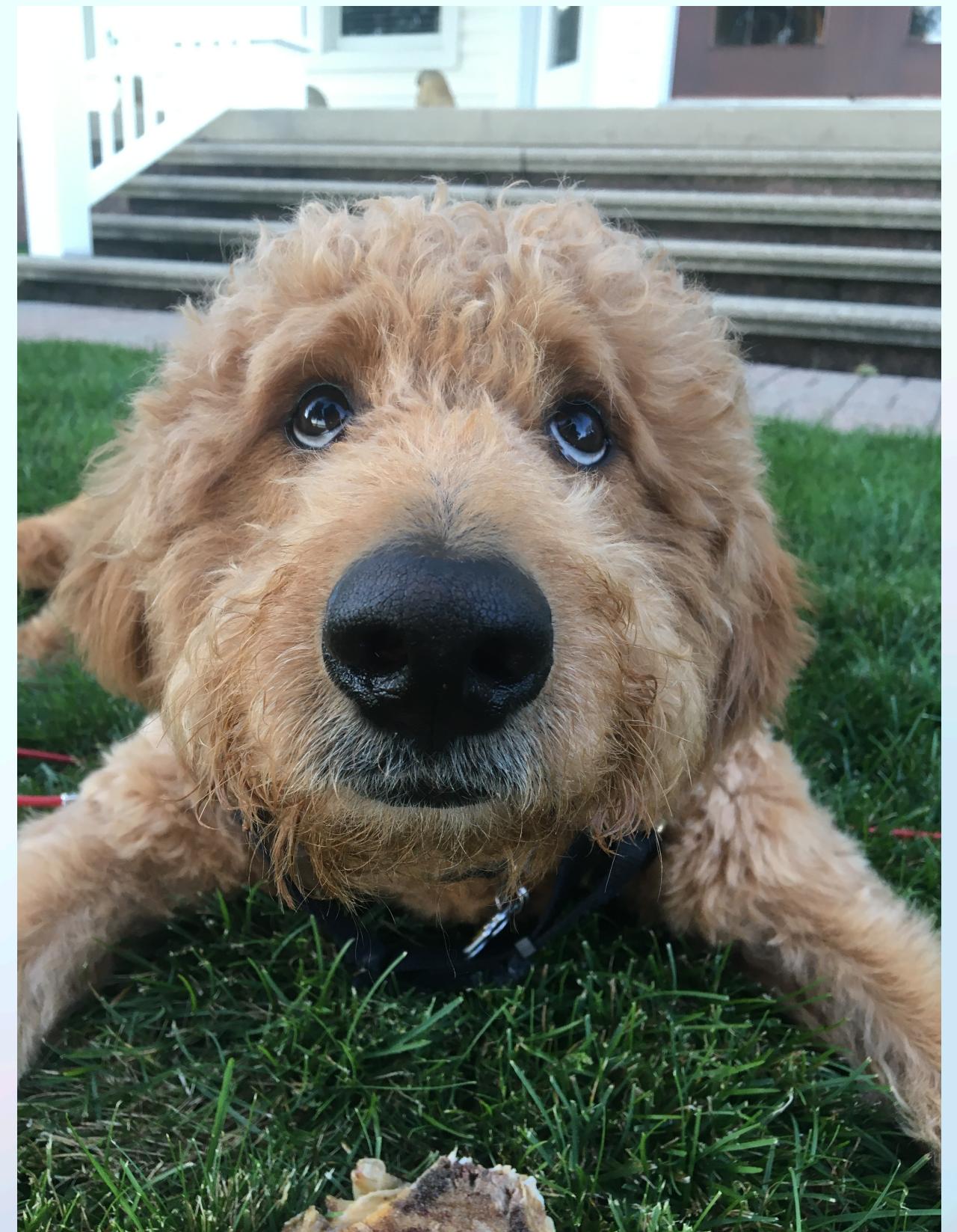


Standardization Via Models

Descriptive Models

- Fitting a model to observed data gives us a pithy description of the data
- Example: median and IQR, say of the time it takes our pets to run across the room
- The statistics not only describe the data, they compress its meaning into a few variables
- Fitting a model to observed data gives us another pithy description of the data
- For example, if our model is $t = D/v$ and we fit velocities v then the v values themselves represent meaning



Descriptive Models

- It becomes easier to infer qualitative differences. This ν parameter we fit for each pet
 - Would look multimodal if we have snakes, cats and dogs
 - Would have missing data if we have fish, too
- If we fit an acceleration a instead, then we may even reduce the *dynamic range* we are examining, since the square root of the data will be in a smaller range
- Fitted values of model parameters also help us find outliers
- Outliers might mean profits!

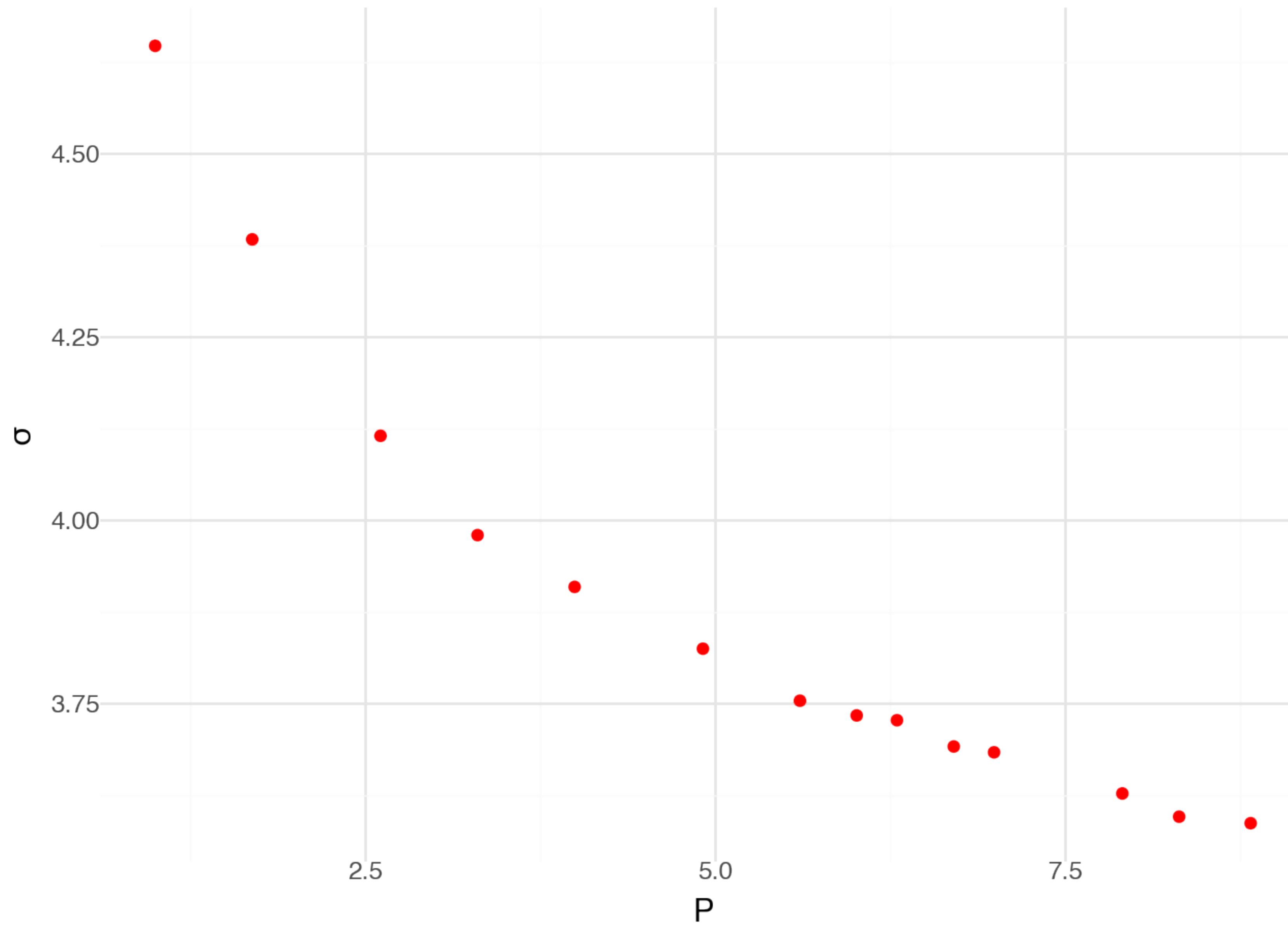


More Complex Example

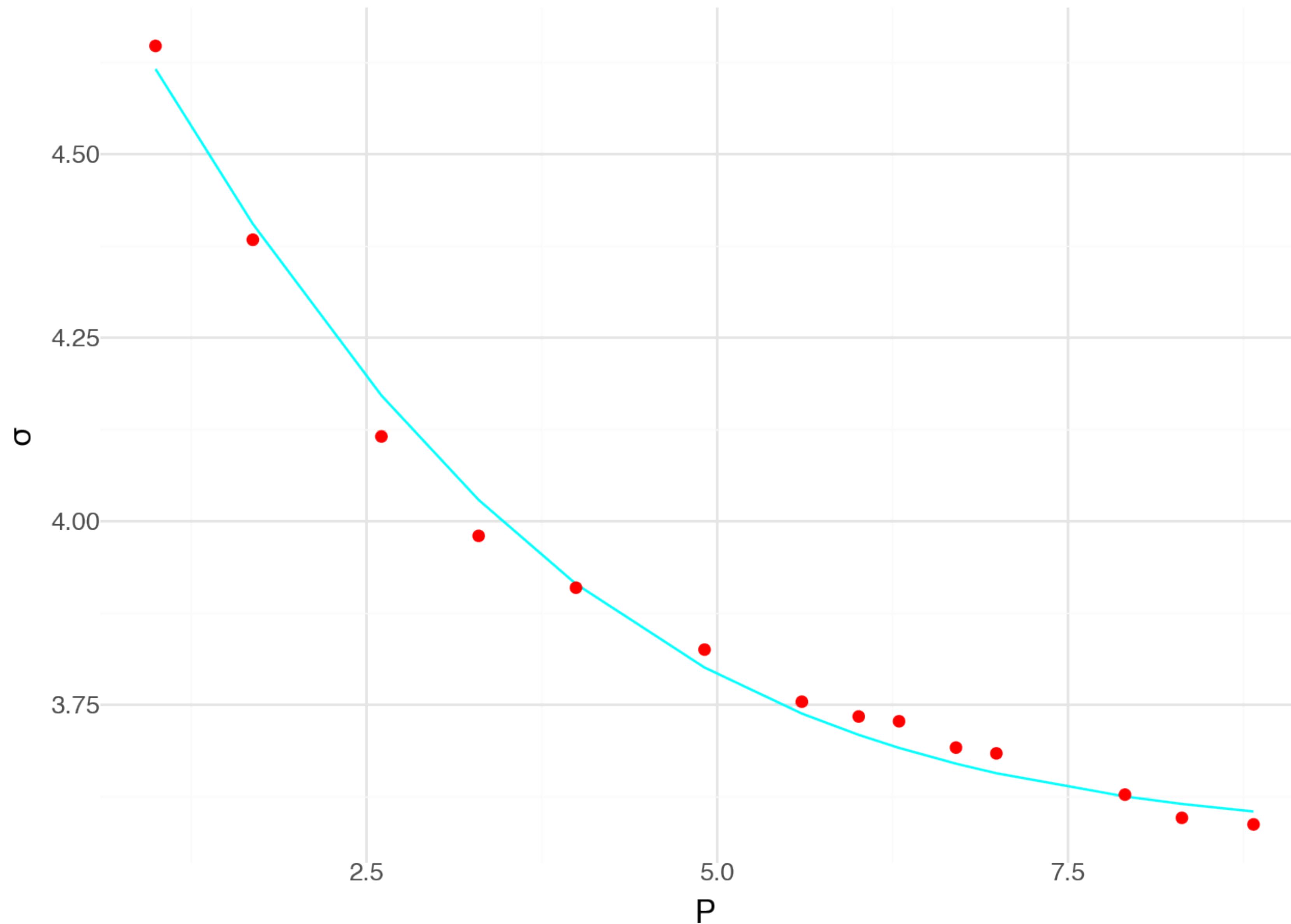
- Let's look at a deterministic, complex system, with measurement error
- We will fit a relatively simple model, one we know is wrong
- Using the model, we can identify unusual aspects of the data
- Our simple model:

$$\sigma = \alpha + \beta\sqrt{P}$$

Model observations



Model prediction versus observation



What Was The Actual System?

Cubes in viscous medium, side length \mathbf{L}

$$F_{\max,i}^{(0)} = \kappa L_i^{3\alpha}, \quad 0 < \alpha < 1, \kappa > 0.$$

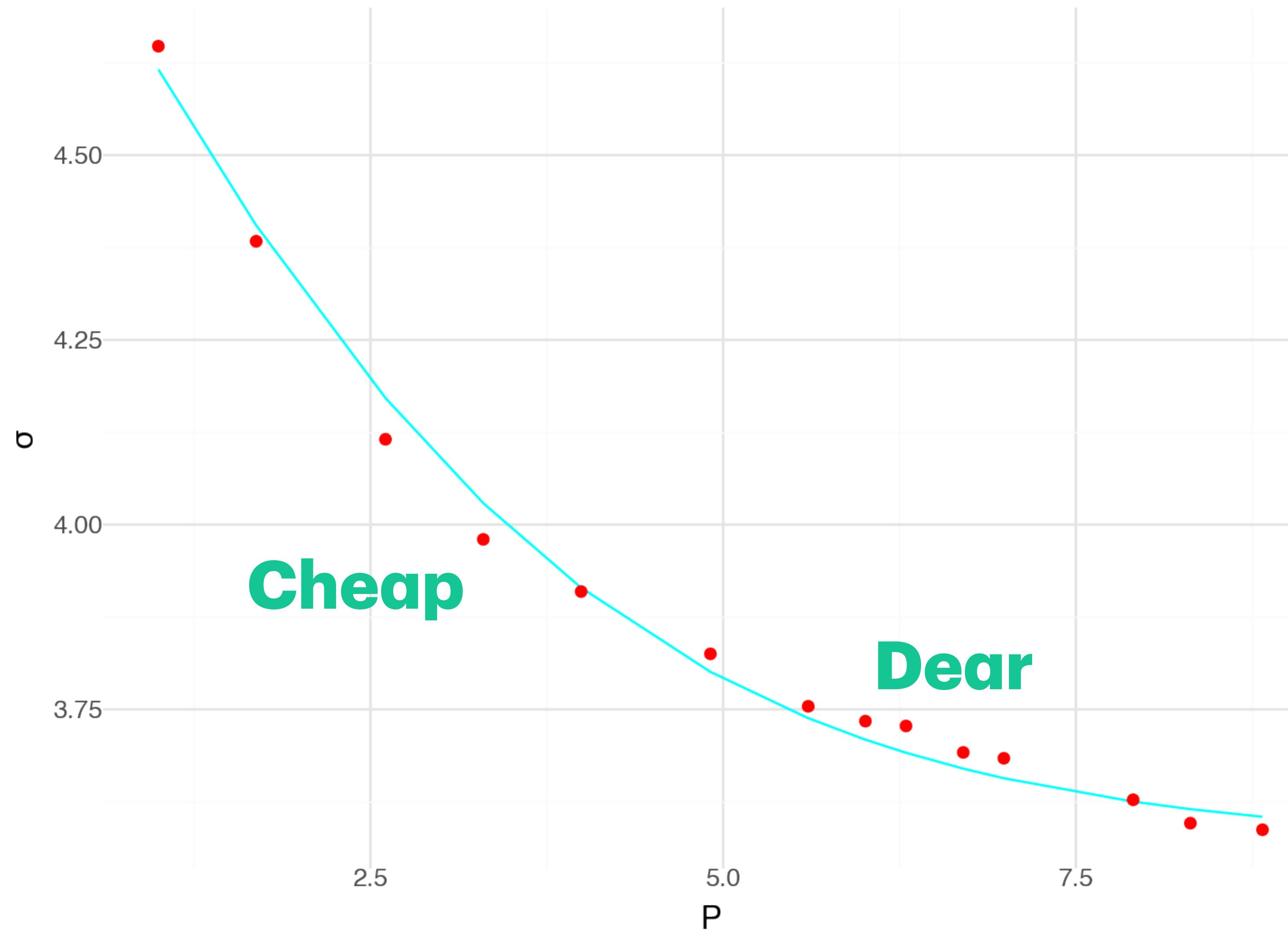
$$F_{d,i}(v_i) = \frac{1}{2} \rho_f C_d A_i v_i |v_i|,$$

What Was The Actual System?

Cubes in viscous medium, side length L

- Errors really differ by cube size
- Examining versus simple fit allows us to find patterns in those errors
- Patterns indicate
 - Measurement error
 - Model error
 - "Market" error

Model prediction versus observation



Option Implied Volatility Skews

TSLA Data

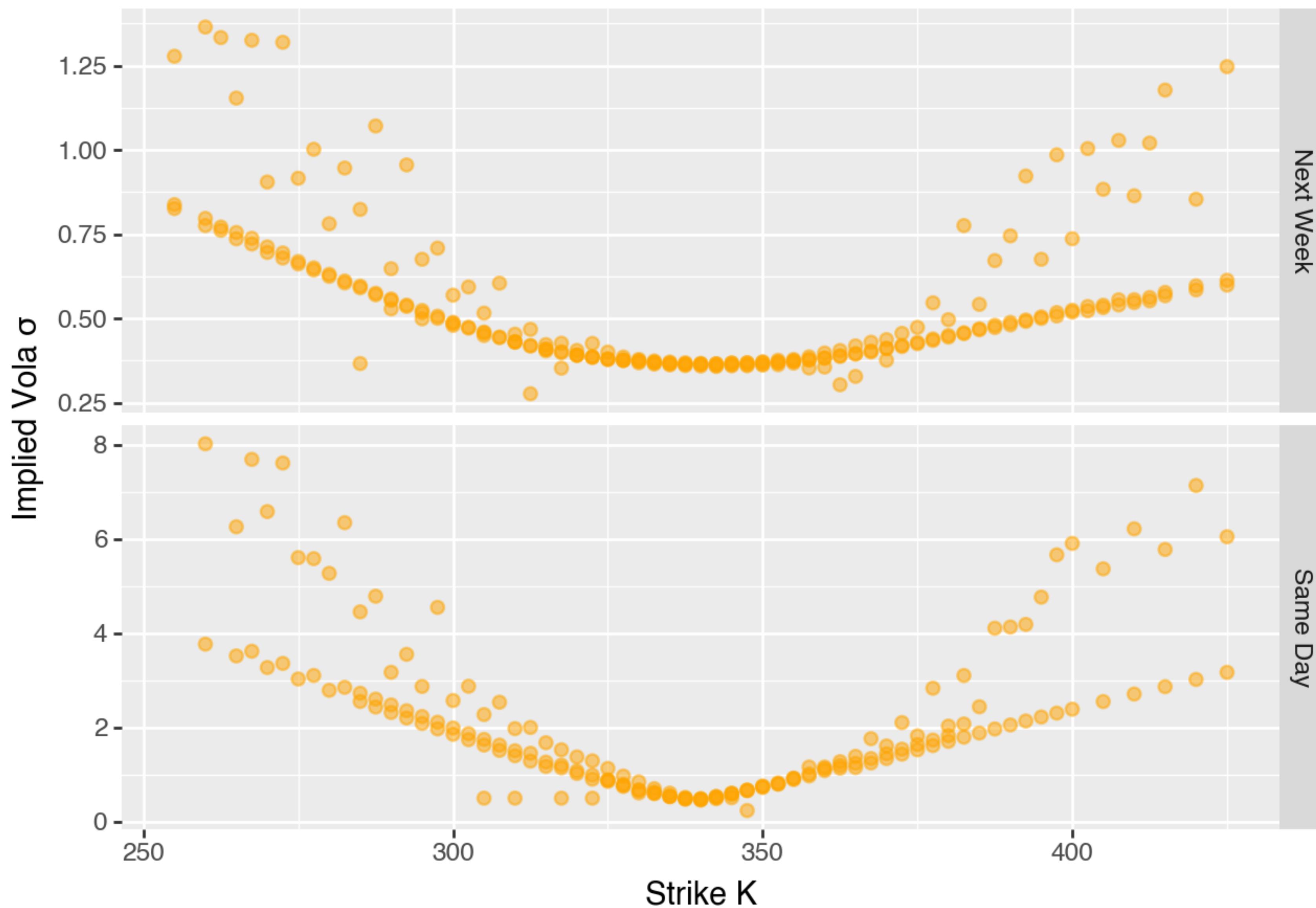
No Bid

Quotes

One Tick

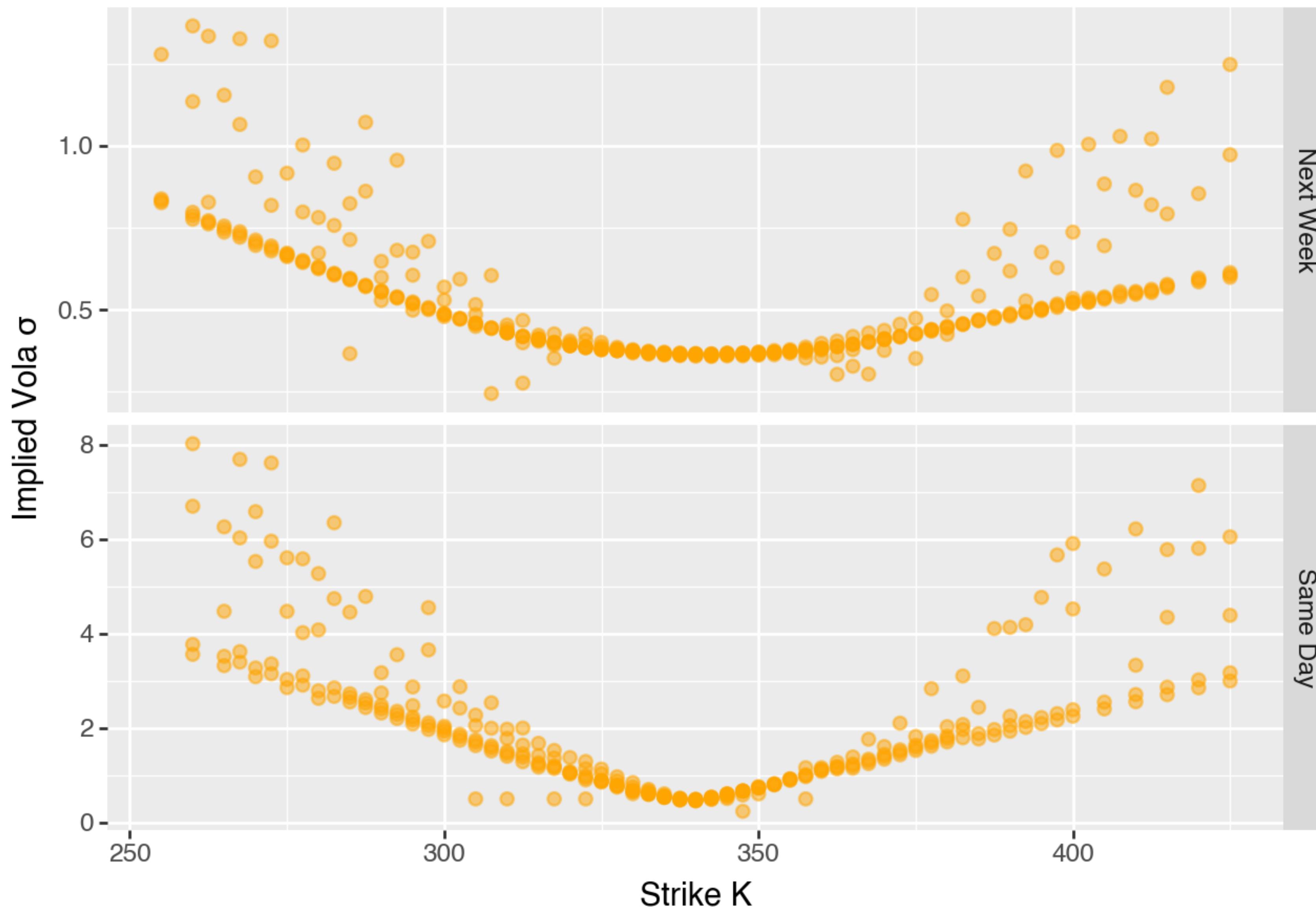
Tenor	Strike	Call Bid	Call Ask	Put Bid	Put Ask
Same Day	260	79.00	81.25	0.00	0.01
Same Day	340	1.19	1.22	1.87	1.90
Same Day	425	0.00	0.01	84.85	87.05
Next Week	260	79.20	81.65	0.08	0.10
Next Week	340	6.75	6.80	7.10	7.20
Next Week	425	0.04	0.05	84.40	88.60

TSLA Options Skew Bids And Offers



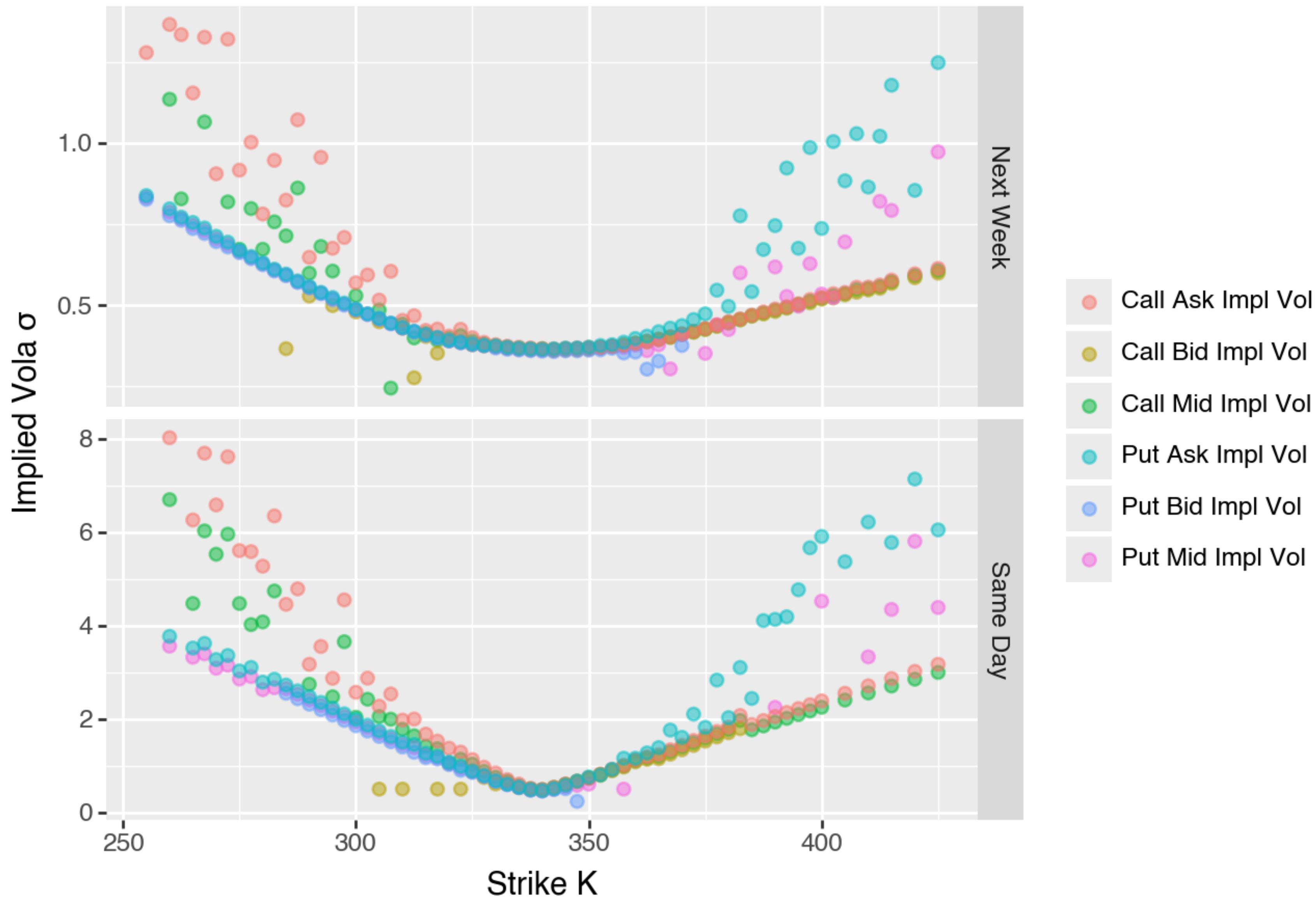
There is structure here, what if we also
imply volatilities from the mid prices?

TSLA Options Implied Volatility Mid Included



No better, maybe worse. Try
finding more categorical structure?

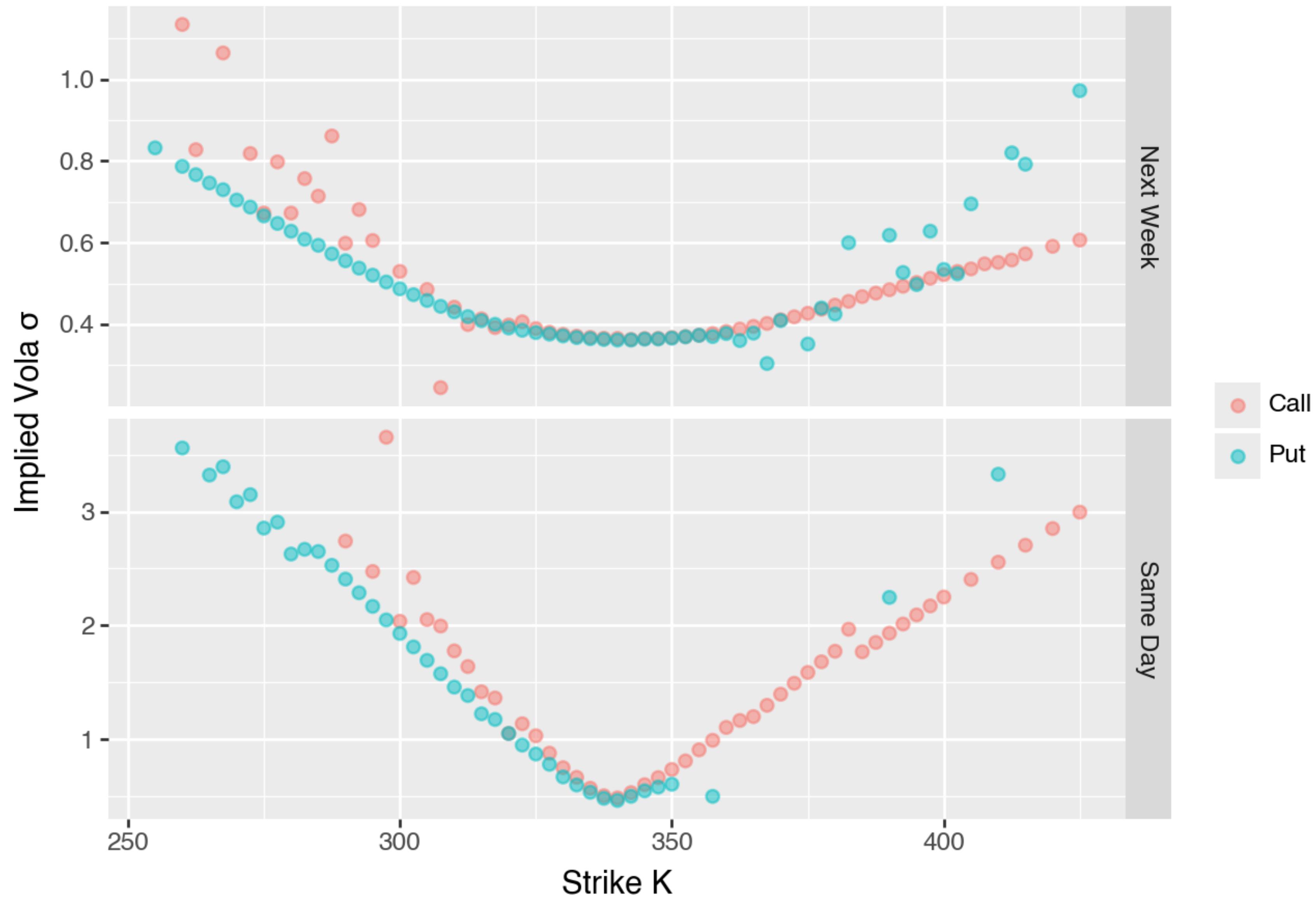
TSLA Options Implied Volatility Categorized Mid Included



Maybe better inference with
mid only?

TSLA Options Skew

Mid Only



What are the empirical
consequences of mistakes?

TSLA Options Skew, Wrong Drift Mid Only

