ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ «НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ «ВЫСШАЯ ШКОЛА ЭКОНОМИКИ»

Международный институт экономики и финансов

Супруненко Георгий Михайлович

Suprunenko Georgii Mikhailovich

ДИНАМИКА ОЦЕНОК ДОХОДНОСТИ: СХОДИМОСТЬ СРЕДНЕГО И ДИСПЕРСИИ НОВЫХ АКЦИЙ И ETF (EVOLUTION OF RETURN ESTIMATES: MEAN AND VARIANCE CONVERGENCE FOR NEWLY LISTED STOCKS AND ETFS)

Курсовая работа - БАКАЛАВРСКАЯ РАБОТА

по направлению подготовки 38.03.01 «Экономика» образовательная программа «Международная программа по экономике и финансам»

Научный руководитель

____

Потапов Д.И.

Content

1. Introduction

In the dynamic world of financial markets, newly listed securities often exhibit behavior markedly different from that of their more established peers. The phenomenon of initial price jumps, heightened trading volume, and volatile swings has been well documented in both academic literature and practitioner reports. Yet, the process by which these young assets' statistical properties—specifically their average returns and volatility—evolve over time remains an area of active inquiry. This paper investigates how the mean and variance of daily returns for U.S.-traded stocks and exchange‑traded funds (ETFs) settle into a steady state during their first 100 post-listing days.

In academic research, mapping the convergence of young assets' return and volatility deepens our understanding of market efficiency and information diffusion by quantifying how quickly new listings transition from elevated initial uncertainty to equilibrium behavior—insights that enrich models of price discovery and inform subsequent empirical studies on lifecycle effects in asset pricing. In real‑world trading, awareness of a 100‑day "settling period" aids traders in calibrating entry and exit rules, exploiting abnormal returns and hedging risks.

This study aims to provide a systematic, data‑driven examination of convergence dynamics. By tracking the daily returns of a broad cross‑section of U.S.-listed equities and ETFs from their first through their hundredth trading day, we chart the trajectory of both mean returns and return variance. The main objective is to retrieve the underlying shape of the convergence path of mean and variance of the returns that would serve as a generalisation for all assets of this class, irrespective of their specifics. According to financial and economic intuition, we expect shapes that would indicate larger price jumps in the first days that later stabilise around some long-term average. We expect this to be visually observable in the first 2 moments of the returns' distributions.

After estimating these convergence functions of time, we will rigorously evaluate their robustness and practical relevance using two complementary strategies. First, we will perform residual-based hypothesis tests to verify whether the fitted convergence patterns are significant for mean and variance separately. Second, we will employ a maximum likelihood test to validate the functional forms of mean and variance simultaneously, for this we will additionally determine the best approximating distribution for assets' returns. Finally, we will synthesize these results to determine whether the observed temporal effects hold consistently across the aggregate sample and at the individual‑asset level. If the effects prove significant and generalizable, we will then explore their predictive power for out‑of‑sample data, assessing whether the established convergence patterns can reliably forecast future return and volatility dynamics for newly listed securities.

2. Literature Review

Newly listed assets (IPOs or first-day ETFs) often exhibit pronounced price and volatility anomalies that fade over time. In equity markets, this includes large first-day “pops” (underpricing) and very high initial volatility, which then typically decline as markets absorb information. Understanding these convergence dynamics is important for pricing, risk management and forecasting. The literature on post-listing behavior draws on both efficient-market and behavioral theories, employs volatility models (ARCH/GARCH, stochastic volatility), and examines U.S. and international cases for stocks and ETFs.

Young stocks often show extreme returns and volatilities immediately after listing. IPOs tend to debut well above their offer price. For example, Lowry, Officer, and Schwert (2010) report that from 1965–2005 the average IPO “pop” was about 22% (with a 55% standard deviation). Such underpricing is generally attributed to information asymmetry, underwriter incentives, or signaling (Bruce, A. and Thilakaratne, P., 2014). Behavioral factors may also play a role: hot markets and investor sentiment can inflate initial prices. Not all new issues go up, however – Lowry et al. note about one-third of IPOs have negative first-day returns. The initial volatility of young assets is found to be high, the researchers link it to the difficulty of pricing IPOs and conclude that “the monthly volatility of IPO initial returns is substantial, fluctuates dramatically over time, and is considerably larger during "hot" IPO markets” (Lowry, Officer, and Schwert, 2010, p.425)

The convergence of returns and volatility in young assets is interpreted through multiple lenses. Traditional models view underpricing as a rational premium for uncertainty (Rock, 1986) or a marketing device (Benveniste, L. M., & Spindt, P. A., 1989). Under these theories, anomalies should dissipate as the firm’s true value and risk become known. Behavioral theories highlight over-optimism, herding or sentiment in hot issue periods. For volatility, standard finance assumes markets quickly incorporate new information, so any excess volatility should revert as new data arrive. Instead, researchers document persistent volatility clustering, suggesting that shocks (including those from hot listings) take time to wash out.

Empirically, GARCH-type models are widely used to study volatility convergence. Beneda and Zhang (2009) employ GARCH(1,1) to study the relation between initial idiosyncratic volatility level and the post-IPO volatility change, i.e. information diffusion. Lowry, Officer, and Schwert (2010) also fit GARCH‑type models to the U.S. IPO return series, finding that initial first‑day variances are on average twice those of seasoned stocks.

Evidence shows that after extreme initial returns and volatilities, young assets’ returns typically revert toward fundamentals while volatility settles. While the literature richly describes initial anomalies (underpricing, high volatility) and theoretical models, explicit studies of the *convergence path* are fewer. Few works systematically trace *both* return and volatility convergence together. Moreover, ETF listings are under-explored – much ETF literature focuses on funds’ impact on underlying securities, not on the ETF’s own post-listing behavior.

3. Practical Implementation
3.1 Full-Sample Fit
3.1.1 Analysis of Mean

We will begin with outlining the functional form of the model. We will examine the time effects on mean and variance of returns based on panel OLS regression with heterogeneous variance. We will consider fixed effect OLS to be able to separate entity and time effects and obtain exact values as functions. Generally, such a model is consistent and unbiased, however, not efficient due to the presence of heterogeneity with respect to time and entity. This drawback will be compensated by large sample size such that estimated parameters will have to have acceptable standard deviation. Additionally, we will use heterogeneity robust tests on means to ensure they are valid. Generally the model will be:

$$
\begin{gathered}
R_{it} = \beta + \lambda_i + \gamma_t + u_{it} \\
\beta - \text{common intercept} \\
\lambda_i - \text{asset effect} \\
\gamma_t - \text{time effect} \\
u_{it} - \text{disturbance term}
\end{gathered}
$$

The dataset that will be used for the purposes of this analysis is a large-volume file containing historical daily prices for over 5000 US stocks and ETFs. The dataset is published by Eric Stanley on Kaggle. According to the author, the prices were retrieved from Yahoo Finance and IEX Cloud API. For this study, we calculated daily returns of assets:

$$R_t = \frac{P_t}{P_{t-1}} - 1$$

Each column represents an asset with the index being the day from listing (relative, not absolute dates) up to a hundredth day. The dataframe takes the form:

	INCR	EURN	NRP	AUB	ROIV	RBA	SBGI	EBON	JCTCF	EXAS	...	PFC	GM	BOTJ	GROV	DRH	ODD	SITM	CPBI	VRTS	TEL
1	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
2	0.358566	0.009836	0.010309	0.100001	-0.042308	0.020711	-0.020725	-0.100000	0.000000	-0.041667	...	0.000000	0.002048	0.000000	0.015228	0.075472	0.051967	0.434615	-0.005495	-0.160952	-0.001287
3	-0.120235	-0.017857	-0.002551	-0.011364	-0.006024	0.037681	0.005291	0.051111	-0.031250	-0.021739	...	0.009091	-0.005254	0.000000	-0.015100	-0.017544	0.051200	-0.032172	0.000000	0.349603	-0.042525
4	-0.033333	0.016529	-0.025576	0.057472	0.001919	-0.027933	0.010526	-0.160677	0.000000	-0.004444	...	-0.009009	-0.024355	0.000000	0.000000	0.000000	-0.054985	-0.049861	0.002210	-0.138772	-0.012113
5	0.000000	-0.051219	0.004725	0.000000	0.008166	0.048850	-0.020833	0.100756	0.032258	0.000000	...	0.018182	0.006918	0.090909	0.000000	-0.013393	-0.027783	-0.008163	-0.007718	-0.032227	-0.020436
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
96	0.000000	0.004008	0.004349	-0.062761	0.000000	-0.004386	-0.029126	-0.108508	0.000000	-0.015018	...	0.008849	0.000000	0.000000	0.001028	-0.009523	0.069542	-0.043127	0.000000	0.014706	0.020178
97	0.000000	-0.015968	0.004764	0.017857	0.001019	-0.030837	0.015000	0.098202	0.000000	0.125561	...	-0.052632	-0.017037	0.000000	0.001541	-0.012238	-0.021125	-0.012676	0.002927	-0.026465	0.018034
98	-0.150000	-0.002704	0.007759	-0.017543	0.000000	-0.031818	-0.044335	-0.080605	0.000000	-0.066932	...	0.000000	-0.024451	0.000000	-0.004103	0.005309	-0.004765	-0.063718	-0.005837	-0.019417	0.019143
99	0.029412	-0.009492	-0.014115	0.071428	0.004073	-0.014084	0.000000	0.008219	0.000000	0.031597	...	-0.027778	-0.023794	-0.023809	0.001030	0.014965	0.063081	0.057389	0.001957	0.019142	-0.014017
100	0.142857	0.008898	0.019523	-0.066667	0.009128	0.009524	-0.041237	0.031250	0.000000	-0.002483	...	0.028572	0.012350	-0.024390	0.000000	0.007805	-0.001324	-0.039385	-0.002930	0.017487	0.005687

100 rows × 5468 columns

Table 1: Assets’ daily returns, first 100 days from listing; Wide format

Additionally, for the convenience of panel regression, we will use the same dataframe in the long format with added constant column:

i	t	ret	const
INCR	2	0.154050	1.0
	3	-0.120235	1.0
	4	-0.033333	1.0
	5	0.000000	1.0
	6	0.017241	1.0
...	...	...	...
TEL	96	0.020178	1.0
	97	0.018034	1.0
	98	0.019143	1.0
	99	-0.014017	1.0
	100	0.005687	1.0

541051 rows × 2 columns

Table 2: Assets' daily returns, first 100 days from listing; Long format

To test the significance of fixed effects we will first run dummy panel regression on the first 100'000 rows of the long format data (as the full regression requires a construction of an excessively large regressors matrix):

$$\hat{R}_{it} = \hat{\beta} + \sum_{i=0}^{N}(\hat{\lambda}_i \times D_i^{Asset}) + \sum_{t=0}^{T}(\hat{\gamma}_t \times D_t^{Time})$$

We will utilise RSS F-test with heteroskedasticity consistent covariance matrix HC3 estimates, which features the highest power amongst HC0-HC3 estimators (see Appendix A)

Thus we will compare 3 models:

1. Baseline - only intercept
2. Entity-only - intercept and dummy for asset
3. Entity+Time - unrestricted full version

RSS F-test yields the following results:

```
Baseline vs Entity-only:
   df_resid         ssr  df_diff   ss_diff         F    Pr(>F)
0   99999.0  152.228191      0.0       NaN       NaN       NaN
1   98989.0  150.365288   1010.0  1.862904  1.214251  0.000003

Entity-only vs Entity+Time:
   df_resid         ssr  df_diff   ss_diff         F    Pr(>F)
0   98989.0  150.365288      0.0       NaN       NaN       NaN
1   98891.0  150.108892     98.0  0.256396  1.723595  0.000012

All three models:
   df_resid         ssr  df_diff   ss_diff         F    Pr(>F)
0   99999.0  152.228191      0.0       NaN       NaN       NaN
1   98989.0  150.365288   1010.0  1.862904  1.215121  0.000003
2   98891.0  150.108892     98.0  0.256396  1.723595  0.000012
```

Table 3: The results of the F-test applied to 3 versions of the regression

According to Table 3, Time and Entity effects are significant jointly and individually at 1% heteroskedasticity-wise. R^2 is 0.014, which is significant at 1% as well.

From now on, we will utilise a demeaning approach to panel regression, as it yields exactly the same results but is much more efficient computationally- and memory-wise:

$$\hat{\beta} = \bar{R}_{\cdot\cdot} = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} R_{it}$$

$$\hat{\lambda}_i = \bar{R}_{i\cdot} - \hat{\beta}$$

$$\hat{\gamma}_t = \bar{R}_{\cdot t} - \hat{\beta}$$

Now we can plot the retrieved gamma values as a function of time (day since listing):

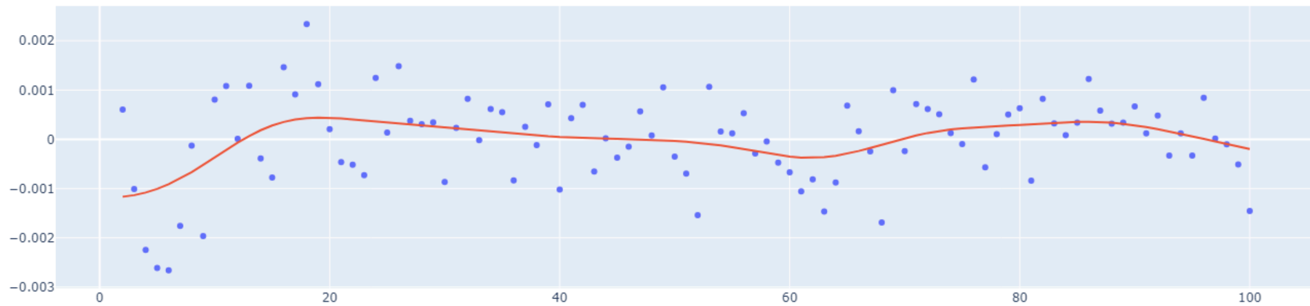

Figure 1: Returns' mean time effects, 100-day dynamics

The red line on the chart is a kernel regression with Gaussian kernel and bandwidth h=5:

$$\widehat{f}(x) = \frac{\sum_{i=1}^{n} K_h(x, x_i)\, y_i}{\sum_{i=1}^{n} K_h(x, x_i)}, K_h(x, x_i) = \exp\left(-\frac{(x - x_i)^2}{2h^2}\right)$$

According to Figure 1, following initial growth on the first trading day, returns are consistently negative over the subsequent seven trading days. After this period the returns stabilise and look approximately stationary around zero. This dynamic can be observed on the chart with expanding-window averaging:

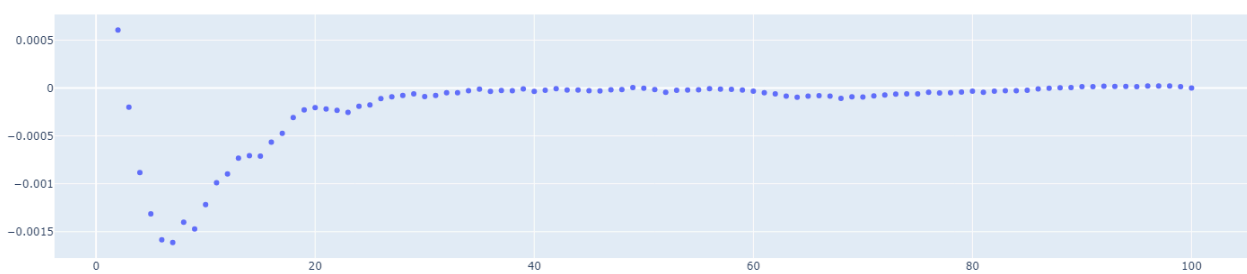

Figure 2: Rolling average of returns' mean time effects, 100-day dynamics

Figure 2 supports the evidence, as after the first positive-return day, the values drop sharply, reflecting the early negative returns. Later they gradually increase, stabilizing around zero, indicating the alleged stationarity of subsequent returns.

To rigorously test that the time effects become stationary over time, we will perform an ADF test (no trend t-test) on the expanding-window subsamples of gamma (Dickey, D. A., & Fuller, W. A., 1979). The number of lags is determined according to Akaike Information Criterion (Akaike, H., 1974):

$$\gamma_j^{\text{sub}} = \{\gamma_1 \dots \gamma_j\} \to \{y_1 \dots y_n\}$$

$$\Delta y_t = a + b y_{t-1} + \sum_{i=1}^{p} c_i \Delta y_{t-i} + \varepsilon_t$$

$$H_0 : b = 0, \text{(non-stationary, unit root present)}$$

$$H_a : b < 0, \text{(stationary)}$$

$$\tau_{st} = \frac{\hat{b}}{\text{SE}(\hat{b})} \sim T^{\text{DF}}(\text{constant only, n})$$

$$\text{AIC} \to \min_k \ln\left(\frac{\text{RSS}_\text{k}}{n}\right) + \frac{2\text{k}}{n}$$

The optimal number of lags is calculated one time, on the entire gamma sample to ensure the stability of all subtests. We create a chart with p-value of the ADF test on the y-axis and the size of the subsample (days since listing) on the x-axis. We will start with 25 days as the smallest subsample to ensure that the test has sufficient power.
AIC optimisation suggests setting lags = 0, i.e. use ordinary DF test

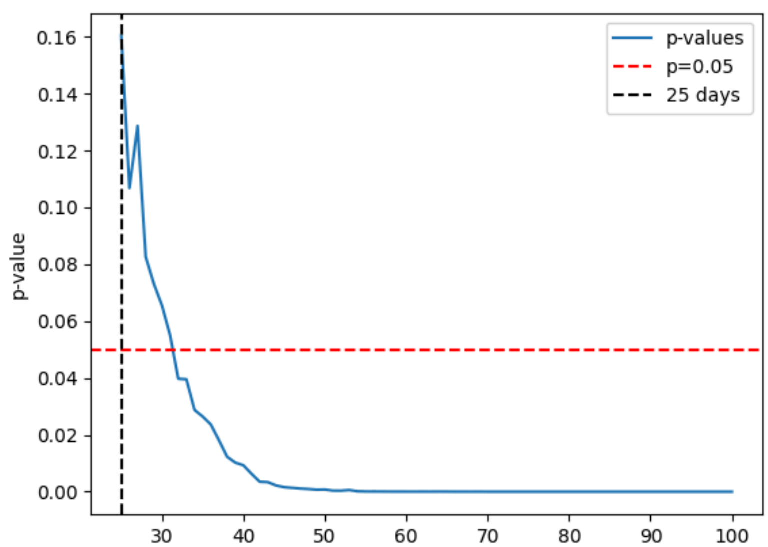

Figure 3: Relation between Dickey-Fuller p-value and the size of the mean time effects

As seen in Figure 3, initially the time series is not stationary but after 32 trading days p-values finally fall below the 5% mark and never rise above that level, indicating that gamma indeed becomes stationary over time.

3.1.2 Analysis of Variance

The variance will be modeled as the interaction term of time and asset effects as we want to make sure it never takes negative values. Additionally, at this moment we will not make any assumptions regarding the distribution of the disturbance term. The model for variance will be:

$$\begin{cases} Var(u_{it}) = k_i \times \sigma_t^2 \\ E(u_{it}) = 0 \end{cases}$$

$$\hat{u}_{it} = R_{it} - \hat{R}_{it}, E(\hat{u}_{it}) = 0 \rightarrow Var(\hat{u}_{it}) = E(\hat{u}_{it}^2)$$
$$\hat{u}_{it}^2 = \hat{k}_i \times \hat{\sigma}_t^2$$

Time and asset effects will be obtained via Iterated Feasible Generalized Least Squares (see Appendix B)

To measure the time effects on variance we will calculate the residuals and squared residuals from the panel regression:

i	t	resid	Squared resid
INCR	2	0.155114	0.024060
	3	-0.117558	0.013820
	4	-0.029419	0.000865
	5	0.004280	0.000018
	6	0.021571	0.000465
...	...	...	...
TEL	96	0.020093	0.000404
	97	0.018780	0.000353
	98	0.020007	0.000400
	99	-0.012748	0.000163
	100	0.007901	0.000062

541051 rows × 2 columns

Table 4: Residuals and squared residuals of the model; Long format

Running the IFGLS converges in 9 iterations, providing the estimates for k and sigma. Notably, min(k) = 7.03e-07, min(σ) = 0.813, so both values are positive which guarantees that all estimates for variance are positive as well.
Breusch-Pagan test results in R^2 for the auxiliary equation being effectively 0 and p-value being effectively 1.0 which signals that the transformed residuals $\hat{v}_{it} = \hat{u}_{it}/(\sqrt{\hat{k}_i} \times \hat{\sigma}_t)$ are homoskedastic at 5% thus the k-sigma decomposition is satisfactory.

We use the same plotting techniques to obtain the charts for raw sigma values (with kernel regression as the smoothing approach) and expanding-window average sigma values:

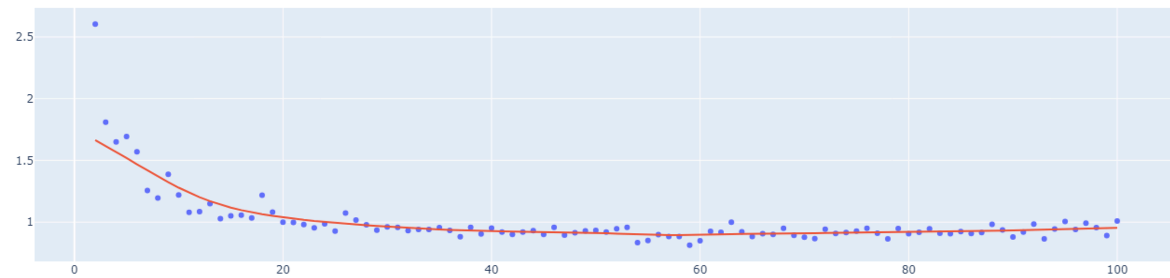

Figure 4: Variance time effects, 100-day dynamics

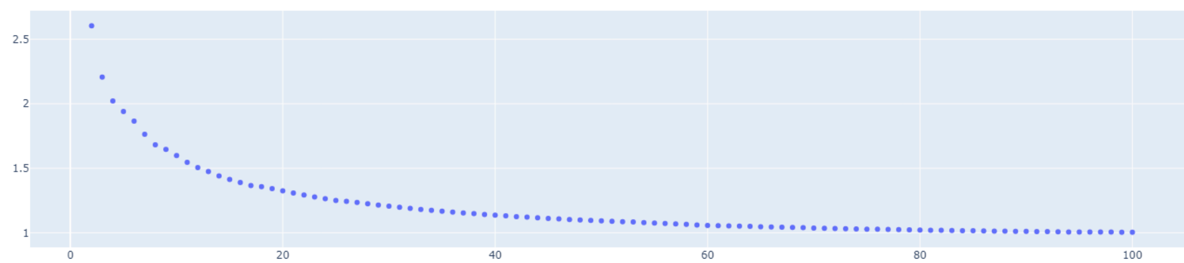

Figure 5: Rolling average of variance time effects, 100-day dynamics

As shown in Figures 4 and 5, there is a significantly elevated variance in the first 20 trading days which then smoothly decays to its long-term value as sigma approaches 1. Applying expanding ADF test (optimal lags = 8) yields the following chart:

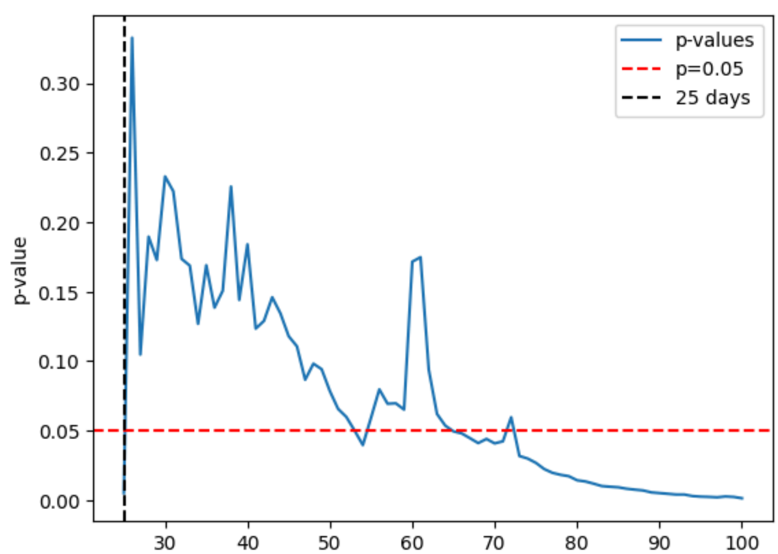

Figure 6: Relation between ADF p-value (lags = 8) and the size of the variance time effects

According to Figure 6, due to the higher number of lags, the decay of p-values is more volatile, nevertheless, the downward trend is evident. After the first 73 trading days the p-values settle below the 5% line and never rise above it again signaling long-term stationarity.

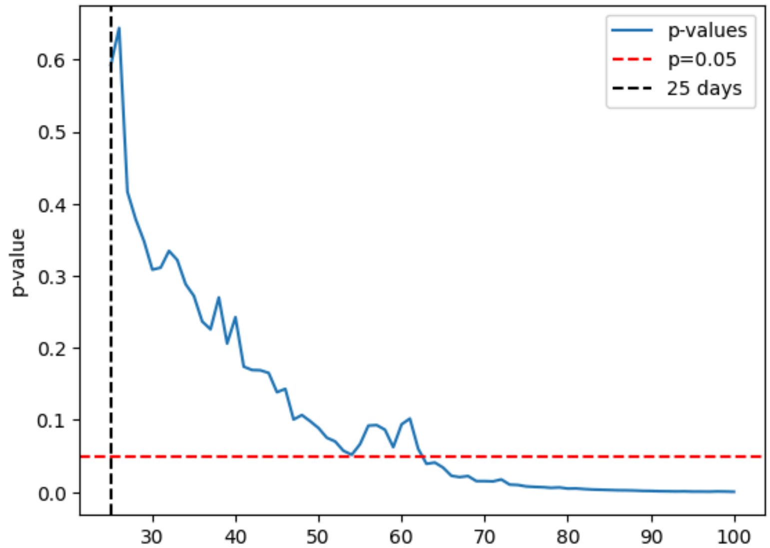

Figure 7: Relation between ADF p-value (lags = 5) and the size of the variance time effects sample

If a smaller number of lags (5) is used, this dynamic is even more explicit (see Figure 7). Here stationarity (at 5%) is achieved 63 days after listing. Both charts show very similar results which additionally proves the robustness of the test.

3.1.3 Joint Analysis of Mean and Variance

To test the overall significance of mean and variance estimates simultaneously we will utilize the Likelihood Ratio test (see Appendix E). For the purposes of this test it is required to approximate the distribution of the disturbance term. We will choose the best parametric distribution by using the Kolmogorov-Smirmov test (see Appendix D). We will first compare 2 most universal approximating distributions for residuals, namely:

$$\text{Student's t-distribution} : f(x) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\,\Gamma\left(\frac{\nu}{2}\right)}\left(1+\frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}$$

$$\text{Generalized normal distribution} : f(x) = \frac{\beta}{2\alpha\,\Gamma\left(1/\beta\right)}\exp\left(-\left|\frac{x-\mu}{\alpha}\right|^{\beta}\right)$$

Both of them allow for excess kurtosis modeling via parameterization which is essential to describe returns.
For each asset and distribution we first choose the best parameters via MLE and then run KS-test for the goodness-of-fit estimation. For GenNorm and T we separately calculate the proportion of assets for which p-value of KS-test is greater than 1%, i.e. for which these distributions are significantly good approximations. The results are 64.46% for GenNorm and 72.75% for Student's T. Thus we will use T as the distribution for the LR test. Further we filter out assets that follow this distribution , this results in a dataset with 3977 stocks and ETFs.

We want to test, for the whole subsample, the null hypothesis that time effects are not significant:

$$H_0 : R_{it} = \beta + \lambda_i + u_{it}; u_{it} \sim \text{Student}_\nu(0, k_i); \text{i.e } \gamma_t = 0 \text{ and } \sigma_t^2 = 1\ \forall$$
$$H_a : R_{it} = \beta + \lambda_i + \gamma_t + u_{it}; u_{it} \sim \text{Student}_\nu(0, k_i\sigma_t^2); \text{i.e } \exists t : \gamma_t \neq 0 \text{ or } \sigma_t^2 \neq 1$$

For each asset we separately calculate the best nu via MLE and sum up individual log-likelihoods.

The test results in a negative test statistic since coefficients were not obtained by maximizing likelihood function, thus not optimal in this context. However, with this number of parameters, a full panel MLE optimization will not be able converge in a reasonable time.
For this reason, we consider running LR test for each asset separately by fitting simple exponential decay functions for time effects of mean and variance to reduce the number of parameters, which will be:

$$R_{it} = \lambda_i + e^{-a_i t + b_i} + u_{it}, u_{it} \sim \text{Student}_\nu(0, k_i \times (1 + e^{-c_i t + d_i}))$$

Against a restricted model:

$$R_{it} = \lambda_i + u_{it}, u_{it} \sim \text{Student}_\nu(0, k_i)$$

The result of this procedure is summarised in the table that includes test statistics, individual p-values and estimates for restricted and unrestricted models:

asset	lr_stat	p_value	λ	a	b	ν_u	log k_u	c	d	λ_r	ν_r	log k_r
A	1.413110	0.841913	0.004281	37.812974	-15.508450	7.801912	-5.975782	1.051285	3.472967	0.004454	10.018690	-5.869764
AA	-1.523728	1.000000	0.000021	4.763827	-0.947284	10.028016	-8.545861	2.141299	-0.035018	0.000103	1174.824821	-8.379525
AACG	0.002738	0.999999	0.003277	3.804880	-1.078518	1.618389	-7.807804	1.833119	-0.049396	0.003273	1.618804	-7.807232
AADI	-0.000093	1.000000	-0.003347	8.254473	-2.603473	1.979226	-7.626372	3.593167	-0.212274	-0.003347	1.979255	-7.626349
AAGR	NaN	NaN	-0.000350	8.613850	-1.924524	10.023844	-11.553722	2.690420	0.003383	0.000232	2.049588	-12.806956
...	...	...	...	...	...	...	...	...	...	...	...	...
ZVIA	0.027181	0.999908	-0.004029	56.617934	42.744877	947.423475	-5.920097	817.725933	-322.632129	-0.004045	104.237232	-5.933426
ZVRA	-0.000352	1.000000	0.002155	6.185191	-1.602546	3.476102	-7.248624	3.749721	-0.341282	0.002155	3.476074	-7.248618
ZWS	9.773997	0.044412	-0.001634	1.002368	-0.837341	3.644264	-8.479276	3.190059	-0.062735	-0.001124	3.503393	-8.404848
ZYME	0.366069	0.985159	-0.008275	1.857389	-0.410827	3.471867	-7.167006	2.470958	-0.316025	-0.007876	3.610517	-7.139016
ZYXI	-0.900568	1.000000	-0.007676	1.810251	-0.215180	9.720052	-6.269915	1.670418	-0.128752	-0.005154	4.200663	-6.540949

3977 rows × 12 columns

Table 5: Results of the Likelihood Ratio test: estimated parameters, test statistics and p-values

According to the data from Table 5, 7.92% of assets of this subsample (315) have significant joint time effects (p-value < 0.05). This shows that even though the proportion is not too large, sometimes time effects can act as an additional estimator for mean and variance of returns.

3.2 Cross-Validation

3.2.1 Validation of the Mean Predictions

At this part of the paper we want to assess the quality of our model on the data that were not used for the fit. For this purpose we will shuffle our dataset asset-wise 100 times and each time split it into train and test subsamples (90% and 10% respectively), this will ensure that the outcome of the tests will be less affected by the particular arrangement of data in each split.

First we perform pooled Diebold-Mariano test for the predictions of mean (see Appendix F), where on the test subsample our models are:

Unrestricted model: $R_t = \beta_1 + \gamma_t + u_{it}$
Restricted model: $R_t = \beta_2 + v_{it}$

To consistently estimate standard deviation of $\hat{L}_t$ and to calculate a consistent statistic for this test will use Newey–West estimator which accounts for possible autocorrelation and heteroskedasticity (Newey, W. K., & West, K. D.,1987):

$$\hat{\sigma}^2_{\hat{L}} = \sum_{j=-K}^{K} \left(1 - \frac{|j|}{K+1}\right) \hat{\gamma}(j)$$

$$\hat{\gamma}(j) = \frac{1}{T} \sum_{t=|j|+1}^{T} \tilde{L}_t \tilde{L}_{t-|j|}$$

Where K is the maximum lag length
For negative lags (j < 0), we set $\hat{\gamma}(j) = \hat{\gamma}(-j)$ to ensure symmetry.

$$\hat{\sigma}_{\hat{L}} = \sqrt{\hat{\sigma}^2_{\hat{L}}}$$

Out of 100 subsamples, 80% demonstrated a positive statistic i.e. loss of the restricted model was higher so the unrestricted model performed better. However, only in 12 subsamples showed p-value < 5% at this test. After pooling the results and performing Nadeau-Bengio t-test (see Appendix G), we get a pooled p-value = 0.0215, which states that on average, for the predictions of mean return the unrestricted model performs better than the restricted (constant) model with ~97.8% confidence.

3.2.2 Validation of the Variance Predictions

To assess the performance of the variance model, we want to:

1. Estimate $\hat{u}_{it} = R_{it} - \hat{\beta} - \hat{\lambda}_i - \hat{\gamma}_t$ for the train subsample
2. Estimate $\hat{v}_{jt} = R_{jt} - \hat{\beta} - \hat{\gamma}_t$ for the test subsample
3. Based on $\hat{u}^2_{it}$ estimate $\hat{k}_i$ and $\hat{\sigma}^2_t$
4. Assess the predictive power of the unrestricted model: $Var(u_{it}) = k_i \times \sigma^2_t$ compared to the restricted model: $Var(u_{it}) = k_i$, based on the test subsample (where k_i is unknown) and thus $\widehat{Var}(v_{jt}) = \hat{k} \times \hat{\sigma}^2_t$ against $\widehat{Var}(v_{jt}) = \hat{k}$

For the restricted model we estimate:

$$\hat{k} = \bar{v}^2_{\cdot\cdot}$$

For the unrestricted model we estimate:

$$\hat{k} = \frac{1}{N_{test}T} \sum_{j=1}^{N_{test}} \sum_{t=1}^{T} \frac{v^2_{jt}}{\hat{\sigma}^2_t}$$

Performing Diebold-Mariano test (see Appendix F) results in 98% of positive J_{st} (loss of the unrestricted model is smaller), however only 6% of subsample p-values were below 0.05. Nadeau-Bengio t-test (see Appendix G) results in effectively zero p-value, which confirms that for the predictions of variance of returns the unrestricted model performs better than the restricted (constant) model with almost 100% confidence.

4. Conclusion

This paper has documented the temporal evolution of mean returns and return variance for U.S. stocks and ETFs over their first 100 trading days, using a fixed‑effects panel framework with heteroskedasticity‑robust inference. Empirically, we observe a pronounced positive return on day 1, followed by systematically negative returns for days 2–8, after which time effects stabilize around zero, overall time effects are significant at 1% according to HC3‑adjusted RSS F‑test; an expanding‑window ADF test rejects non‑stationarity of time-effects of mean returns only once $t \geq 32$ ($p\text{-value} < 0.05$). Volatility is elevated through roughly day 20 before decaying toward its long‑run level, with stationarity attained only after day 73 ($p\text{-value} < 0.05$) under an expanding ADF (lags = 8) . The obtained variance time-effects are significant at 1% for eliminating heteroskedasticity according to Breusch–Pagan test. At the same time only ~7.9% of individual assets exhibit significant joint time‑effects under an exponential‑decay specification ($p\text{-value} < 0.05$)

Out‑of‑sample validation via pooled Diebold–Mariano tests indicates that models incorporating time-effects achieve lower mean‑forecast loss in 80% of splits (pooled Nadeau–Bengio $p\text{-value} \approx 0.022$), and lower variance‑forecast loss in 98% of splits (pooled Nadeau–Bengio $p\text{-value} \approx 0$)

Overall, we find clear evidence of persistent time effects in both mean returns and return variance during the initial trading period of new listings. These effects dissipate over the first few months, confirming the existence of a definable “settling window.” Portfolio managers can harness this knowledge by incorporating time-varying adjustments into entry timing and risk forecasts: delaying full allocation until the bulk of the settling period has passed, scaling position sizes to account for elevated early volatility, and dynamically updating volatility forecasts to reflect decaying variance effects. By explicitly modelling these time effects, practitioners can achieve more accurate return expectations and more robust risk management for newly issued assets.

The current analysis relies on simplified exponential‑decay functions and sequential estimation for simplicity and transparency. A full panel MLE that jointly estimates mean–variance convergence remains a promising extension, as does broadening the sample to international listings or alternative asset classes (e.g. cryptocurrencies). Further work could also incorporate firm‑specific covariates (sector, market‑cap, listing mechanism) to explain heterogeneity in convergence speed, thereby refining both theoretical models of information diffusion and practical forecasting of young‑asset dynamics.

5. List of References

1. LOWERY, M., OFFICER, M. S., & SCHWERT, G. W. (2010). The Variability of IPO Initial Returns. *The Journal of Finance*, *65*(2), 425–465.
2. Bruce, A. and Thilakaratne, P. (2014) IPO Stocks Performance Imperfection: A Review of Models and Empirical Works. *American Journal of Industrial and Business Management*, 155-166.
3. Rock, K. (1986). Why new issues are underpriced. *Journal of Financial Economics, 15*(1), 187-212.
4. Benveniste, L. M., & Spindt, P. A. (1989). How investment bankers determine the offer price and allocation of new issues. *Journal of Financial Economics, 24*(2), 343-361.
5. Beneda, N., & Zhang, Y. (2009). Heterogeneous relationship between IPO return and risk across idiosyncratic variance characteristics. *The Quarterly Review of Economics and Finance, 49*(4), 1298-1316
6. MacKinnon, J. G. & White, H. (1985). "Some Heteroskedasticity‐Consistent Covariance Matrix Estimators with Improved Finite Sample Properties." *Journal of Econometrics*, 29(3), 305–325.
7. P. A. V. B. Swamy, & Arora, S. S. (1972). The Exact Finite Sample Properties of the Estimators of Coefficients in the Error Components Regression Models. *Econometrica*, *40*(2), 261–275.
8. Breusch, T. S. & Pagan, A. R. (1979). "A Simple Test for Heteroscedasticity and Random Coefficient Variation." *Econometrica*, 47(5), 1287–1294
9. Kolmogorov, A. N. (1933). "Sulla determinazione empirica di una legge di distribuzione." *Giornale dell'Istituto Italiano degli Attuari*, 4, 83–91.
10. Stanley, E. (2024). US Stock Market History Data. *Kaggle*. Retrieved April 20, 2025, from https://www.kaggle.com/datasets/ericstanley/us-stock-market-history-data-csv
11. Dickey, D. A., & Fuller, W. A. (1979). "Distribution of the Estimators for Autoregressive Time Series with a Unit Root." *Journal of the American Statistical Association*, 74(366), 427–431
12. Akaike, H. (1974). "A New Look at the Statistical Model Identification." *IEEE Transactions on Automatic Control*, 19(6), 716–723.
13. Allan Timmermann & Yinchu Zhu (2019). *Comparing Forecasting Performance with Panel Data*. SSRN Working Paper
14. Newey, W. K., & West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3), 703–708
15. Nadeau, C., & Bengio, Y. (2003). *Inference for the Generalization Error*. Advances in Neural Information Processing Systems, 16, 307–313.

6. Appendices

A. Test on the joint significance of time effects, HC3 standard errors (MacKinnon, J. G. & White, H., 1985)

$$H_0 : \gamma_1 = \gamma_2 = \cdots = \gamma_T = 0$$
$$H_a : \exists \gamma_t \neq 0$$

$$F_{st} = \frac{(RSS_r - RSS_u)/(df_r - df_u)}{s^2_{\text{HC3}}} \sim F(T, n-p)$$

$$\hat{u}_{it} = R_{it} - \hat{R}_{it}, \quad h_{it} = \text{leverage for observation } (i,t)$$
$$\hat{u}_{it,\text{HC3}} = \frac{\hat{u}_{it}}{1 - h_{it}}$$
$$s^2_{\text{HC3}} = \frac{1}{n-p} \sum_{i,t} \left(\hat{u}_{it,\text{HC3}}\right)^2$$

$$\hat{R} = HR$$
$$H = X(X'X)^{-1}X \rightarrow h_{it} = x_{it}(X'X)^{-1}x_{it}$$

B. Iterated Feasible Generalized Least Squares (P. A. V. B. Swamy, & Arora, S. S., 1972)

IFGLS procedure implies sequentially separating and refining asset and time effects until the estimates converge. This approach ensures that both effects and final estimates are non-negative as required for the variance modeling:

1. $\hat{\sigma}^2_t = \frac{1}{N} \sum_{i=1}^{N} \hat{u}^2_{it},$
2. $\hat{k}_i = \frac{1}{T} \sum_{t=1}^{T} \frac{\hat{u}^2_{it}}{\hat{\sigma}^2_t},$
3. $\hat{\sigma}^2_t = \frac{1}{N} \sum_{i=1}^{N} \frac{\hat{u}^2_{it}}{\hat{k}_i},$
4. Repeat steps 2-3 until $\max\left|\hat{k}_i^{(r)} - \hat{k}_i^{(r-1)}\right|, \max\left|\hat{\sigma}_t^{2\,(r)} - \hat{\sigma}_t^{2\,(r-1)}\right| < \varepsilon.$

C. Breusch-Pagan test with dummy variables (Breusch, T. S. & Pagan, A. R., 1979)

To test the significance of the functional form of variance, we will assess if the sample of error terms compensated for time and asset effects is homoskedastic

$$Var(u_{it}) = k_i \times \sigma_t^2 \Rightarrow Var(\frac{u_{it}}{\sqrt{k_i} \times \sigma_t}) = 1 \Rightarrow E(\frac{u_{it}^2}{k_i \times \sigma_t^2}) = 1$$

$$\frac{\hat{u}_{it}}{\sqrt{\hat{k}_i} \times \hat{\sigma}_t} = \hat{v}_{it} \Rightarrow \frac{\hat{u}_{it}^2}{\hat{k}_i \times \hat{\sigma}_t^2} = \hat{v}_{it}^2$$

$$\hat{v}_{it}^2 = \alpha + \sum_{i=1}^{N-1} \delta_i \times D_i^{Asset} + \sum_{t=1}^{T-1} \theta_t \times D_t^{Time} + \varepsilon_{it},$$

$$H_0 : \text{homoskedasticity}; \delta_i = 0 \text{ and } \theta_t = 0 \ \forall i, t$$
$$H_a : \text{heteroskedasticity}; \exists \delta_i \neq 0 \text{ or } \theta_t \neq 0$$

$$\chi_{st}^2 = n_{\text{aux}} \times R_{\text{aux}}^2 \sim \chi^2(N + T - 2)$$

D. Kolmogorov-Smirmov test (Kolmogorov, A. N., 1933)

$$H_0 : F(x) = F_0(x), \quad \forall x$$
$$H_a : F(x) \neq F_0(x) \text{ for some } x$$

$$D_{st} = \sup_{x \in \mathbb{R}} \left| F_n(x) - F_0(x) \right|$$

D_{st} measures the maximum vertical deviation between F_n and F_0

E. Likelihood Ratio test

Likelihood Ratio test will compare the complete unrestricted model specification with the restricted one based on the chosen approximating distribution:

$$H_0 : \text{Time Effects} = 0 \ \forall t$$
$$H_a : \text{Time Effects} \neq 0 \text{ for some } t$$

$$\Lambda_{st} = -2\left[\ell_R - \ell_{UR} \right] \sim \chi^2(\# \text{ Time Parameters})$$

F. Modified Diebold-Mariano test Allan (Timmermann & Yinchu Zhu, 2019)

$$H_0 : \mathbb{E}[L_{m_1}] = \mathbb{E}[L_{m_2}]$$
$$H_a : \mathbb{E}[L_{m_1}] > \mathbb{E}[L_{m_2}]$$

where:
- m_2 is the unrestricted model
- m_1 is the restricted model

Let L_m denote the pooled average squared-error loss for forecasting model m, computed over N cross-sectional units and T time periods:

$$L_m = \frac{1}{NT} \sum_{t=1}^{T} \sum_{i=1}^{N} \left(y_{i,t} - \hat{y}_{i,t}^{(m)} \right)^2$$

Define the loss differential for unit i at time t as:

$$\Delta L_{i,t} = \left(y_{i,t} - \hat{y}_{i,t}^{(m_1)} \right)^2 - \left(y_{i,t} - \hat{y}_{i,t}^{(m_2)} \right)^2$$

Compute the cross-sectional average loss differential at each time t:

$$\Delta L_t = \frac{1}{N} \sum_{i=1}^{N} \Delta L_{i,t}$$
$$\hat{L}_t = \sqrt{N} \cdot \Delta L_t$$
$$J_{st} = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \frac{\hat{L}_t - \bar{\hat{L}}}{\hat{\sigma}_{\hat{L}}} \sim \mathcal{N}(0,1)$$

G. Modified NB t-test (Nadeau, C., & Bengio, Y., 2003)

This test accounts for the correlation between test subsamples taken from one general sample. It adjusts the variance of test statistics J calculated on each random subsample.

$$H_0 : \mathbb{E}[\bar{J}] = 0$$
$$H_a : \mathbb{E}[\bar{J}] > 0$$

$$\widehat{\mathrm{Var}}(\bar{J}) = (\frac{1}{n} + \frac{\rho}{1-\rho}) \times \widehat{\mathrm{Var}}(J_i)$$

$$t_{st} = \frac{\bar{J}}{\sqrt{\widehat{\mathrm{Var}}(\bar{J})}} \sim t_{n-1}$$

$$\rho \approx \text{test size}$$