1 Choosing the Model for the Temporary Impact

In modeling temporary market impact $g_t(x)$, which represents the slippage cost incurred when executing a trade of size x, a common simplification is the linear approximation, x. While this is analytically convenient, real order book data often demonstrates that slippage does not scale linearly with order size. In fact, the true behavior is often nonlinear and saturating. To evaluate this empirically, I analyzed minute-level order book data across 21 trading days for the three tickers provided in the assignment: CRWV, FROG, and SOUN. For each 1-minute snapshot, I simulated market buy orders of varying sizes, computed the volume-weighted average execution price, and calculated slippage relative to the mid-price. I then fit a variety of models to the observed slippage curves, which included:

1. Linear function
2. Square root function
3. Logarithmic function
4. Power law function
5. Sigmoid function

I chose to test the square-root and logarithmic functions as candidate models for the temporary market impact function $g_t(x)$ because they naturally exhibit saturation behavior — slippage increases with order size but at a diminishing rate. This reflects a key empirical feature observed in real market data, especially for moderately sized orders that do not fully consume the book. In addition, I included the sigmoid function, which also captures saturation but with the added benefit of being nearly flat for small values of x. This subtle curvature at the low end matches the behavior I consistently saw in the data, where small orders often resulted in negligible slippage. To compare these models, I fit each function to slippage curves computed for every 1-minute interval across the dataset. For each snapshot, I calculated the RMSE between the predicted and observed slippage, then averaged these RMSE values across all time periods. This process allowed me to empirically identify the function that best fits the data. Among the tested models, the sigmoid function consistently produced the lowest average RMSE, supporting its use as a more accurate and flexible representation of temporary impact compared to simpler functional forms. The results for the CRWV data can be summarized in the below table:

Model	**Average RMSE**
Linear	0.015822
Log	0.026517
Square-root	0.017302
Power	0.013171
Sigmoid	0.008150

2 Mathematical Framework

We want to find x_t at time t such that the total temporary impact of the executed orders is minimized. Essentially, we can write a 'Lagrangian' functional of the following form:

$$L(g_t(x_t), \lambda) = \sum_{t=1}^{N} g_t(x_t) + \lambda \left(\sum_{t=1}^{N} x_t - S \right), \tag{1}$$

where λ is a Lagrange multiplier to ensure that the total number of orders is S as required. To find x_t that minimizes the total temporary impact, we take a derivative of L with respect to x_t, and set it to zero.

$$\frac{\partial L}{\partial x_t} = \frac{\partial g_t(x_t)}{\partial x_t} + \lambda = 0 \tag{2}$$

$$\frac{\partial L}{\partial \lambda} = \sum_{t=1}^{N} x_t - S = 0 \tag{3}$$

This is a system of $N + 1$ equations which can be used to solve for x_t and λ depending on the model used for $g_t(x_t)$.

2.1 Assuming g_t is a power law function

To make calculations easy and obtain an analytical solution, I will assume g_t is a power function of x, but the same process would work for the sigmoid function, except we would have to solve the equations numerically rather than analytically.

Suppose $g_t = \beta_t x_t^{\alpha}$. Plugging this into equation 2, we can write:

$$\alpha \beta_t x_t^{\alpha-1} + \lambda = 0 \rightarrow x_t = \left(\frac{-\lambda}{\alpha \beta_t} \right)^{1/(\alpha-1)}. \tag{4}$$

Let $D_t = \left(\frac{1}{\beta_t} \right)^{1/(\alpha-1)}$ and $C = \left(\frac{-\lambda}{\alpha} \right)^{1/(\alpha-1),}$ then we can write:

$$x_t = C D_t. \tag{5}$$

Plugging this into equation 3, we get:

$$\sum_{t=1}^{N} C D_t = S \rightarrow C = \frac{S}{\sum_{t=1}^{N} D_t}, \tag{6}$$

and we can finally write:

$$x_t = \frac{S D_t}{\sum_{t=1}^{N} D_t}. \tag{7}$$

A similar process would follow for the different models, except some of them would require numerical solutions.