In order to analyse the effect of passive ownership on volatility around earnings dates, we use an event study specification with fixed effects as shown below:

$$\text{Volatility}_{i,t} = \delta \text{Passive}_{i,t} + \sum_{k \neq -1} \beta_k D_{i,k} + \sum_{k \neq -1} \gamma_k D_{i,k} \times \text{Passive}_i + \mathbb{X}_{i,t}\Theta + \alpha_i + \epsilon_{i,t}$$

Where, the unit i is a stock on day t, $\text{Volatility}_{i,t}$ is either 3, 5 or 7 day volatility. $\text{Passive}_{i,t}$ is the share of a stock passively held. $D_{i,t}$ is a dummy variable equal to 1 at time t. $\mathbb{X}$ are controls such as total assets, revenue, volume traded and earnings per share as they are correlated to volatility around earnings dates. α_i are entity fixed effects, which are included to absorb time invariant firm characteristics not included in the controls.

It is important to note that the variable Passive is quarterly, so it does not change day to day. This is also the case for some of the control variables.

β_k captures the average change in volatility on event day k relative to day -1 for all firms, while γ_k captures how this effect differs with passive ownership. A negative γ_k before earnings and a positive γ_k after earnings would indicate that higher passive ownership reduces pre-announcement volatility but amplifies post-announcement reactions.

Since simultaneous causality is a potential endogeneity issue, I use a two stage least squares specification. The key identification assumption is that, conditional on firm characteristics such as size and liquidity, inclusion in the S&P 500 affects volatility only through its effect on passive ownership.

The first stage is:

$$\text{Passive}_{i,t} = \alpha + \beta \text{ S\&P inclusion}_{i,t} + \mathbb{W}_{i,t}\Theta + \nu_{i,t}$$

Using the first stage we compute $\widehat{\text{Passive}}$ and estimate the second stage:

$$\text{Volatility}_{i,t} = \delta \widehat{\text{Passive}}_{i,t} + \sum_{k \neq -1} \beta_k D_{i,k} + \sum_{k \neq -1} \gamma_k D_{i,k} \times \widehat{\text{Passive}}_i + \mathbb{W}_{i,t}\Theta + \alpha_i + \epsilon_{i,t}$$

Similar to the first model, the coefficients β_k and γ_k, $\forall k \in \{-30, \ldots -2, 0, 1 \ldots 15\}$ are the coefficients of interest. However, by substituting the variable Passive with $\widehat{\text{Passive}}$ into the event study specification the coefficients capture the causal response of volatility to exogenous variation in passive ownership.