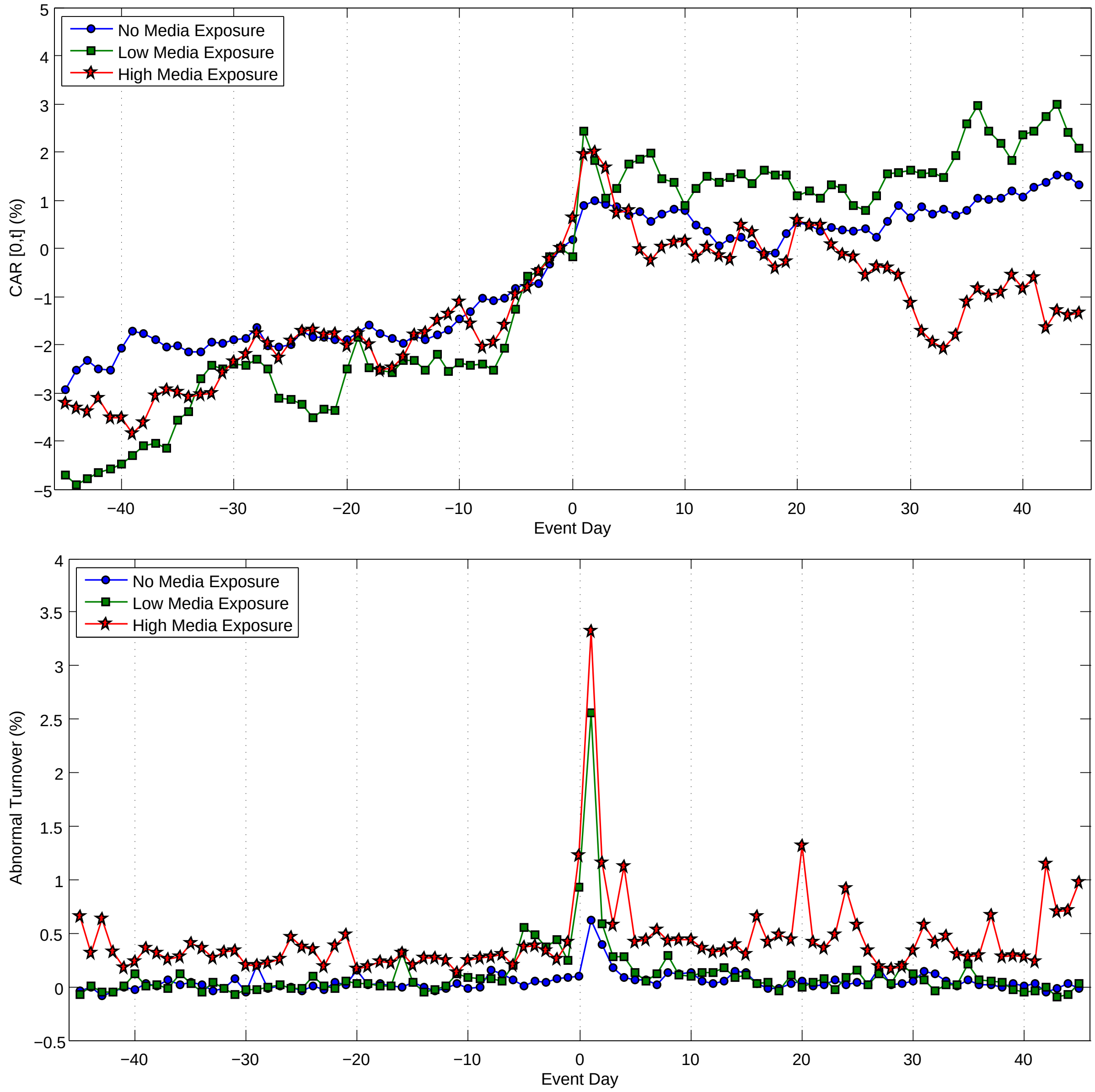
Plotted is a regression of the 1998 dollar price per column-inch of advertising on circulation measured in thousands of readers for U.S. daily newspapers. Labeled are the *Philadelphia Inquirer, Boston Globe, Chicago Sun-Times, San Francisco Chronicle, Houston Chronicle, Washington Post, New York Times, USA Today,* and *Wall Street Journal*.

Figure 4: Average Cumulative Abnormal Returns and Media Exposure for New Drug Approvals



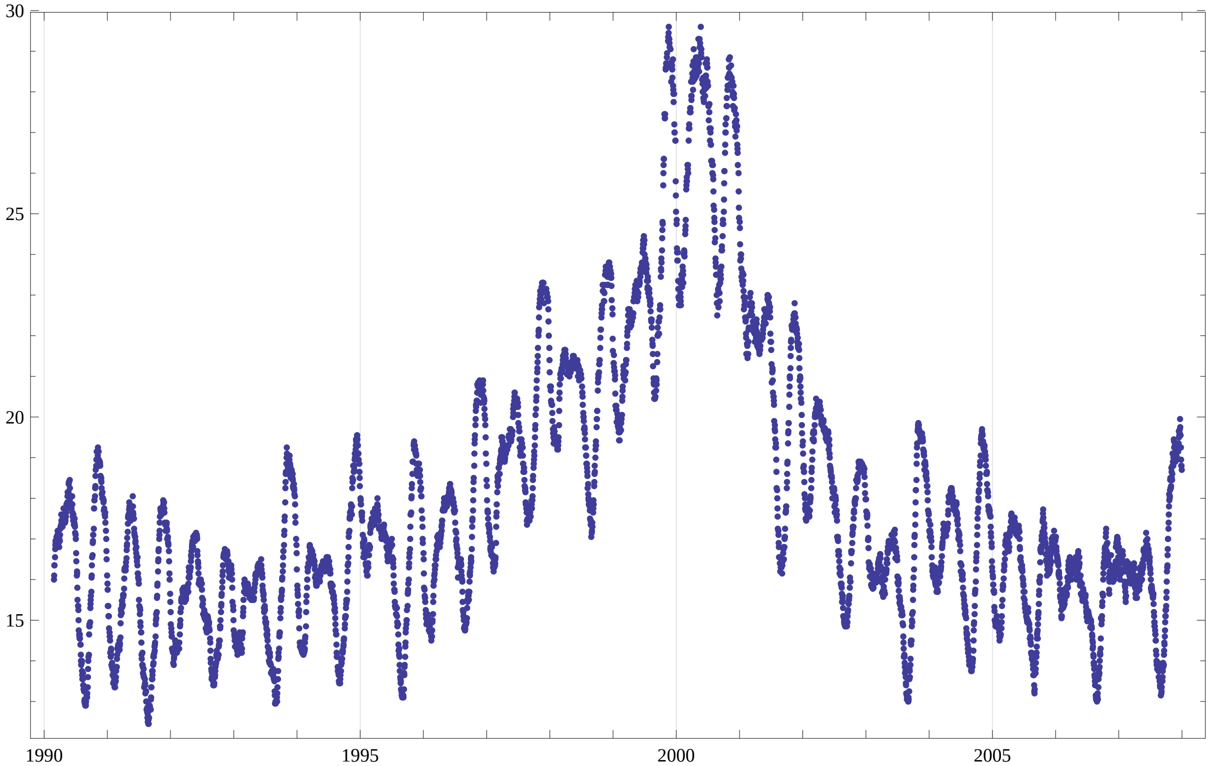
The stacked bars are the average number of articles covering a new drug approval by the FDA. The dark bars are for articles for which advertising rates are available. Light bars for other publications mostly consist of news services such as newswires and newsletters. The solid line represents CAR over days [0,t]. The dotted lines are the 90% confidence interval.

Figure 5: Media Exposure Sub-Samples



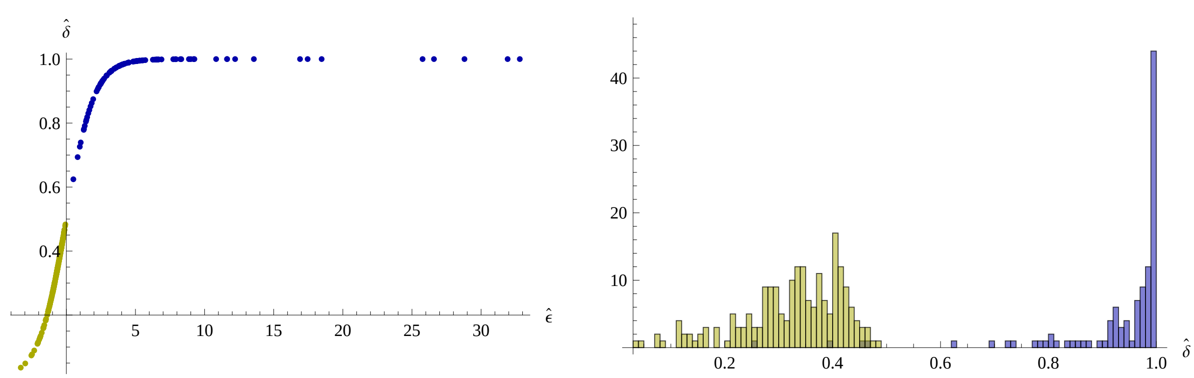
High media exposure sub-sample is the top half of drug approvals with positive media exposure on days 0 and 1 containing 56 observations. Low contains the bottom 57 observations with positive media exposure. No media exposure contains the remaining 207 observations. The plotted average cumulative abnormal returns are normalized to zero one day before the approval. Abnormal returns and turnover are in excess of those of the value-weighted market portfolio.

Figure 6: WSJ News Pressure 1990-2007



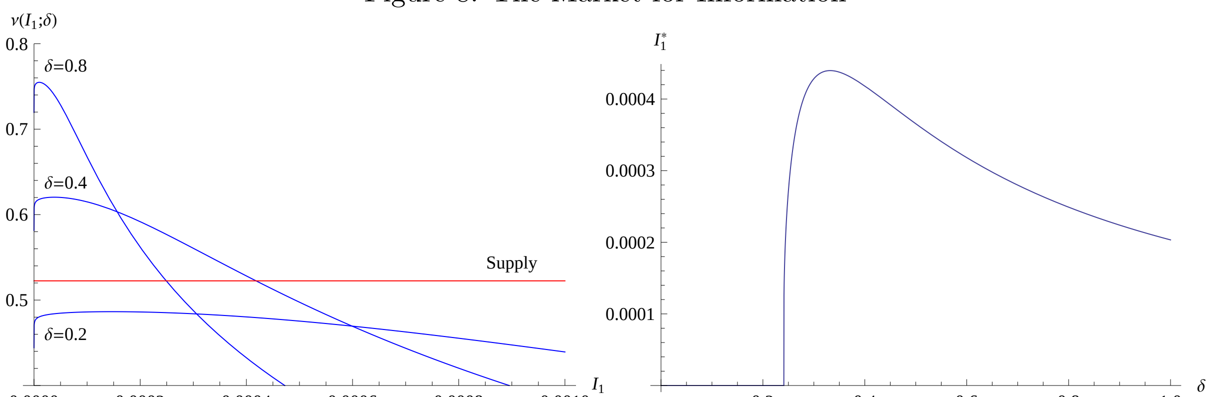
WSJ News Pressure is the backward-looking 40-day moving average of the number of pages in section A of the *Wall Street Journal*.

Figure 7: Estimated Transmission Rates



The plot on the left shows the logistic mapping from latent residuals to transmission rates of the censored media coverage model as specified in (30). The histogram on the right describes the distribution of latent transmission rates. In both plots, dark colors represent uncensored ($m_j > 0$), and light colors represent censored ($m_j = 0$) observations.

Figure 8: The Market for Information



The plot on the left shows demand for information $v(I_1;\delta)$ as a function of the percent of population informed I_1 for three different values of δ. The horizontal line indicates the constant supply price. The plot on the right shows the equilibrium informed percent of population. The parameters are those estimated in section 4.

Table 1: New Drug Approvals Sample Summary Statistics

	Mean	Std	Min	25^{th} %	Median	75^{th} %	Max	N
Media Exposure	1.50	4.31	0	0	0	0.78	31.96	320
Media Exposure (Count)	3.58	9.22	0	0	0	2.00	70.00	320
Subsequent Media Exposure	2.17	3.39	0	0	1	2.63	20.92	320
Preceding Media Exposure	0.14	0.58	0	0	0	0	5.18	320
WSJ News Pressure	18.57	3.44	13.25	16.29	17.60	20.38	29.20	320
Market Cap (Millions $)	28064	34622	18	1774	16401	39211	194823	320
Number of Analysts	19.62	13.86	0.00	7.00	19.50	31.00	48.00	320
Analysts Estimates Dispersion	0.08	0.20	0.00	0.01	0.02	0.06	2.41	297
Priority Drug	0.45	0.50	0	0	0	1	1	320
Orphan Drug	0.19	0.39	0	0	0	0	1	320
Cancer Drug	0.16	0.37	0	0	0	0	1	170
HIV/AIDS Drug	0.06	0.24	0	0	0	0	1	170
Friday Approval	0.32	0.47	0	0	0	1	1	320
Idiosyncratic Risk	2.15	1.33	0.70	1.31	1.69	2.51	9.32	309
Patent Months Remaining	192.21	47.90	34.3	166.6	187.4	223.8	343.3	235
Exclusivity Months Remaining	57.45	56.97	0.0	0.0	60.0	90.0	213.4	235

The sample includes 320 Original New Drug Approvals over the years 1990-2007 marked by the FDA as New Molecular Entity applications. **Media Exposure** is the sum of all articles on the approval day and the following day, weighted by an adjacent advertising rate and presented in thousands of dollars. **Subsequent Media Exposure** is calculated similarly for days 2 to 6, as is **Preceding Media Exposure** for days -5 to -1. **Media Exposure (Count)** is the number of articles. **WSJ News Pressure** is the 40-day moving average of the number of pages in section A of the *Wall Street Journal* one day after the approval. **Market Cap** is the sponsoring firm's market capitalization in millions of 1990 dollars one year before the event. If data are not available for that time then the first day with data within that year is used instead. **CAR** [a,b] is cumulative abnormal percent return over event trading days a to b, where abnormal return is return in excess of the value-weighted market portfolio. **CATO** [a,b] is cumulative abnormal percent turnover, where abnormal turnover is turnover in excess of market portfolio turnover. **Pre-Approval** window is [-5,-1]. **Approval** window is [0,1]. **Post-Approval** window is [2,6]. **Historical Turnover** is the security's past year average daily percent turnover up to day t-12. **Number of Analysts** and **Analyst Estimates Dispersion** are from I/B/E/S unadjusted summary file for one year earnings estimates valid one month before the approval. Dispersion, defined as standard deviation over absolute value of the mean, is undefined when the mean is zero or when only one analyst is covering the firm. **Priority Drug** and **Orphan Drug** are dummy variables set according to the drug's review classification (both can be true). **Idiosyncratic Risk** is the standard deviation of the residuals from a market model regression of past year daily stock returns on the value-weighted market portfolio. **Patent** and **Exclusivity Months Remaining** are calculated as the difference between their expiry date as it appears in the FDA's Electronic Orange Book files and the approval date.

Table 2: Returns and Volume around Approvals

	Mean	t-statistic	Std	Min	Median	Max	Per-day	N
Pre-Approval CAR (R_1)	1.33	3.75	6.34	-18.07	0.63	39.13	0.27	320
Approval CAR (R_2)	1.35	4.32	5.59	-18.61	0.60	47.08	0.68	320
Post-Approval CAR (R_3)	-0.53	-1.62	5.86	-30.00	-0.14	22.72	-0.11	320
Historical Turnover	0.52	14.12	0.58	0.04	0.31	5.43	0.52	248
Pre-Approval CATO (T_1)	0.56	2.36	3.79	-3.79	-0.56	28.28	0.11	255
Approval CATO (T_2)	2.02	4.92	6.55	-1.41	-0.13	54.82	0.40	255
Post-Approval CATO (T_3)	1.29	3.53	5.84	-3.91	-0.48	52.27	0.26	255

CAR [a,b] is cumulative abnormal percent return over event trading days a to b, where abnormal return is return in excess of the value-weighted market portfolio. **CATO** [a,b] is cumulative abnormal percent turnover, where abnormal turnover is turnover in excess of market portfolio turnover. Turnover data is not available for ADRs. **Pre-Approval** window is [-5,-1]. **Approval** window is [0,1]. **Post-Approval** window is [2,6]. **Historical Turnover** is the security's past year average daily percent turnover up to day -12. **Per-day** is the mean divided by the number of cumulation days.

Table 3: Effects of Transmission Rates on Returns

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent Variable:	m	R_1	R_2	R_3	R_1	R_2	R_3
Transmission Rate $\hat{\delta}$		0.14	-0.03	-0.23***			
		[0.12]	[0.13]	[0.06]			
Media Exposure $\hat{m}$					-0.02	0.88*	0.20
					[0.63]	[0.48]	[0.71]
WSJ News Pressure	-0.54**						
	[0.21]						
Firm Size	0.09	-0.73***	-0.51**	0.44**	-0.73***	-0.59***	0.42*
	[0.11]	[0.23]	[0.20]	[0.20]	[0.24]	[0.23]	[0.23]
R-squared	0.10	0.14	0.10	0.13	0.14	0.11	0.10
N	320						

Media Exposure m is the sum of all articles on the approval day and the following day, weighted by an adjacent advertising rate and presented in thousands of dollars. Cumulative abnormal returns R_1, R_2 and R_3 are measured over the pre-approval [-5,-1], approval [0,1], and post-approval [2,6] windows respectively. WSJ News Pressure is the 40-day moving average of the number of pages in section A of the *Wall Street Journal* one day after the approval. Firm Size is the sponsoring firm's log market capitalization in millions of 1990 dollars one year before the event. The predicted value of Media Exposure $\hat{m}$ and the Transmission Rate $\hat{\delta}$ are estimated from regression (1). All regressions include an intercept and year and month fixed-effects. Standard errors clustered by approval day and adjusted using GMM for estimation error in $\hat{\delta}$ are in brackets. *, **, and *** indicate significance at 10, 5, and 1 percent respectively.

Table 4: Alternative Proxies for Transmission Rates

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable:	R_1		R_2		R_3	
Priority Drug	-0.32	-0.70	0.97	1.11*	-1.90***	-1.41**
	[0.77]	[0.82]	[0.62]	[0.64]	[0.70]	[0.67]
Orphan Drug	-0.31	-0.35	0.93	0.95	3.05***	3.11***
	[0.98]	[0.97]	[1.05]	[1.05]	[1.14]	[1.13]
Media Exposure		0.16		-0.06		-0.21***
		[0.13]		[0.14]		[0.06]
R-squared	0.14	0.15	0.11	0.11	0.14	0.16
N	320					

Cumulative abnormal returns R_1, R_2 and R_3 are measured over the pre-approval [-5,-1], approval [0,1], and post-approval [2,6] windows respectively. The FDA grants Priority Drug review status to drugs that promise a significant advance over existing treatments, and Orphan Drug status if the treatment is for a disease affecting fewer than 200,000 Americans. Both indicators can be 1. Media Exposure is the sum of all articles on the approval day and the following day, weighted by an adjacent advertising rate and presented in thousands of dollars. Firm Size is the sponsoring firm's log market capitalization in millions of 1990 dollars one year before the event. All regressions include an intercept, firm size, and year and month fixed-effects. Standard errors clustered by approval day are in brackets. *, **, and *** indicate significance at 10, 5, and 1 percent respectively.

Table 5: Effects of Transmission Rates on Turnover

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable:	T_1	T_2	T_3	T_1	T_2	T_3
Transmission Rate $\hat{\delta}$	0.12***	0.23***	0.19**			
	[0.04]	[0.06]	[0.08]			
Media Exposure $\hat{m}$				-0.41	-0.22	-0.46
				[0.33]	[0.41]	[0.55]
Firm Size	-0.70***	-1.08***	-0.96***	-0.67***	-1.07***	-0.93***
	[0.12]	[0.17]	[0.16]	[0.12]	[0.17]	[0.16]
R-squared	0.33	0.42	0.31	0.32	0.40	0.29
N	255					

Cumulative abnormal turnover T_1, T_2 and T_3 is measured over the pre-approval [-5,-1], approval [0,1], and post-approval [2,6] windows respectively. WSJ News Pressure is the 40-day moving average of the number of pages in section A of the *Wall Street Journal* one day after the approval. Firm Size is the sponsoring firm's log market capitalization in millions of 1990 dollars one year before the event. The predicted value of Media Exposure $\hat{m}$ and the Transmission Rate $\hat{\delta}$ are estimated from regression (1) in Table (3). All regressions include an intercept and year and month fixed-effects. Standard errors clustered by approval day and adjusted using GMM for estimation error in $\hat{\delta}$ are in brackets. *, **, and *** indicate significance at 10, 5, and 1 percent respectively.

Table 6: Censored Media Coverage Model ML Estimates

$$m_j = \max\left\{0, \boldsymbol{x}_j^{\mathbf{T}}\boldsymbol{\beta}_m + \gamma z_j + logit\left(\delta_j\right)\right\}$$
$$\delta \sim Beta(\alpha_\delta, \beta_\delta)$$

γ	α_δ	β_δ	N
-0.71**	0.30	0.25***	320
[0.36]	[0.24]	[0.02]	

Reported are maximum likelihood estimates from the censored media coverage model. Controls in $\boldsymbol{x}_j$ include a constant, firm size, and year and month fixed-effects. z_j is the WSJ News Pressure one day after the approval, which proxies for the availability of other newsworthy material. Standard errors are in brackets.

Table 7: Indirect Estimation Results

ϕ^{-1}	$E[\theta]$	$\sigma[x_t]$	$\sigma[\theta]$	c	n	Objective
0.5	1.008	0.019	0.028	4.080	0.163	4.129
		[0.001]	[0.000]	[1.188]	[0.055]	
1.0	1.013	0.013	0.021	0.566	0.099	4.165
		[0.023]	[0.001]	[23.220]	[3.494]	
2.5	1.028	0.014	0.019	0.201	0.109	4.269
		[0.010]	[0.000]	[3.850]	[2.133]	
5.0	1.058	0.010	0.034	0.522	0.014	2.892
		[0.001]	[0.008]	[0.118]	[0.007]	
10.0	1.122	0.015	0.044	0.150	0.057	3.451
		[0.001]	[0.008]	[0.010]	[0.024]	

Reported are indirect inference parameter estimates each using five different values of absolute risk tolerance ϕ. Each optimization uses the 255 observations times 300 simulated shock draws. The mean payoff $E[\theta] = \mu_0$ is constrained given the other parameters so that $P_0 = 1$. The standard deviation of the the unobservable component is normalized to $\sigma[\epsilon] = 0.1$. $\sigma[x_t]$ is the standard deviation of the supply shocks. $\sigma[\theta]$ is the standard deviation of the signal θ at time 0. c is the fixed cost of information. n is the elasticity of the incidence probability $\Gamma(I_1; \delta)$ w.r.t to the fraction informed at time 1. Objective is the value of the minimized quadratic objective. Standard errors are in brackets.

Table 8: Moments Comparison ($\phi^{-1} = 5.0$)

	Sample	Model	Error
$E[R_1^e]$	1.669	1.641	0.028
	[0.424]	[0.127]	[0.442]
$E[R_2^e]$	1.469	0.766	0.702
	[0.383]	[0.094]	[0.394]
$E[R_3^e]$	-0.635	0.227	-0.861
	[0.395]	[0.640]	[0.752]
$E[T_1]$	0.556	0.922	-0.366
	[0.237]	[0.034]	[0.240]
$E[T_2]$	2.017	1.219	0.798
	[0.410]	[0.038]	[0.412]
$E[R_1^e\delta]$	0.989	1.108	-0.119
	[0.283]	[0.084]	[0.295]
$E[R_2^e\delta]$	1.042	0.294	0.748
	[0.306]	[0.041]	[0.309]
$E[R_3^e\delta]$	-0.550	0.008	-0.557
	[0.269]	[0.392]	[0.475]
$Cov[R_1^e,\delta]$	0.091	0.225	-0.134
$Cov[R_2^e,\delta]$	0.253	-0.118	0.371
$Cov[R_3^e,\delta]$	-0.208	-0.114	-0.094

Reported are sample moments and their simulated counterparts from the model. The numerical procedure's goal is to minimize the discrepancy between the two, which I report in the last column. The three covariances are calculated based on the matched moments in the top part of the table. Standard errors are in brackets.

Table 9: Regressions of Post-Approval Returns

$$R_3^e = \beta_0 + \beta_1 R_1^e + \beta_2 R_2^e + \beta_3 T_3 + \pi_3 \delta + \upsilon$$

	Sample	Model
1	0.630	2.200
	[0.791]	[0.433]
R_1^e	-0.129	0.118
	[0.058]	[0.119]
R_2^e	0.021	-0.277
	[0.065]	[0.205]
T_3	-0.101	0.994
	[0.068]	[0.688]
δ	-1.766	-3.373
	[1.307]	[0.599]
R^2	0.042	0.016
Obs	255	2550

The dependent variable is post-approval cumulative abnormal percent return over days 2 to 6 for the sample regression. In the regressions simulated using the model, it corresponds to the per-share excess percent return. T_t in the model corresponds to cumulative abnormal turnover over the same window in the sample. δ in the model is drawn randomly from $Beta(\hat{\alpha}_{\delta}^{MLE}, \hat{\beta}_{\delta}^{MLE})$, whereas in the sample, it is $\hat{\delta}$ estimated from (30) using MLE . Standard errors are in brackets.

Table 10: Regressions of Post-Approval Turnover

$$T_3 = \beta_0 + \beta_1 T_1 + \beta_2 T_2 + \pi_3 \delta + \upsilon$$

	Sample	Model
1	0.070	0.098
	[0.480]	[0.005]
T_1	0.437	-0.199
	[0.090]	[0.004]
T_2	0.478	0.460
	[0.053]	[0.004]
δ	0.023	-0.089
	[0.802]	[0.005]
R^2	0.586	0.861
Obs	255	2550

The dependent variable is post-approval cumulative abnormal percent turnover over days 2 to 6 for the sample regression. In the regressions simulated using the model it corresponds to percent trade volume T_3. δ in the model is drawn randomly from $Beta(\hat{\alpha}_\delta^{MLE}, \hat{\beta}_\delta^{MLE})$, whereas in the sample, it is $\hat{\delta}$ estimated from (30) using MLE . Standard errors are in brackets.