The Puzzle of Index Option Returns

Minhao Li Yun Hu Zean Li Zilin Song

The Nature of Replication

Beyond standard linear factor methodology in explaining SP 500 Index call and put option returns, this academic paper hypothesized several crisis-related factors, including price jumps, volatility jumps, and liquidity to explain the cross-sectional variation in returns. In conclusion, their hypothesis was not rejected, and it has shown that the alphas of short-maturity out-of-money puts become economically and statistically insignificant.

In our replication project, we automatically downloaded option data from WRDS, and then filtered out certain rows based on the criteria given in the paper. This data filtering process contains 3 levels of filters, and each level contains three to five subfilters. As a result, we replicated the numbers in Table B.1 quite successfully (with roughly 0.93 million rows in total and only 30 thousands rows in difference). After we replicated the Table B.1, we used the resulting DataFrame to replicate Table 2. In the end, we were also able to reproduce the numbers close enough.

We also automated the whole workflow by leveraging doit, which utilize the dodo.py file to replicate our tables without manual works.

How the project went: success, challenges, data sources

We found our success in replicating the numbers closely while at the same time utilizing some knowledge in option pricing. For example, when writing the last two filters, we need to first write a program to calculate the implied volatility and the implied market interest rate, and fit the calculated implied volatility to a normal distribution. Then, remove the samples that deviate too much. Calculating the implied volatility requires solving the function inversely through the pricing formula, and it is necessary to avoid numerous issues such as data missing, data mismatch, and problems with cleaning the original data. This not only uses a lot of statistical knowledge but also requires analysis combining option pricing theory.

We also found challenges when teammates are not familiar with what each of our code does. Specifically, when one of our teammates was working on the dodo file, he was unable to know which files to use, and what is the correct

order to process, and what are the inputs and outputs of these py files. Hence it is also valuable to enhance communications betweeen group members.

The data source we used comes from OptionMetrics module in WRDS, which is a comprehensive database of historical option price, underlying security information, implied volatility, and sensitivity information for the entire US listed and equity options markets from 1996 till present.

Table 1: Summary Statistics for Calls and Puts, April 1986 to January 2021 (Table 2 in Paper)

	Calls	CallPercentage	Puts	PutPercentage
Found	342519	0.9961	344036	0.9958
Missing	724	0.0021	783	0.0023
Expired	619	0.0018	669	0.0019

Table 2: Filters (Table B.1 in Paper)

Filter	Level	Deleted	Remaining
Starting	Calls		1704128
	Puts		1706268
All			3410580
Level 1 filters	CleanOptionsData	10	3410570
	RemoveDuplicateQuotes	0	3410570
	DeleteZeroBidFilter	272078	3138492
Level 2 filters	DaystoMaturityFilter	1311128	1827364
	ExtremeIVFilter	202133	1625231
	Moneyness(0.8 1.2)Filter	465052	1160179
	ImpliedInterestRateFilter	177246	982933
Level 3 filters	IVFilter	4096	978837
	PutCallParityFilter	46138	932699

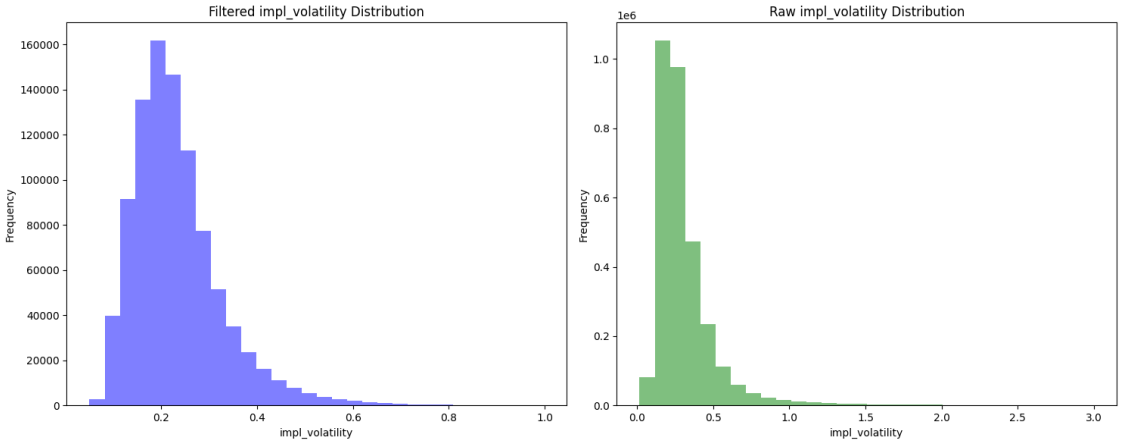

Figure1. Implied volatility distribution change after filtering.
We remove all options with fewer than seven or more than 180 calendar days to expiration. The short maturity options tend to move erratically close to expiration and the long maturity options lack volume and open interest.

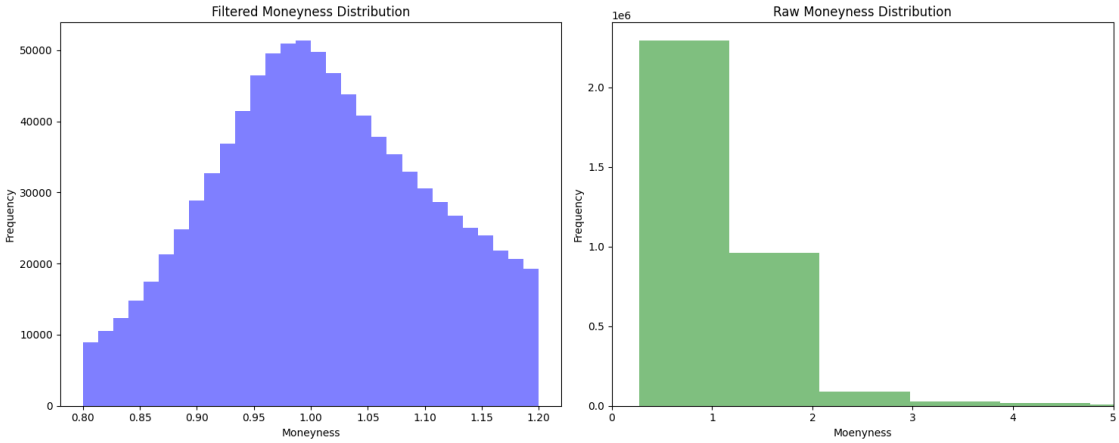

Figure1. Implied volatility distribution change after filtering.
We remove all option quotes with moneyness, the ratio of strike price to index price, below 0.8 or above 1.2. These options have little value beyond their intrinsic value and are also very thinly traded.

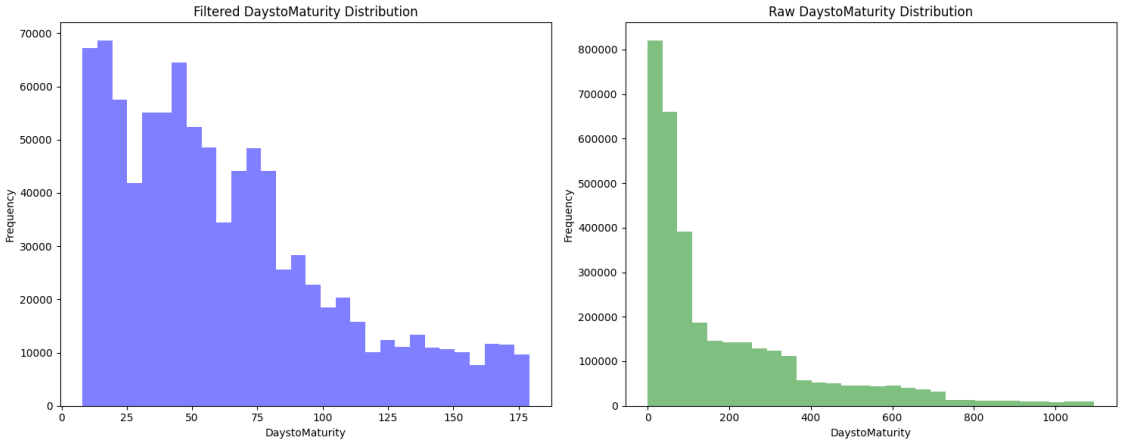

Figure1. Implied volatility distribution change after filtering.
We remove all options with fewer than seven or more than 180 calendar days to expiration. The short maturity options tend to move erratically close to expiration and the long maturity options lack volume and open interest.