A COMPARISON OF THE PIN AND APIN MODELS USING NYSE DATA

AUTHOR: YUE WU
SUPERVISOR: DANIEL PREVE

Department of statistics
Uppsala University
A COMPARISON OF THE PIN AND APIN MODELS USING NYSE DATA

YUE WU

June 6, 2012

Abstract

This master thesis reviews the probability of informed trading (PIN) model proposed by Easley et al. (2002) and its recent extension, the adjusted PIN (APIN) model proposed by Duarte and Young (2009). The models are then applied to high frequency data from the New York Stock Exchange (NYSE). Our empirical results indicate that the APIN model with symmetric order-flow shock provides a better fit than the PIN model. The less flexible PIN model appears to overestimate the asymmetric information.¹

¹Key words: probability of informed trading, probability of symmetric order-flow shock, market microstructure, transaction data.
1 Introduction

The information risk or information asymmetry problem is one of the main problems in market microstructure research. When there is asymmetric information in the market, which is part of traders relative to other traders who have private information of the true value of the assets, this part of the informed traders will make a profit using of the private information of informed trading. Information risk which means investors may suffer a loss due to asymmetric information on certain assets is a measure of the degree of information asymmetry. It often uses to measure by probability of informed trading. Accurate measure of the risk of stock information is great significance both for asset pricing, risk management, or the measure of market performance.

As informed trading can not be observed directly from the market, the early literature from the indirect perspective often use some substitute variables to measure information risk or information asymmetry in the market. Such as Bagehot (1971), Jaffe and Winkler (1976), they proposed the use of the bid-ask spread as a simple measure of asymmetric information. But these methods do not explicitly measure the information asymmetry faced by the traders in the market, and these results are not standardized. People can not directly compare the information risk on different markets or different levels of asymmetric information.


Recently, Duarte & Young (2009) included an order-flow shock component in the PIN model and proposed the adjusted PIN (APIN) model. On a normal trading day, the trader will face either bad news, good news or no news. We use the PIN model to estimate the probability of informed trading and the probability of good-, bad- and no-news. In the APIN model, we will also consider the probability of symmetric order-flow shock (PSOS). In this thesis, we compute empirical PIN, APIN and PSOS using high-frequency transaction data. We use maximum likelihood estimation to estimate model parameters and Akaike’s information criterion (AIC) to choose the model with the best fit.

The general structure of this thesis is as follows. In Section 2, we briefly introduce the PIN and APIN models. In Section 3, firstly, we describe the characteristics of the data. Secondly, we discuss our empirical results. Section 4 concludes the paper.
2 Methods

2.1 The classical PIN model

With the classical PIN model of asymmetric information by Easley et al. (1997b), the authors introduce the probability of informed trading measure. Since then, the PIN model has been a popular model of information asymmetry and it has been widely applied in market microstructure research. We here briefly introduce the classical PIN model.

2.1.1 Assumptions of the PIN model

In the PIN model, Easley et al. (1997b) assumes that the investors on the market is divided into informed traders and uninformed traders. Investors and market makers transact with risk assets and cashes. In each trading day, it is assumed that the probability of news is \( \theta_E \), hence, the probability of no news is \( 1 - \theta_E \). Uninformed traders will trade whether news has been released or not. In each trading day, the aggregate number of buy or sell orders are assumed to be Poisson distribution with expectations \( \lambda_1 \) and \( \lambda_{-1} \). Informed traders will only trade if news is released, these days the buy and sell intensity increases by a constant, \( \delta \). When news is released, there are two different situations: bad news and good news. Assume that the probability of bad news is \( \theta_B \), so the probability of good news is \( 1 - \theta_B \). When there is bad news, informed traders will choose to sell. In contrast, when there is good news, informed traders will choose to buy. The trading tree for a day is outlined in Figure 1.

In Figure 1, news arrive with probability \( \theta_E \). Given news, the news are either good or bad. If the news is good, the buy intensity is \( \lambda_1 + \delta \) and the sell intensity is \( \lambda_{-1} \). Similarly, if the news is bad, the sell intensity increases by \( \delta \) and the buy intensity is \( \lambda_1 \) (the same as on a no news day).

2.1.2 PIN model measure of information asymmetry

Probability of informed trading, PIN, is defined as

\[
PIN = \frac{\theta_E \delta}{\lambda_1 + \lambda_{-1} + \theta_E \delta},
\]

and can be interpreted as the relative intensity of informed trades to all trades. As can be seen from Equation (1), PIN depends on the arrival rate of informed and uninformed traders and on the probability of information being released.

2.1.3 PIN model parameter estimation

We only need to know the total number of buy and sell orders within a trading day when we estimate parameters. If we do not know what type of events in the transaction, we can calculate the likelihood function according to the weighted probability of occurrence of each part. In other words, since the probability of
no news, the conditional probability of bad and good news is \((1 - \theta_E)\), \(\theta_E\theta_B\) and \(\theta_E(1 - \theta_B)\) respectively, then we can write the daily contribution to the likelihood function as

\[
L_d = L_d(\theta_E, \theta_B, \lambda_1, \lambda_{-1}, \delta \mid B_d, S_d)
\]

\[
= (1 - \theta_E) \frac{\lambda_1 B_d}{B_d!} e^{-\lambda_1} \frac{\lambda_{-1} S_d}{S_d!} e^{-\lambda_{-1}} + \theta_E \theta_B \frac{\lambda_1 B_d}{B_d!} e^{-\lambda_1} \frac{(\lambda_{-1} + \delta) S_d}{S_d!} e^{-(\lambda_{-1} + \delta)}
\]

\[
+ \theta_E (1 - \theta_B) \frac{(\lambda_1 + \delta) B_d}{B_d!} e^{-(\lambda_1 + \delta)} \frac{\lambda_{-1} S_d}{S_d!} e^{-\lambda_{-1}},
\]

(2)

where \(B_d\) is the number of buy orders in a trading day and \(S_d\) is the number of sell orders in a trading day. As trading days are assumed to be independent, the likelihood function for \(D\) observation days is

\[
L = \prod_{d=1}^{D} L_d.
\]

(3)

We use the maximum likelihood method to estimate the parameters.

### 2.1.4 Characteristics of the PIN model

Because of the Poisson assumption,

\[
E(B_d) = \theta_E \theta_B \lambda_1 + \theta_E (1 - \theta_B)(\lambda_1 + \delta) + (1 - \theta_E) \lambda_1
\]

\[
= \lambda_1 + \theta_E (1 - \theta_B) \delta,
\]

(4)

where \(E(B_d)\) is the expectation of the daily number of buy orders. Similarly, the expectation of sell orders in PIN model is

\[
E(S_d) = \theta_E \theta_B (\lambda_{-1} + \delta) + \theta_E (1 - \theta_B) \lambda_{-1} + (1 - \theta_E) \lambda_{-1}
\]

\[
= \lambda_{-1} + \theta_E \theta_B \delta,
\]

(5)

where \(E(S_d)\) is the expectation of the daily number of sell orders. Thus, the daily expected total number of trades is \(E(B_d + S_d) = \lambda_1 + \lambda_{-1} + \theta_E \delta\). The variance of the buy orders is \(Var(B_d) = E(B_d^2) - E^2(B_d)\). By the same way that we calculated \(E(B_d)\), we can get \(E(B_d^2)\). Then
\[ \text{Var}(B_d) = \lambda_1^2 + \theta_E (1 - \theta_B) \delta^2 + 2\theta_E (1 - \theta_B) \lambda_1 \delta - (\lambda_1 + \theta_E (1 - \theta_B) \delta)^2 \]
\[ = \theta_E (1 - \theta_B) (1 - \theta_E (1 - \theta_B)) \delta^2 \]  
\[ \text{Var}(S_d) = \lambda_{-1}^2 + \theta_E \theta_B \delta^2 + 2\theta_E \theta_B \lambda_{-1} \delta - (\lambda_{-1} + \theta_E \theta_B \delta)^2 \]
\[ = \theta_E \theta_B (1 - \theta_E \theta_B) \delta^2 \]

Similarly, we can calculate the variance of the sell orders

The covariance between buy and sell orders is \( \text{Cov} = E(B_d S_d) - E(B_d)E(S_d) \), which is

\[ \text{Cov}(B_d, S_d) = -\theta_B (1 - \theta_B) (\theta_E \delta)^2 \]  

From Equation (8), we can see that the covariance between the daily number of buy and sell orders is always negative (\( \theta_B \) is between 0 and 1). This means that the correlation between buy and sell orders is always negative in the PIN model. As observed by Duarte & Young (2009), this property is not supported by actual data.

2.2 The adjusted PIN model

Recently, Duarte & Young (2009) improved the PIN model. They increased trading motivation within the entire market called order-flow shock. Because even there is no news, both buyer and seller will do some tradings to increase order flows. At least two reasons produce order-flow shock. One possibility is that public information events lead to investor divergence. This kind of public information divergence increases both buy and sell orders. As Kandel and Pearson (1995) found, if investors have different views to explain information, they will divide even observing the same information. Another possibility for order-flow shock is that traders only reduce transaction costs in a special trading day which are discussed in detail in Admati and Pfleiderer (1988). Although there are many possibilities for the existence of the order-flow shock, this thesis will not attempt to distinguish between these different reasons. Our aim is to find empirical evidence that the inclusion of the order-flow shock in the classical PIN model increases its explanatory power on actual data, which has important implications for a more accurate measure of information asymmetry.

2.2.1 Assumptions of the APIN model

There are two main differences between the adjusted PIN model and the classical PIN model. First of all, when informed events happens, we distinguish
the different arrival rates by buying and selling. Let $\delta_1$ denote the buyer informed trading arrival rate when good news happened and $\delta_{-1}$ denote the seller informed trading arrival rate when bad news happened. The reason for this is that buy-order flow has a larger variance than sell-order flow. To do this change, the adjusted model is more consistent with the characteristics of the actual data. Secondly, we add order-flow shock into adjusted PIN model. That common shock increase the number of buy and sell orders, in order to better match the characteristics of the actual data. In the adjusted PIN model, we use $\theta_C$ to denote the probability of a common shock. In the event of a common shock, let $\Delta_1$ denote the increase of buy intensity and $\Delta_{-1}$ denote the increase of sell intensity.

Figure 2 shows a entire trading process in a trading day. After we identified the specific informed event, the news are divided into common shock and no common shock.

\subsection{APIN model measure of information asymmetry}

Similar to the PIN model, information asymmetry can be measured using the adjusted probability of informed trading (APIN). In the APIN model, this probability is defined as

$$APIN = \frac{\theta_E[(1 - \theta_B)\delta_1 + \theta_B\delta_{-1}]}{\lambda_1 + \lambda_{-1} + \theta_E[(1 - \theta_B)\delta_1 + \theta_B\delta_{-1}] + \theta_C(\Delta_1 + \Delta_{-1})}.$$  \hspace{1cm} (9)

At the same time, we also can calculate the probability of symmetric order flow shock (PSOS). In the APIN model, this probability is defined as

$$PSOS = \frac{\theta_C(\Delta_1 + \Delta_{-1})}{\lambda_1 + \lambda_{-1} + \theta_E[(1 - \theta_B)\delta_1 + \theta_B\delta_{-1}] + \theta_C(\Delta_1 + \Delta_{-1})}.$$  \hspace{1cm} (10)

\subsection{APIN model parameter estimation}

Similar to the classical PIN model, we can write down the likelihood function of the APIN model, then maximize it. The likelihood function is

$$L[\theta_E, \theta_B, \theta_C, \lambda_1, \lambda_{-1}, \delta_1, \delta_{-1}, \Delta_1, \Delta_{-1} | B_d, S_d (d = 1, \ldots, D)]$$

$$= \prod_{d=1}^{D} \left[ (1 - \theta_E) \theta_C \frac{\lambda_1 + \Delta_1}{B_d} e^{-(\lambda_1 + \Delta_1)\frac{B_d}{B_d!}} e^{-(\lambda_1 + \Delta_1)\frac{S_d}{S_d!}} e^{-\lambda_1 - \Delta_1} \right. $$

$$+ (1 - \theta_E)(1 - \theta_C) \lambda_1 \lambda_{-1} \frac{B_d}{B_d!} e^{-\lambda_1 - \Delta_1} $$

$$+ \theta_E \theta_B \theta_C \frac{(\lambda_1 + \Delta_1)B_d}{B_d!} e^{-(\lambda_1 + \Delta_1)\frac{(\lambda_1 + \Delta_1)S_d}{S_d!}} e^{-(\lambda_1 + \Delta_1)} $$

$$+ \theta_E \theta_B \theta_C \frac{(\lambda_{-1} + \Delta_{-1})B_d}{B_d!} e^{-(\lambda_{-1} + \Delta_{-1})\frac{(\lambda_{-1} + \Delta_{-1})S_d}{S_d!}} e^{-(\lambda_{-1} + \Delta_{-1})}.$$
In the APIN model, the expectation of the daily number of buy orders is 2.2.4 Characteristics of the APIN model

The expectation of the daily number of sell orders is

\[ E(B_d) = \theta_E \theta_B \theta_C (\lambda_1 + \Delta_1) + \theta_E \theta_B (1 - \theta_C) \lambda_1 + \theta_E (1 - \theta_B) \theta_C (\lambda_1 + \Delta_1) + \theta_E (1 - \theta_B)(1 - \theta_C) \lambda_1 \]

\[ = \lambda_1 + \theta_E (1 - \theta_B) \delta_1 + \theta_C \Delta_1. \] (12)

The expectation of the daily number of sell orders is

\[ E(S_d) = \theta_E \theta_B \theta_C (\lambda_{-1} + \delta_{-1} + \Delta_{-1}) + \theta_E \theta_B (1 - \theta_C) (\lambda_{-1} + \delta_{-1}) + \theta_E (1 - \theta_B) \theta_C (\lambda_{-1} + \delta_{-1} + \Delta_{-1}) + (1 - \theta_E)(1 - \theta_C) \lambda_{-1} \]

\[ = \lambda_{-1} + \theta_E \theta_B \delta_{-1} + \theta_C \Delta_{-1}. \] (13)

The variance of the buy orders is

\[ Var(B_d) = \lambda_1^2 + \theta_E (1 - \theta_B) \delta_1^2 + \theta_C \Delta_1^2 + 2 \theta_E (1 - \theta_B) \lambda_1 \delta_1 + 2 \theta_C \lambda_1 \Delta_1 \]

\[ + 2 \theta_E (1 - \theta_B) \theta_C \delta_1 \Delta_1 - (\lambda_1 + \theta_E (1 - \theta_B) \delta_1 + \theta_C \Delta_1)^2 \]

\[ = \theta_E (1 - \theta_B)(1 - \theta_E (1 - \theta_B)) \delta_1^2 + \theta_C (1 - \theta_C) \Delta_1^2. \] (14)

The variance of the sell orders is

\[ Var(S_d) = \lambda_{-1}^2 + \theta_E \theta_B \delta_{-1}^2 + \theta_C \Delta_{-1}^2 + 2 \theta_E \theta_B \lambda_{-1} \delta_{-1} + 2 \theta_C \lambda_{-1} \Delta_{-1}. \]
\[ + 2\theta_E \theta_B \theta_C \delta_{-1} \Delta_{-1} - (\lambda_{-1} + \theta_E \theta_B \delta_{-1} + \theta_C \Delta_{-1})^2 \]

\[ = \theta_E \theta_B (1 - \theta_E \theta_B) \delta_{-1}^2 + \theta_C (1 - \theta_C) \Delta_{-1}^2. \quad (15) \]

The covariance between the daily number of buy and sell orders is

\[ \text{Cov}(B_d, S_d) = E(B_d S_d) - E(B_d) E(S_d) = \lambda_1 \lambda_{-1} + \theta_E \theta_B \lambda_1 \delta_{-1} + \theta_C \lambda_1 \Delta_{-1} \]

\[ + \theta_E (1 - \theta_B) \lambda_{-1} \delta_1 + \theta_E (1 - \theta_B) \theta_C \delta_1 \Delta_{-1} + \theta_C \lambda_{-1} \Delta_1 + \theta_E \theta_B \theta_C \delta_{-1} \Delta_1 \]

\[ + \theta_C \Delta_1 \Delta_{-1} - [\lambda_1 + \theta_E (1 - \theta_B) \delta_1 + \theta_C \Delta_1] (\lambda_{-1} + \theta_E \theta_B \delta_{-1} + \theta_C \Delta_{-1}) \]

\[ = \theta_C (1 - \theta_C) \Delta_1 \Delta_{-1} - \theta_E^2 \theta_B (1 - \theta_B) \delta_1 \delta_{-1}. \quad (16) \]

It can be seen from Equation (16) that, in contrast to the PIN model, the correlation between buy and sell orders can be positive in the APIN model (e.g. let \( \theta_C = \theta_B, \Delta_1 = \delta_1 \) and \( \Delta_{-1} = \delta_{-1} \)). Thus the APIN model allows for the pervasive positive correlation observed in actual data, Duarte & Young (2009).
3 Empirical results

3.1 Data

The intraday data used in this thesis is from the paper by Preve and Tse (2012). It includes the trading date, trading time, and bid-ask prices for some stocks traded on the NYSE over the period Jan 05 - Dec 07, a total of 754 trading days. In order to look for trends in estimated PIN and APIN, the data was divided into three sample periods: Jan 05 - Dec 05, Jan 05 - Dec 06 and Jan 05 - Dec 07. The sizes of the three samples were 252, 503 and 754, respectively. We selected four stocks as our object of study, General Electric (GE), International Business Machines (IBM), Procter & Gamble (PG) and Walmart (WMT). These four companies belong to the world's top one hundred companies by revenue. They are all leading companies in the industry. It facilitates our study that there is a lot of news released on these well-known companies.

The New York Stock Exchange daily opening time is from half past nine to four o'clock. In a previous study, Engle and Russell (1998) calculated the average trading volume of each trading period. At the opening, trading is very active. There is a transaction on average every ten seconds. After the opening period, trading time interval started increasing and frequency of transactions declined. Until about one o'clock, where the transaction is in an inactive state. The next transaction would happen on average 35 seconds. After this, transaction time interval decreed. Until closing on four o'clock, one transaction would happen every 20 seconds. Daily opening, trading is very frequently mainly due to the opening auction, but part of transactions often occurred delay. Between 9:30 and 9:50 which has just opened 20 minutes, opening auction decided stock price. So we deleted this unstable 20 minutes data.

Trade direction (buy- or sell-order was classified using the Lee and Ready (1991) algorithm. More than half of the trades for all stocks over the longest sample period were sell orders.)

3.2 Maximum Likelihood Estimation of the models

Maximum likelihood estimation of the PIN and APIN models was performed using the Matlab functions ga (genetic algorithm) and fmincon, with the interior-point algorithm and numerical derivatives. To search for a global optimum, we ran the likelihood optimization ten times for each data set. We then selected the maximum of these ten optimizations. The accompanying maximum likelihood estimates were then used to compute the PIN, APIN and PSOS measures. See tables 1-4, and 5-8, respectively.

For all four stocks, the (recursive) estimates of $\theta_E$ (the probability of information being released on day $d$) in the PIN model are lower than those for the APIN model. The estimates of $\theta_E$ for the PG stock are systematically higher than those for the other three stocks over the different time periods. The same is also true for the estimates of $\theta_B$ (the probability of bad news, conditional on the release of news) for PG, except for the APIN model corresponding to the period
Jan 05 - Dec 06. In general, the results indicate that the probability of bad news is much higher than the probability of good news for PG in 2005-2007. During this time Procter & Gamble’s reputation was, indeed, damaged: In December 2005, meeting concerning osteoporosis drug Actonel is still disputed in PG’s pharmaceutical department. On September 14, 2006, China General Administration of Quality Supervision, Inspection and Quarantine confirmed that SK Series cosmetics were detected to contain prohibited substances made in Japan Procter & Gamble company. PG refused to admit, accepted returns and exited the Chinese market. The firm had gone through several stages, the quality of their products and crisis management capabilities being questioned. In October 2007, in Georgia many users claimed that their teeth were stained and they lost taste feeling after using mouthwash manufactured by PG. Many evidences prove that between 2005-2007, the bad news about PG company influenced seriously on investment decisions. Let us observe the parameter $\theta_c$, which only appears in the APIN model. This parameter stands for the probability of a common shock. With the amount of data increasing, the probabilities of common shock occurrence of all firms are reduced. The greatest change is General Electric and the smallest change is Procter & Gamble.

### 3.3 Estimates of PIN, APIN and PSOS

We applied equations (1), (9) and (10) to calculate empirical PIN, APIN and PSOS values for our four companies. In order to look for trends, data from 2005 to 2007 was divided into three sample periods: Jan 05 - Dec 05, Jan 05 - Dec 06, and Jan 05 - Dec 07. As seen in tables 5-8, the PIN and APIN values appear to increase as the sample size used for parameter estimation increases. Only the WMT APIN value for Jan 05 - Dec 05 is slightly lower than the corresponding value for Jan 05 - Dec 06, but the decline is quite small. So we think that the overall trend is positive. We then compared the corresponding PIN and APIN estimates and found that APIN generally is lower than PIN, except for GE over the period Jan 05 - Dec 05. This is also depicted in Figure 3. All PIN and APIN values, which represent the proportion of informed trading to total trading, are between 0 and 0.1. PSOS values, which represent the proportion of the trading caused by symmetric order-flow shock to total trading, are almost always larger than 0.1.

Of course, it is also possible that the observed trends in empirical PIN and APIN over the period Jan 05 - Dec 07 is due to changes in the market that took place in this period. Specifically, more trades were done in limit order than through specialists (i.e. market makers). This change makes the identification/classification of buy/sell orders very difficult.

---

2The year of 2006 represents an important shift for the NYSE from specialist market to limit order market, and in 2007 the limit order market became far more important than the specialist market.
3.4 Model selection

For model selection, we use Akaike’s Information Criterion (AIC) and Schwartz’s Bayesian Criterion (SBC), given by

\[ AIC = -2\hat{l} + 2k, \]  

and

\[ SBC = -2\hat{l} + \ln(D)k, \]  

where \( \hat{l} \) is the value of the log-likelihood evaluated at the maximum (i.e. at the MLE), \( k \) is the number of parameters of the estimated model (i.e. 5 for the PIN model and 9 for the APIN model), and \( D \) is the number of observations (i.e. days) in the sample. The smaller the AIC or SBC, the better the model fit. It can be seen from tables 1-4 that all AIC and SBC values for the APIN model are smaller than those for the PIN model. Consequently, the APIN model is preferred over the PIN model.

We use the AIC and SBC as measures of the relative goodness of fit of the PIN and APIN models instead of the likelihood ratio test statistic as the standard regularity conditions for this test to be asymptotically chi-squared distributed under the null are not satisfied when testing the restricted PIN model against the unrestricted APIN model.

4 Conclusions

This thesis reviewed the classical PIN model and its recent extension, the adjusted PIN model. Compared with the classical PIN model, the adjusted PIN model allows for the arrival rate of informed sellers to be different from the arrival rate of informed buyers, and for a symmetric order-flow shock. We calculated PIN, APIN and PSOS over three different time periods for the GE, IBM, PG and WMT stocks. Finally, we used the AIC and SBC criterion to evaluate the fit of the models. After comparison, we found that the APIN model provides the best fit for our data.
References


Table 1: Recursive parameter estimates for the PIN and APIN models and the GE stock.

<table>
<thead>
<tr>
<th>Est</th>
<th>Jan 05 - Dec 05</th>
<th>Jan 05 - Dec 06</th>
<th>Jan 05 - Dec 07</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PIN</td>
<td>APIN</td>
<td>PIN</td>
</tr>
<tr>
<td>( \theta_E )</td>
<td>0.3646</td>
<td>0.5428</td>
<td>0.4023</td>
</tr>
<tr>
<td>( \theta_B )</td>
<td>0.4425</td>
<td>0.2903</td>
<td>0.6000</td>
</tr>
<tr>
<td>( \lambda_1 )</td>
<td>1867.1</td>
<td>1598.6</td>
<td>1935.1</td>
</tr>
<tr>
<td>( \lambda_{-1} )</td>
<td>2311.7</td>
<td>2068.4</td>
<td>2360.0</td>
</tr>
<tr>
<td>( \delta )</td>
<td>434.8799</td>
<td>541.8377</td>
<td>1584.9</td>
</tr>
<tr>
<td>( \delta_{C} )</td>
<td>0.6289</td>
<td>0.4557</td>
<td>0.2598</td>
</tr>
<tr>
<td>( \delta_{1} )</td>
<td>299.0984</td>
<td>566.7462</td>
<td>948.0710</td>
</tr>
<tr>
<td>( \delta_{-1} )</td>
<td>376.7765</td>
<td>370.5389</td>
<td>631.7460</td>
</tr>
<tr>
<td>( \Delta_1 )</td>
<td>382.9385</td>
<td>377.7622</td>
<td>1346.5</td>
</tr>
<tr>
<td>( \Delta_{-1} )</td>
<td>405.3601</td>
<td>595.3876</td>
<td>1702.8</td>
</tr>
<tr>
<td>AIC</td>
<td>15995.8</td>
<td>10866.4</td>
<td>45932</td>
</tr>
<tr>
<td>SBC</td>
<td>16013.3</td>
<td>10877.9</td>
<td>45953</td>
</tr>
</tbody>
</table>
Table 2: Recursive parameter estimates for the PIN and APIN models and the IBM stock.

<table>
<thead>
<tr>
<th>Est</th>
<th>Jan 05 - Dec 05</th>
<th>Time Period</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PIN</td>
<td>APIN</td>
</tr>
<tr>
<td>θE</td>
<td>0.3512</td>
<td>0.5026</td>
</tr>
<tr>
<td>θB</td>
<td>0.3578</td>
<td>0.5093</td>
</tr>
<tr>
<td>λ_1</td>
<td>1834.7</td>
<td>1644.1</td>
</tr>
<tr>
<td>λ_-1</td>
<td>2039.1</td>
<td>1770.6</td>
</tr>
<tr>
<td>δ</td>
<td>607.3587</td>
<td>683.2669</td>
</tr>
<tr>
<td>θC</td>
<td></td>
<td>0.4524</td>
</tr>
<tr>
<td>δ_1</td>
<td>485.3261</td>
<td>493.3153</td>
</tr>
<tr>
<td>δ_-1</td>
<td>258.0794</td>
<td>407.0265</td>
</tr>
<tr>
<td>Δ_1</td>
<td>465.2085</td>
<td>675.2018</td>
</tr>
<tr>
<td>Δ_-1</td>
<td>610.3756</td>
<td>716.1394</td>
</tr>
<tr>
<td>AIC</td>
<td>27014</td>
<td>15074.6</td>
</tr>
<tr>
<td>SBC</td>
<td>27031.5</td>
<td>15106.1</td>
</tr>
</tbody>
</table>

Table 3: Recursive parameter estimates for the PIN and APIN models and the PG stock.

<table>
<thead>
<tr>
<th>Est</th>
<th>Jan 05 - Dec 05</th>
<th>Time Period</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PIN</td>
<td>APIN</td>
</tr>
<tr>
<td>θE</td>
<td>0.4348</td>
<td>0.5654</td>
</tr>
<tr>
<td>θB</td>
<td>0.7007</td>
<td>0.3941</td>
</tr>
<tr>
<td>λ_1</td>
<td>1760.0</td>
<td>1520.3</td>
</tr>
<tr>
<td>λ_-1</td>
<td>1660.2</td>
<td>1540.6</td>
</tr>
<tr>
<td>δ</td>
<td>506.4925</td>
<td>617.2061</td>
</tr>
<tr>
<td>θC</td>
<td></td>
<td>0.4334</td>
</tr>
<tr>
<td>δ_1</td>
<td>318.0011</td>
<td>435.7759</td>
</tr>
<tr>
<td>δ_-1</td>
<td>363.2760</td>
<td>420.5001</td>
</tr>
<tr>
<td>Δ_1</td>
<td>452.6840</td>
<td>528.3535</td>
</tr>
<tr>
<td>Δ_-1</td>
<td>446.4459</td>
<td>735.2203</td>
</tr>
<tr>
<td>AIC</td>
<td>19899</td>
<td>12458.8</td>
</tr>
<tr>
<td>SBC</td>
<td>19916.5</td>
<td>12490.3</td>
</tr>
</tbody>
</table>
Table 4: Recursive parameter estimates for the PIN and APIN models and the WMT stock.

<table>
<thead>
<tr>
<th>Est</th>
<th>Jan 05 - Dec 05</th>
<th>Jan 05 - Dec 06</th>
<th>Jan 05 - Dec 07</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PIN</td>
<td>APIN</td>
<td>PIN</td>
</tr>
<tr>
<td>θ₁</td>
<td>0.4139</td>
<td>0.5142</td>
<td>0.3994</td>
</tr>
<tr>
<td>θ₂</td>
<td>0.3224</td>
<td>0.1936</td>
<td>0.3452</td>
</tr>
<tr>
<td>λ₁</td>
<td>1871.2</td>
<td>1628.2</td>
<td>1910.1</td>
</tr>
<tr>
<td>λ₋₁</td>
<td>2303.5</td>
<td>2207.0</td>
<td>2475.2</td>
</tr>
<tr>
<td>δ</td>
<td>511.0065</td>
<td>591.3072</td>
<td></td>
</tr>
</tbody>
</table>

Table 5: Empirical PIN, APIN and PSOS for GE calculated over the periods Jan05-Dec05, Jan05-Dec06 and Jan05-Dec07.

| | PIN | APIN | PSOS |
| | Jan 05 - Dec 05 | 0.0366 | 0.0403 | 0.1143 |
| | Jan 05 - Dec 06 | 0.0483 | 0.0446 | 0.0983 |
| | Jan 05 - Dec 07 | 0.0821 | 0.0502 | 0.1462 |

Table 6: Empirical PIN, APIN and PSOS for IBM calculated over the periods Jan05-Dec05, Jan05-Dec06 and Jan05-Dec07.

| | PIN | APIN | PSOS |
| | Jan 05 - Dec 05 | 0.0522 | 0.0454 | 0.1191 |
| | Jan 05 - Dec 06 | 0.0572 | 0.0464 | 0.1093 |
| | Jan 05 - Dec 07 | 0.0685 | 0.0574 | 0.1243 |

Table 7: Empirical PIN, APIN and PSOS for PG calculated over the periods Jan05-Dec05, Jan05-Dec06 and Jan05-Dec07.

| | PIN | APIN | PSOS |
| | Jan 05 - Dec 05 | 0.0605 | 0.0522 | 0.1071 |
| | Jan 05 - Dec 06 | 0.0693 | 0.0554 | 0.1073 |
| | Jan 05 - Dec 07 | 0.0808 | 0.0705 | 0.1337 |
Table 8: Empirical PIN, APIN and PSOS for WMT calculated over the periods Jan05-Dec05, Jan05-Dec06 and Jan05-Dec07.

<table>
<thead>
<tr>
<th>Period</th>
<th>PIN</th>
<th>APIN</th>
<th>PSOS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan 05 - Dec 05</td>
<td>0.0473</td>
<td>0.0439</td>
<td>0.0993</td>
</tr>
<tr>
<td>Jan 05 - Dec 06</td>
<td>0.0511</td>
<td>0.0418</td>
<td>0.1119</td>
</tr>
<tr>
<td>Jan 05 - Dec 07</td>
<td>0.0609</td>
<td>0.0516</td>
<td>0.1207</td>
</tr>
</tbody>
</table>

Figure 1: Outline of the PIN model: $\theta_E$ is the probability of news. $\theta_B$ is the probability that the news is bad. $B_d$ and $S_d$ are the daily total number of buy and sell orders, assumed to be Poisson distributed. If the news is bad, the sell intensity increases by $\delta$. If the news is good, the buy intensity increases by $\delta$. 

17
Figure 2: Outline of the APIN model: $\theta_E$ is the probability of news. $\theta_B$ is the probability that the news is bad. $\theta_C$ is the probability of a common shock. $B_d$ and $S_d$ are the daily total number of buy and sell orders, assumed to be Poisson distributed. If the news is bad, the sell intensity increases by $\delta_1$. If the news is good, the buy intensity increases by $\delta_l$. If there is a common shock, the buy and sell intensity increases by $\Delta_1$ and $\Delta_{-1}$, respectively.
Figure 3: Recursive PIN and APIN estimates for the GE stock.

Figure 4: Recursive PIN and APIN estimates for the IBM stock.
Figure 5: Recursive PIN and APIN estimates for the PG stock.

Figure 6: Recursive PIN and APIN estimates for the WMT stock.