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 1996-01 to 2012-01 as well as extending this analysis forward from 2012-01 to 2019-12. Our analysis can be readily found on our Github [1].

*Taught by Jeremy Bejarano, jbejarano@uchicago.edu

[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. Furthe discussion relating to the distribution of implied volatility and moneyness through the level three filter can be found in the Appendix A.

1.1 Level 1

Two issues arose when applying this set of filters. Firstly, there are quite a few options in the "Identical but Price" filter that have no implied volatility. This will be taken care in level two, but knowledge of the implied volatility is required to select the option's price that has volatility closest to the "at the money" volatility. This is limited to about 5 in the 1996-2012 dataset, and most of these options will fall out in the level 2 filters (days to expiration <7 or out of the money. In the 2012 to 2019 dataset we have observed around 355,896 options with no implied volatility. The distribution of expiration time and moneyness peak is given in. For the purpose of this filter, if an implied volatility is not given it will not be choosen as the option with volatility closest to the at the money,as we will be unable to calculate it and it will fall out in level 2. Additionally, if an option family's "at the money" member cannot have their implied volatilitiy calculated all options will be inevitably dropped due to the "Unable to compute IV" filter.

Secondly, 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:2012-01 to 2019-12		Total	
		Deleted	Remaining	Deleted	Remaining	Deleted	Remaining
Starting	Calls		1,704,220		7,922,884		9,627,104
	Puts		1,706,360		7,922,410		9,628,770
	All		3,410,580		15,845,294		19,255,874
Level 1 filters	Identical	0		277,102		277,102	
	Identical except price	10		2,557,330		2,557,340	
	Bid = 0	272,078		1,072,421		1,344,499	
	All		3,138,492		11,938,441		15,076,933
Level 2 filters	Days to expiration <7 or >180	1,297,729		3,096,182		4,393,911	
	IV <5% or >100%	16,432		63,952		80,384	
	K/S <0.8 or >1.2	550,227		1,997,234		2,547,461	
	Implied interest rate < 0	592,726		4,434,332		5,027,058	
	Unable to compute IV	38,434		207,340		245,774	
	All		642,944		2,139,401		2,782,345
Level 3 filters	IV filter	33,706		133,029		166,735	
	Put-call parity filter	61,212		372,658		433,870	
	All		548,026		1,633,714		2,181,740

Number of observations that are removed upon application of appendix B filters.

2 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

	Calls				Puts			
Observations	1996-01 to 2012-01		2012-01 to 2019-12		1996-01 to 2012-01		2012-01 to 2019-12	
All trading days								
Found	222,525	81%	527,594	69%	222,525	81%	527,594	69%
Missing	4,058	1%	6,371	1%	4,058	1%	6,371	1%
Expired	47,355	17%	232,476	30%	47,355	17%	232,476	30%
Last trading day of the month								
Found	7,170	82%	210,826	82%	7,170	82%	210,826	82%
Interpolated	1,538	18%	45,943	18%	1,538	18%	45,943	18%

Tracking the instances options are found, missing or expired.

3 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

4 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.

A Level Three Filter

A.1 1996-01 to 2012-01

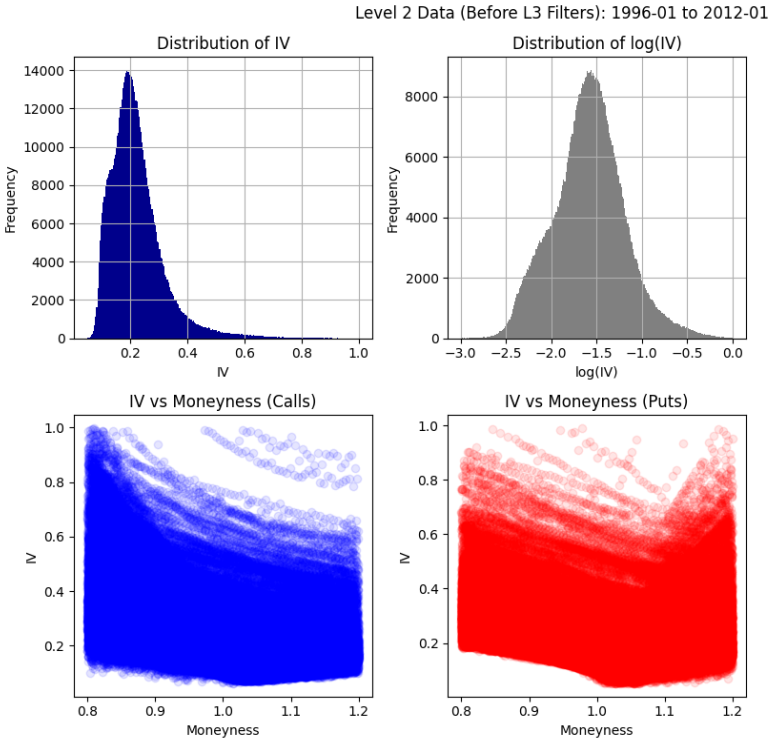

Figure 1: Your caption here

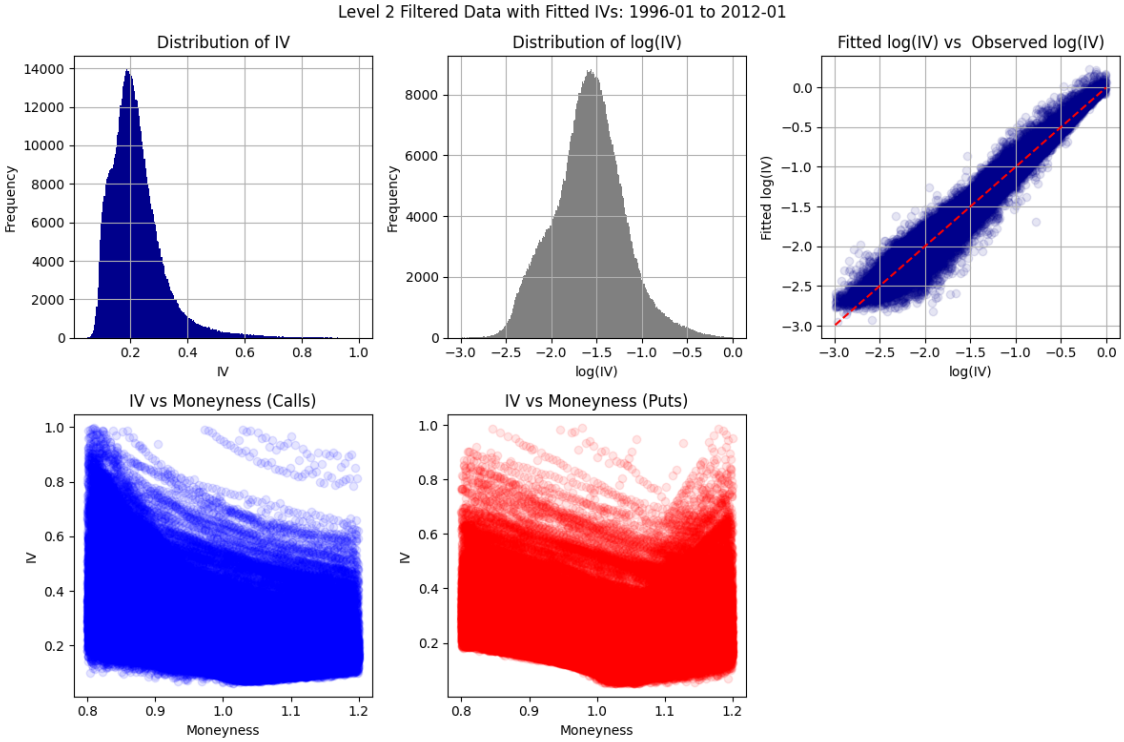

Figure 2: Your caption here

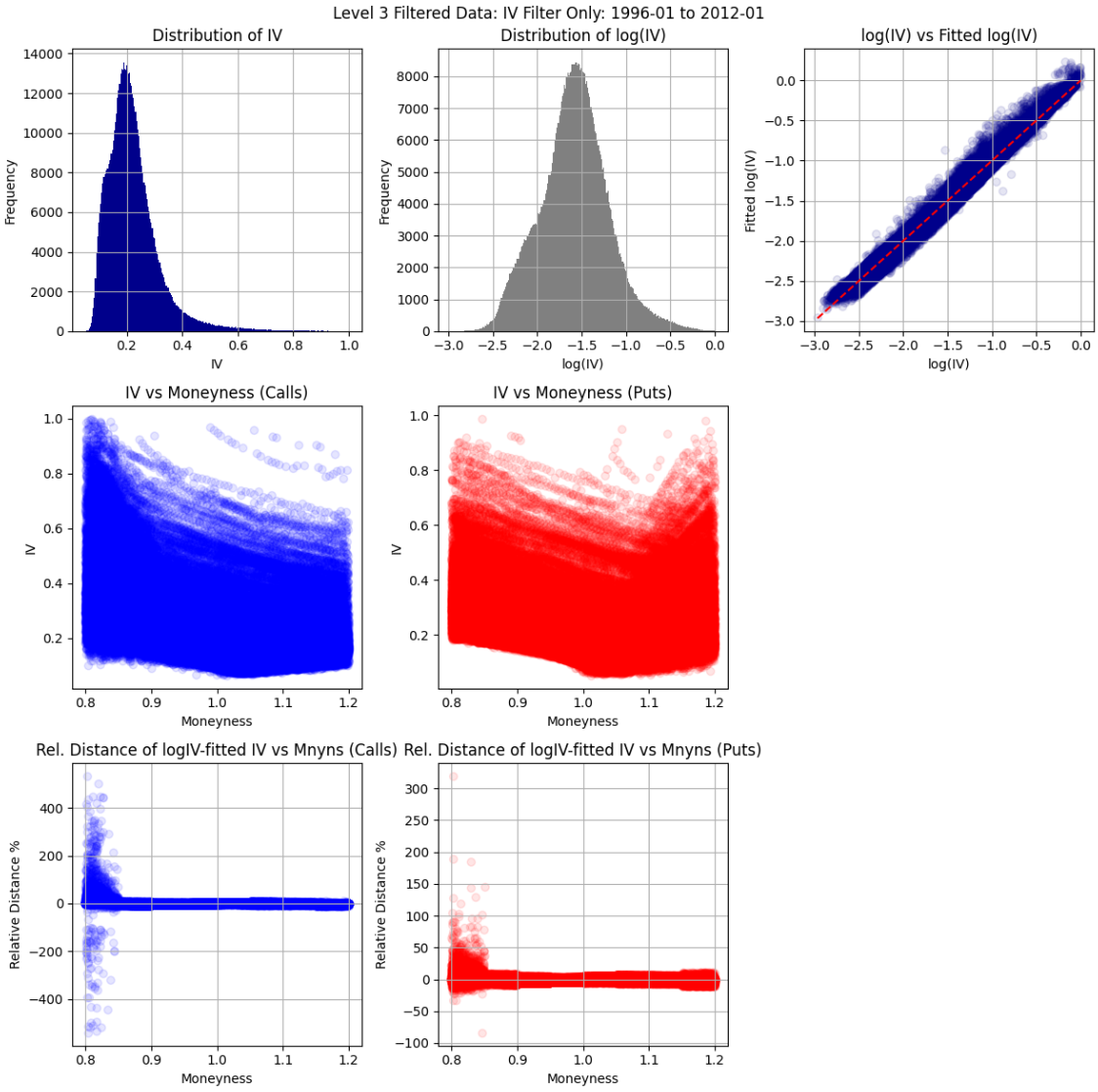

Figure 3: Your caption here

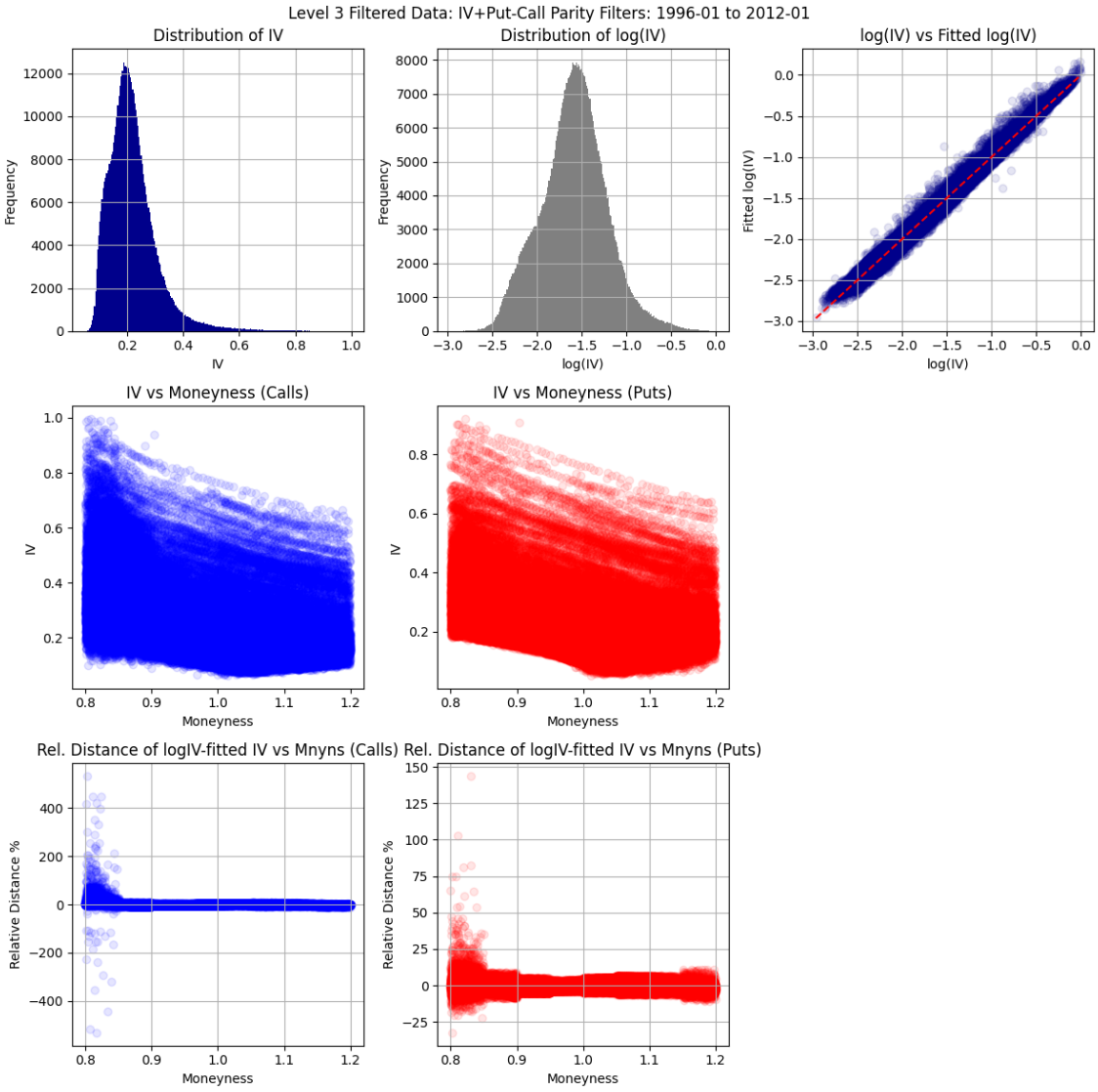

Figure 4: Your caption here

A.2 2012-01 to 2019-12

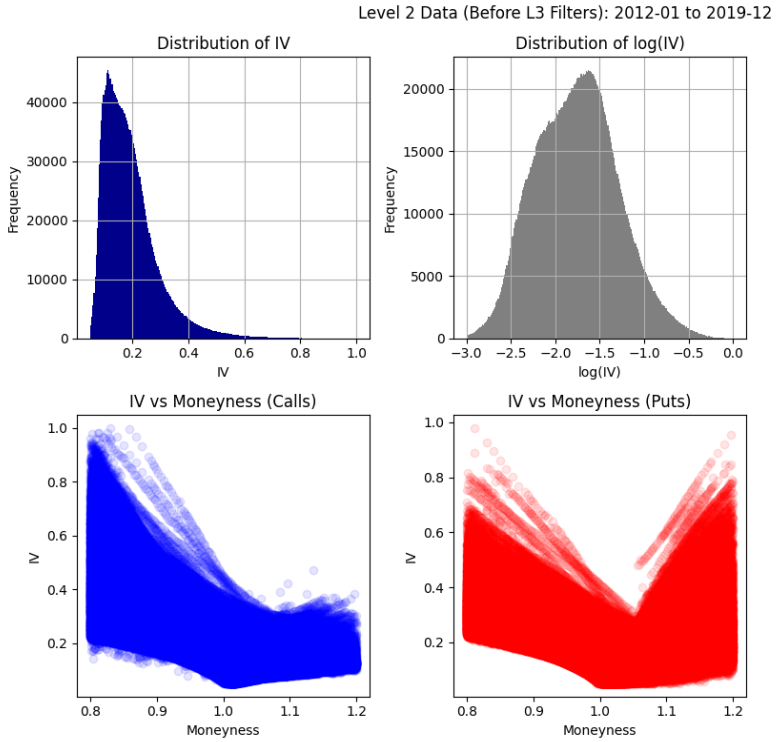

Figure 5: Your caption here

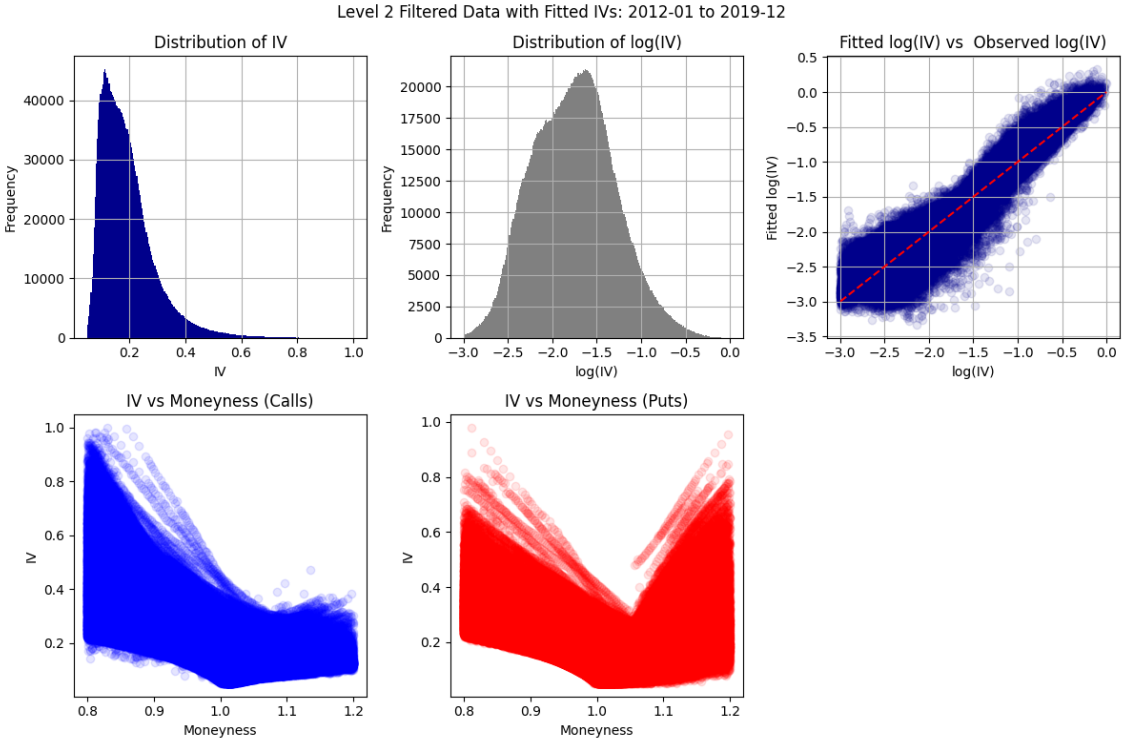

Figure 6: Your caption here

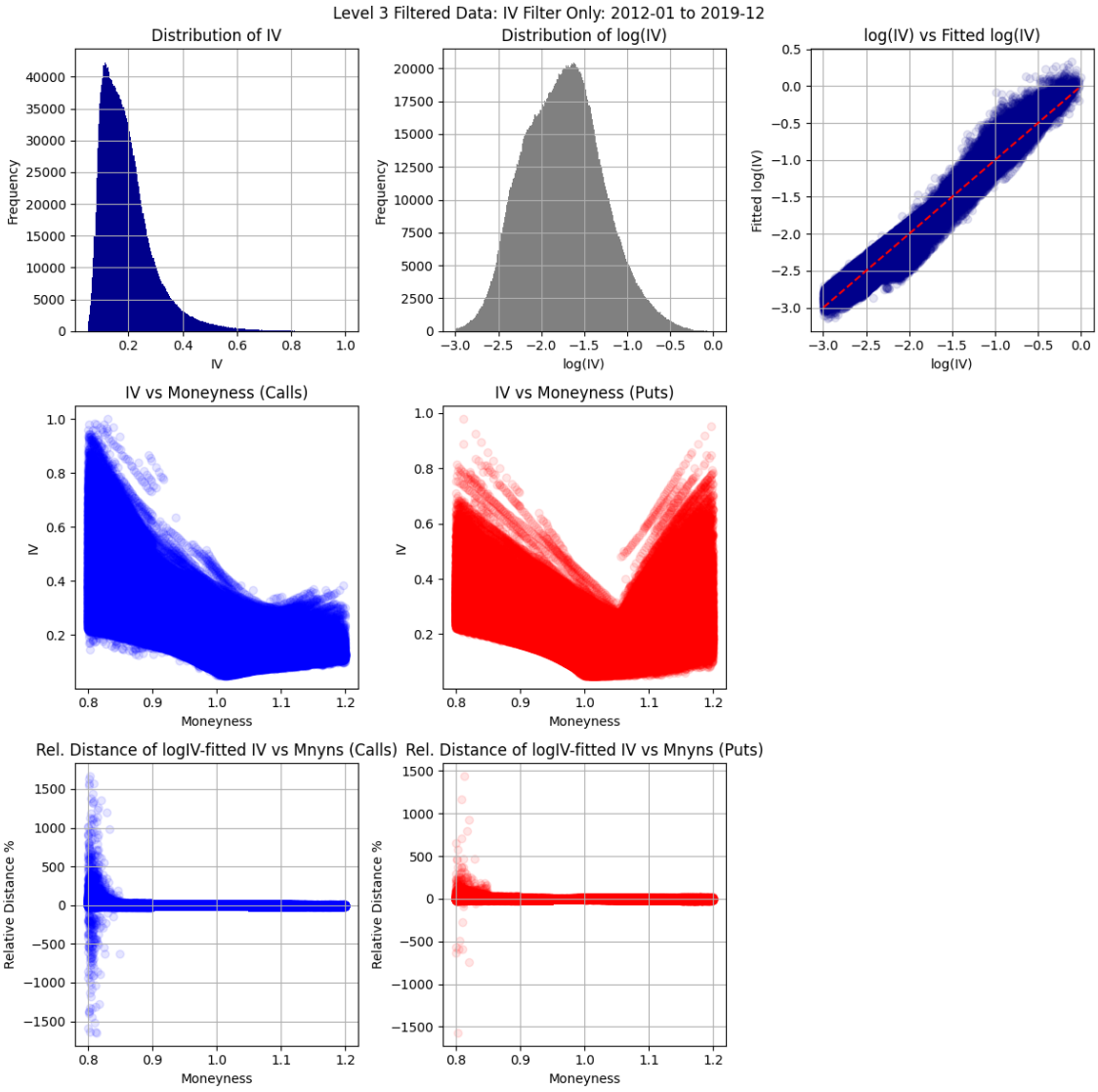

Figure 7: Your caption here

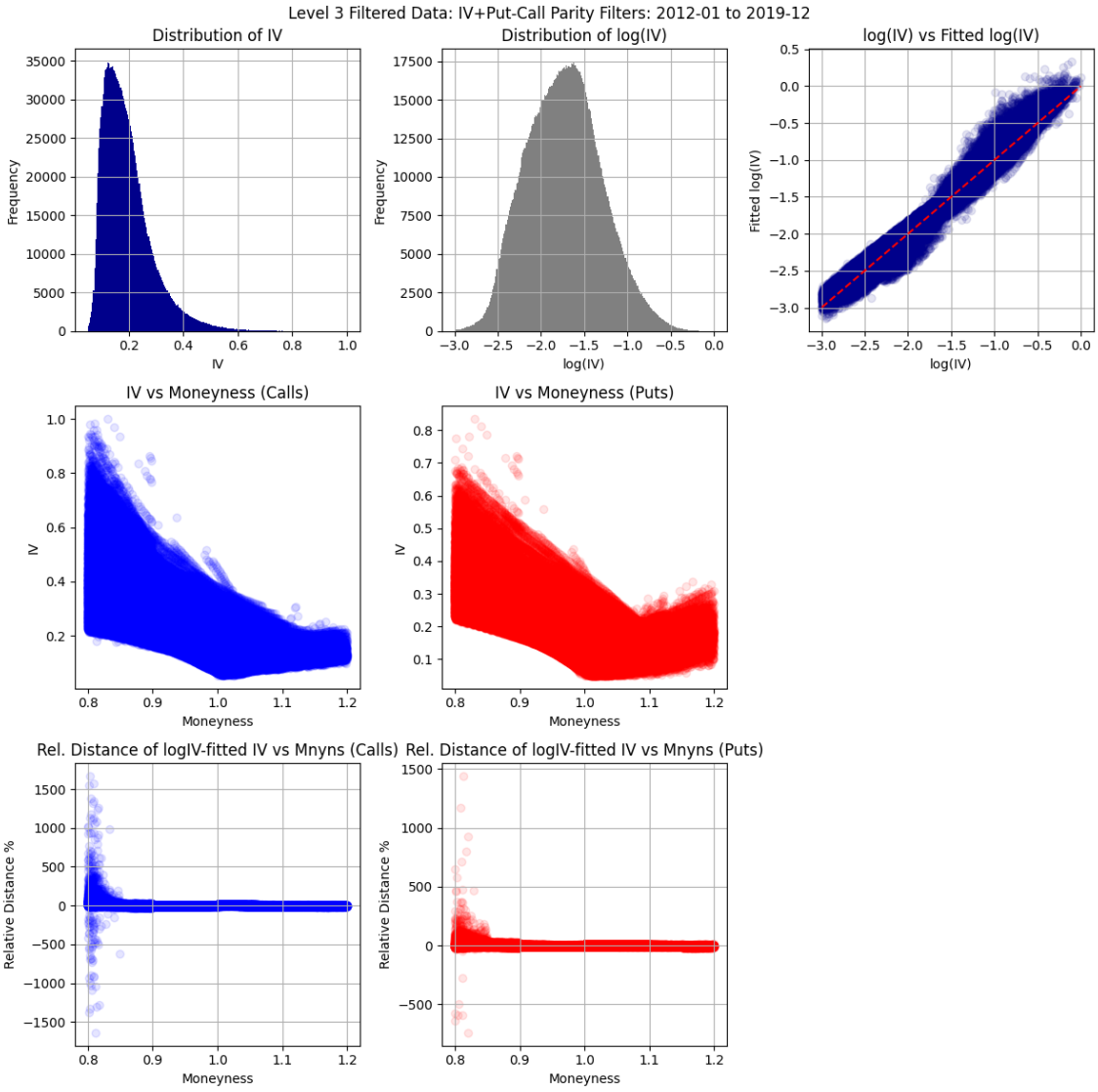

Figure 8: Your caption here