

New drugs and stock market: how to predict pharma market reaction to clinical trial announcements

Semen BUDENNY^{1,2,*}, Alexey KAZAKOV¹, Elizaveta KOVTUN¹, and Leonid ZHUKOV³

¹Sber AI Lab, Moscow

²Artificial Intelligence Research Institute (AIRI), Moscow

³Higher School of Economics University, Moscow

*Corresponding author: sanbudenny@sberbank.ru

Supplementary materials

Post-announcement period calculation

To estimate the duration of the announcement effect, we analyze trading volume peaks in the companies' history. These peaks do not exclusively mean the reaction to the trial results but are rather associated with all events related to the company. We calculate the duration in days of all available trading volume peaks. The results are presented in Figure 1. As we can see, the duration of 90% peaks constitutes 20 trade days. Thus, we take 20 days as the post-announcement time period we are interested in.

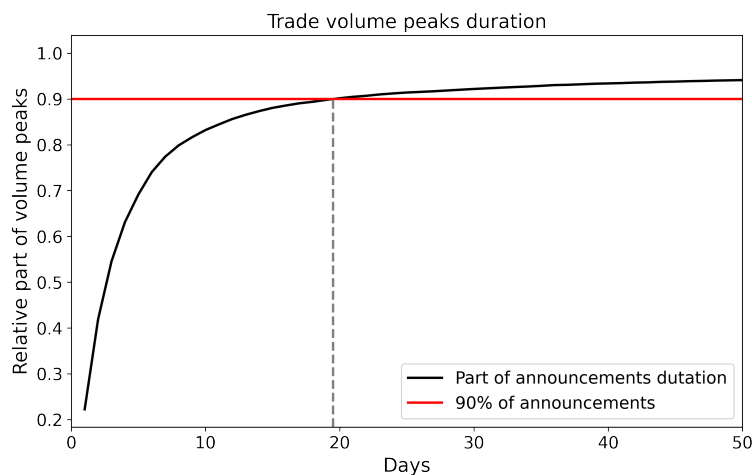


Figure 1. The correspondence between the share of considered trading volume peaks with the peak duration. Redline is the 90% of the total number of peaks.

Other relationships between announcement impact and company characteristics

The dependence of the number of FDA news announcements on the year for the public companies is given in Figure 2. It demonstrates the increase in the number of FDA announcements in open sources by approximately ten times from 2017 to 2021.

Figures 3, 4 and 5 show the dependence of stock price changes caused by announcements on the company age since entering the IPO at the announcement moment, historical volatility, and the company's 30-day stock price trend. Volatility is defined as the relation of the standard deviation of stock price to the median stock price in a 200-day period before the announcement. We can conclude that the younger company is, the more sensitive it is to the announcements. Greater intrinsic company volatility results in more extreme price changes. In addition, if the company trend of a stock price is about zero in the pre-event period, then the induced price change is not significant.

Mismatch analysis between announcement polarities and price change

To realize the mismatch between events sentiment in accordance with FDA announcements and events sentiment in accordance with the actual price change, we build a confusion matrix shown in Figure 6. The sentiment polarity derived from stock price

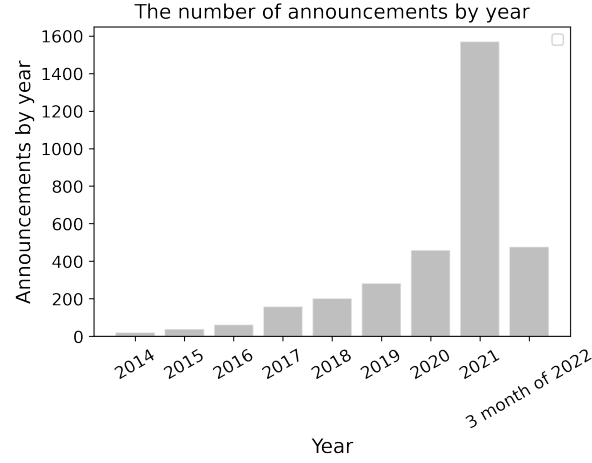


Figure 2. The number of FDA announcements of public companies in different years.

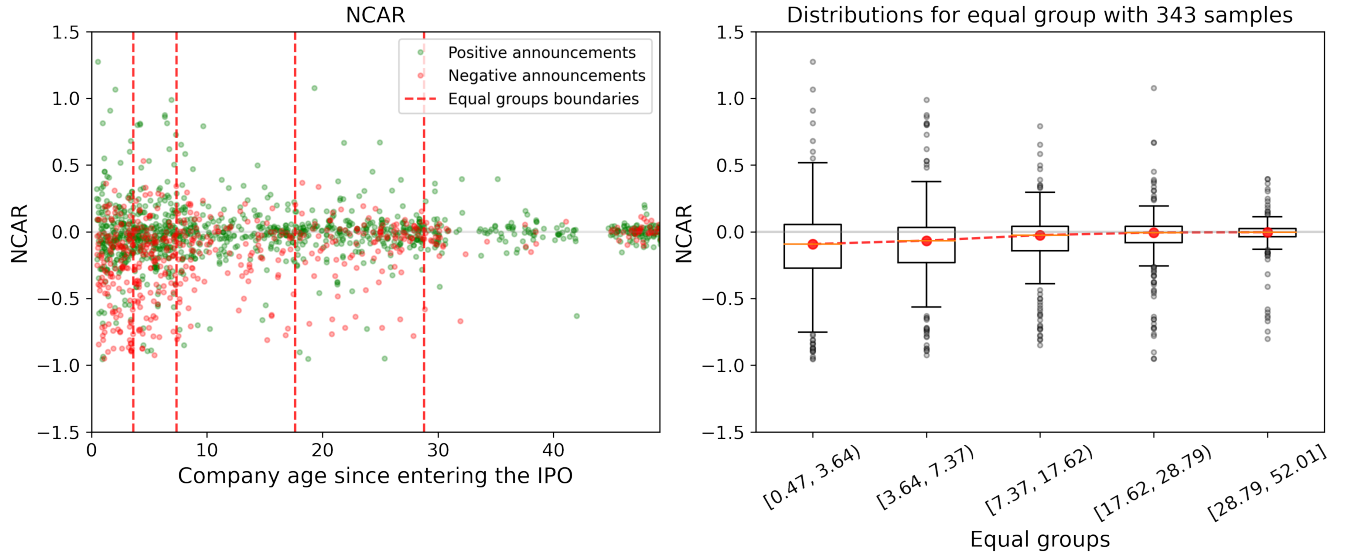


Figure 3. Dependence of the stock price changes on the company age. In the left part, the red dashed lines divide announcements into equal groups (with the same number of announcements inside). In the right part, the red dashed line goes through the median values of each group's stock price changes.

change, namely $NCAR_{20}$, is defined as follows: the event is set to be positive or negative one in accordance with $NCAR_{20}$ if its price change is more than $\sigma_{neut}/2 = 11.5\%$ (σ_{neut} - standard deviation of $NCAR_{20}$ for neutral announcements) or less than $-\sigma_{neut}/2$, respectively. If the price change lies within $[-\sigma_{neut}/2; +\sigma_{neut}/2]$, then we define such event as a neutral one in accordance with $NCAR_{20}$. As a result, the highest rates in the confusion matrix constitute 23.7% and 24.1% and relate to the cases in which negative announcements have negative or neutral price changes.

Statement on computational resources and environmental impact

We used a NVIDIA GeForce RTX 2070 SUPER GPU and NVIDIA A100 80GB PCIe GPU to train the classifiers and the BERT model, correspondingly.

This work contributed 1.34 kg and 6 g of equivalent CO_2 emissions during the classifiers and BERT training, respectively. The carbon emissions information was generated using the open-source library eco2AI (<https://github.com/sb-ai-lab/Eco2AI>).

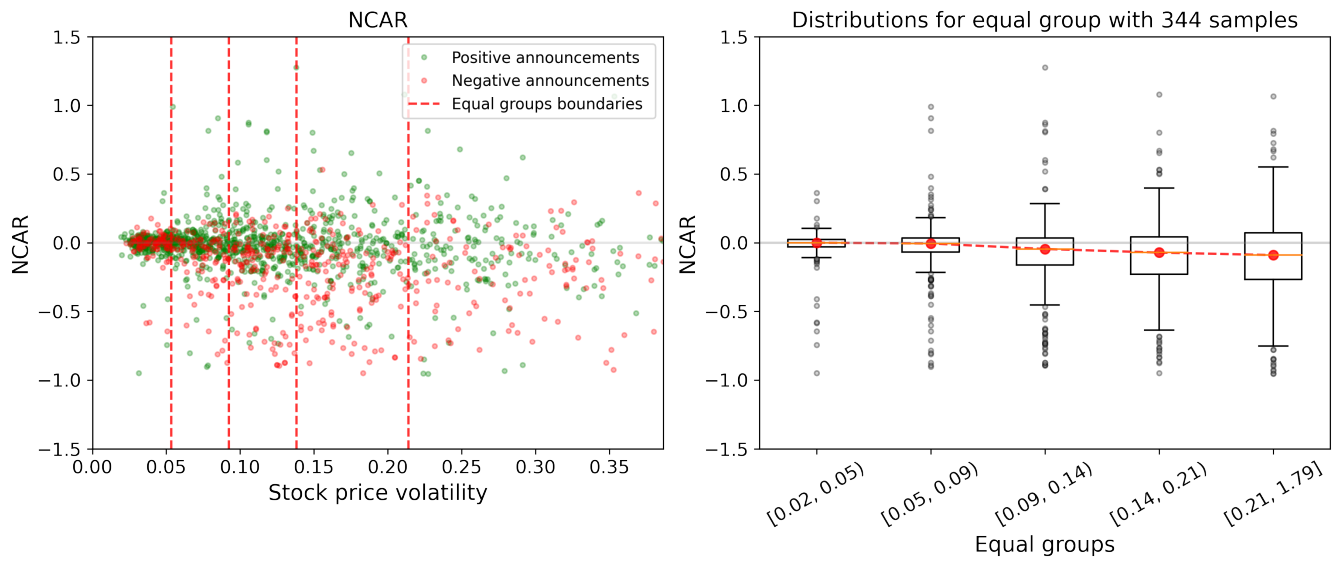


Figure 4. Dependence of the stock price changes on the company volatility. In the left part, the red dashed lines divide announcements into equal groups (with the same number of announcements inside). In the right part, the red dashed line goes through the median values of each group's stock price changes.

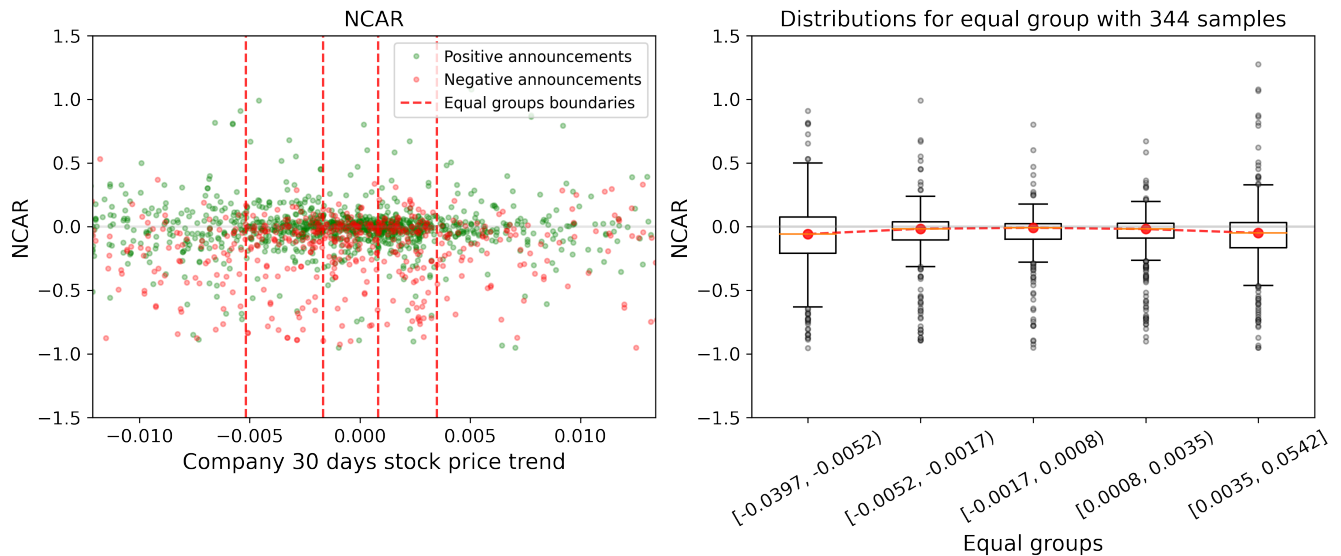


Figure 5. Dependence of the stock price changes on the 30-day trend before the announcement. In the left part, the red dashed lines divide announcements into equal groups (with the same number of announcements inside). In the right part, the red dashed line goes through the median values of each group's stock price changes.

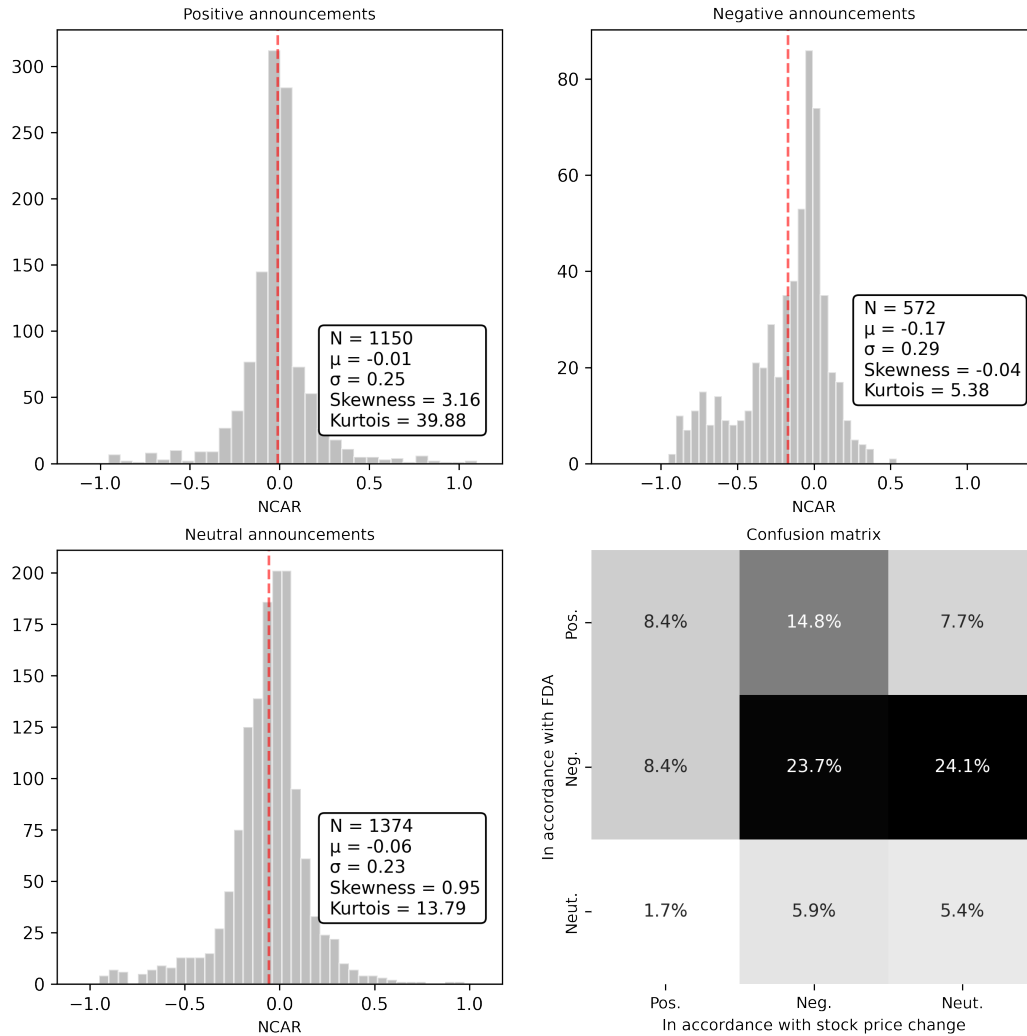


Figure 6. The price change distribution ($NCAR_{20}$) for positive (upper left), negative (upper right), neutral (lower left) FDA news with statistical parameters provided in the boxes (number of events N , mean value μ , standard deviation σ , skewness and kurtosis coefficients). The mean value is depicted with the red vertical dashed line. The confusion matrix (lower right) of mismatch between sentiments in accordance with FDA announcements and polarities in accordance with the stock price change ($NCAR_{20}$).