The Puzzle of Filtering Index Options

UChicago WI 24: FINN 329[*]

Viren Desai, Harrison Holt, Ian Hammock

March 4, 2024

Abstract

In this article we will summarize our efforts to replicate the filtering described in appendix B of *The Puzzle of Index Option Returns* by Constantinides, Jackwerth, and Savov (2013). We provide additional insight on how these filters shape the distribution on implied volatility and moneyness. Moreover, due to the unavailability of index option data from 1985 to 1995, we focus our comparison on the dataset of 1/1996-1/2012 as well as extending this analysis forward from 2/2012 to 12/2019. Our analysis can be readily found on our Github [1].

---

[1]https://github.com/harrypandas/finm-32900_final_project.git

1 Replicating Table B1

In the appendix B of Constantinides, Jackwerth, and Savov (2013), three levels of filters are described with the intent to minimize quoting errors in the construction of their portfolios. In this section we will summarize our implementation and briefly discuss the differences. Our results are summarized in Table 1.

1.1 Level 1

This application of filters is fairly straight forward. However, an unexplainable difference occurs upon the application of the Volume = 0 filter. In Table B1 of Constantinides, Jackwerth, and Savov (2013), no options have a Volume = 0 in their dataset. However, we observe 2,093,744 options with a Volume = 0. Unfortunately, no more details are given in the manuscript describing this step. In order to not diverge from their data pool we choose to drop 0 options, this is reflected in Table 1.

1.2 Level 2

1.3 Level 3

Table 1: Table B.1

		OptionMetrics: 1996-01 to 2012-01		OptionMetrics:2012-02 to 2019-12		Total	
		Deleted	Remaining	Deleted	Remaining	Deleted	Remaining
Starting	Calls		1,704,220		7,901,901		9,606,121
	Puts		1,706,360		7,901,427		9,607,787
	All		3,410,580		15,803,328		19,213,908
Level 1 filters	Identical	0		277,102		277,102	
	Identical except price	10		2,557,330		2,557,340	
	Bid = 0	272,078		1,069,116		1,341,194	
	All		3,138,492		11,899,780		15,038,272
Level 2 filters	Days to expiration <7 or >180	1,297,729		3,080,910		4,378,639	
	IV <5% or >100%	16,432		63,639		80,071	
	K/S <0.8 or >1.2	550,227		1,987,486		2,537,713	
	Implied interest rate < 0	642,940		2,053,938		2,696,878	
	Unable to compute IV	37,733		385,913		423,646	
	All		593,431		4,327,894		4,921,325
Level 3 filters	IV filter	38,568		312,899		351,467	
	Put-call parity filter	92,973		850,793		943,766	
	All		461,890		3,164,202		3,626,092

Number of observations that are removed upon application of appendix B filters.
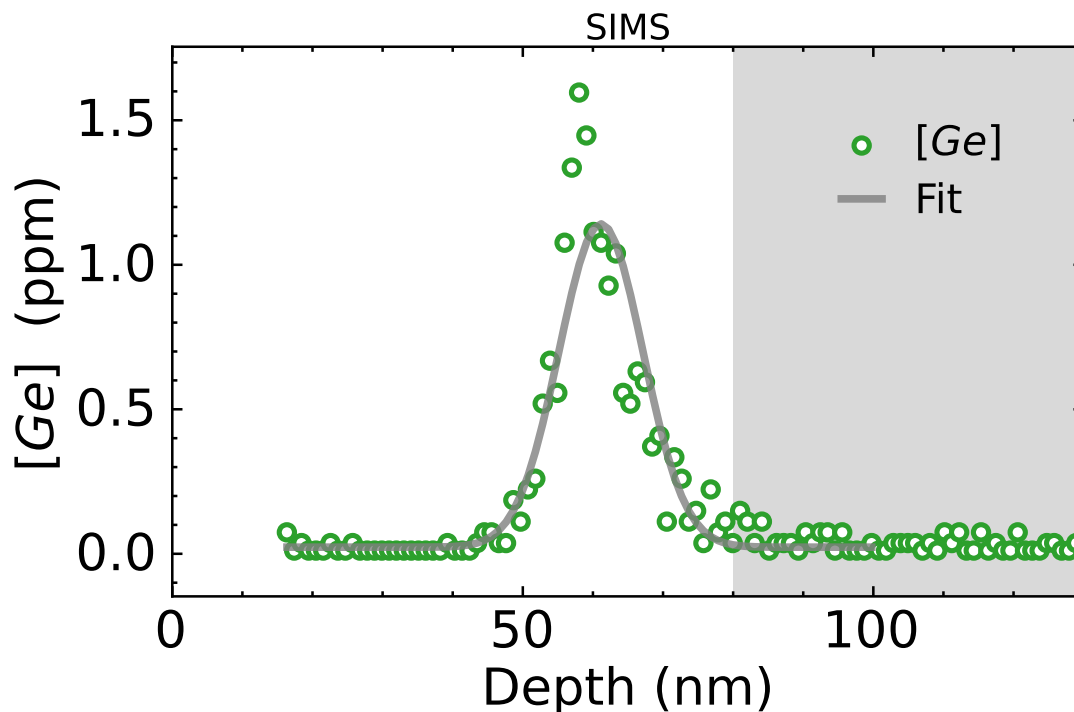
2   B Filters, Implied Volatility, Moneyness

Figure 1: Your caption here

3   Replicating Table2

Table 2 describes how many options are found, go missing, or expire in the dataset. An option is found if it reappears the next trading day. An option is missing for if it does not reappear the next trading day. Multiple days missing, counts as multiple options missing. Lastly, if an option is lost and expires this is noted as expired.

We would like to note an interesting aspect of this dataset. Over 80% of the options expire on a Saturday or a non-trading day. To handle this, we push the expiration day to the nearest Friday, presumably the nearest trading day. However, there are quite a few edge cases which would explain the discrepancy between our analysis and Constantinides, Jackwerth, and Savov (2013). Further investigation is required.

Table 2: Table 2 Sample

Observations	Calls				Puts			
	1996-01 to 2012-01		2012-02 to 2019-12		1996-01 to 2012-01		2012-02 to 2019-12	
All trading days								
Found	176,225	79%	1,185,069	76%	176,225	79%	1,185,069	76%
Missing	6,867	3%	21,026	1%	6,867	3%	21,026	1%
Expired	40,298	18%	352,870	23%	40,298	18%	352,870	23%
Last trading day of the month								
Found	19,126	82%	280,486	80%	19,126	82%	280,486	80%
Interpolated	4,104	18%	69,727	20%	4,104	18%	69,727	20%

Tracking the instances options are found, missing or expired.

4 Data

Our option data is queried from OptionMetrics provided by Wharton Research Data Services (WRDS). We limit the query to SECID = 108105, S&P 500 Index - SPX. We use the three month Tbill as our interest rate, this is from the Federal Reserve Board's H15 report supplied by WRDS.

In comparison to their data, we have pulled 184 more options than them. It is unclear where the discrepancy lies. We assumed we were off by a day however this will truncate or elongate the dataset by over 300 points. We credit the discrepancy to OptionMetrics updating their data to be more accurate.

The following links contain the documentation and helpful links for the WRDS database:

- Option Metrics Overview

- Option Metric Keys

- Option Metrics Query

- Federal Reserve Report

5 References

Constantinides, George M., Jens Carsten Jackwerth, and Alexi Savov. 2013. "The Puzzle of Index Option Returns." *The Review of Asset Pricing Studies* 3 (2):229–257. URL https://doi.org/10.1093/rapstu/rat004.