

Too Fast or Too Slow? Determining the Optimal Speed of Financial Markets*

Daniel Fricke[†] Austin Gerig[‡]

How fast should a security trade? To answer this question, we model the trading of a security via periodic batch auctions and study how market quality is affected as the clearing frequency is changed. In the model, the optimal clearing frequency depends on three factors: (1) the volatility of the security, (2) the intensity of trading in the security, and (3) the correlation of the security's value with other securities. Using rough estimates for these values, we determine that the ideal interval of trade for a typical U.S. stock is currently 0.2 to 0.9 seconds. Our analysis suggests that speed is important in financial markets and that time delays of even a fraction of a second can harm market quality. On the other hand, our results also suggest that for many securities, milli- and microsecond speeds are unnecessary.

Keywords: call auctions; clearing frequency; co-location; high-frequency trading; latency; liquidity.

JEL Classification: G14, G19.

*An earlier version of this manuscript was entitled *Liquidity Risk, Speculative Trade, and the Optimal Latency of Financial Markets*. This work was supported by the European Commission FP7 FET-Open Project FOC-II (no. 255987) and the European Union Seventh Framework Programme FP7/2007-2013 under grant agreement CRISIS-ICT-2011-288501.

[†]Institute for New Economic Thinking and Saïd Business School, University of Oxford; Kiel Institute for the World Economy; email: daniel.fricke@sbs.ox.ac.uk

[‡]U.S. Securities and Exchange Commission; email: geriga@sec.gov; The Securities and Exchange Commission, as a matter of policy, disclaims responsibility for any private publication or statement by any of its employees. The views expressed herein are those of the author and do not necessarily reflect the views of the Commission or of the author's colleagues on the staff of the Commission.

I Introduction

Over the last two decades, technological innovations have dramatically increased the speed of trading in global financial markets. The time it takes for messages to travel within and between financial markets is now measured in micro- and milliseconds and is limited, for the most part, only by the speed of light. Although most would agree that extremely long delays in markets are undesirable, it is unclear if millisecond or microsecond speeds are really necessary – perhaps they are even harmful.

How does market speed affect investors? Are faster speeds always more desirable, or are there limits to the benefits of speed? Is there an optimal speed of trading, and if so, what determines this speed?

To help answer these questions, we model the trading of a security via periodic batch auctions and study how market quality is affected as the clearing frequency is changed. In the model, the optimal clearing frequency depends on three factors: (1) the volatility of the security, (2) the intensity of trading in the security, and (3) the correlation of the security's value with the market security. All else equal, a security should be traded faster if its volatility is higher, slower if its intensity of trade is lower, and faster if its correlation with the market is higher. Using rough estimates of these values, we determine that the optimal time interval of trade for a typical U.S. stock is currently 0.2 to 0.9 seconds.

Our model is based on the batch auction model of Garbade and Silber (1979). Although their paper is primarily focused on studying liquidity, the model they develop lends itself nicely to an analysis of speed in markets. In their model, the liquidity of the market is maximized at intermediate auction intervals, when markets are neither too fast or too slow. The reason behind this intermediate result is rather intuitive. If markets are too fast, then very few orders are mixed in each market clearing and transaction prices will not coincide with equilibrium values. If markets are too slow, then orders will sit for long periods of time and prices will have shifted by the time

the orders clear. In either case, because speeds are not set appropriately, market quality is harmed.

We modify the model in Garbade and Silber (1979) in two ways. First, we assume that investors' reservation prices are normally distributed around current equilibrium prices (instead of a future equilibrium price as in the original model). We therefore can derive market quality as a function of the arrival time of an investor. Second, we add a market security. The addition of the market security is especially important. Liquidity providers can use information about the market to push the price of the original security closer to its equilibrium value, which allows the security to trade faster. For example, a security that has correlation $\rho = 0.5$ with the market will have an optimal speed that is 11% faster than if the market security were not present. The effect becomes even more pronounced for higher correlations: a security with $\rho = 0.85$ will have an optimal speed 729% faster compared to the case without a market security.

To present analytic results, we assume that the market security is infinitely liquid and therefore trades at infinitely short intervals. Interestingly, we find that this continuous trading property can be transmitted to other securities. Specifically, we find that it is optimal for a security to also trade at infinitely short intervals if its correlation with the market is above or below a critical threshold of $\rho^c = \pm\sqrt{3/4}$. This continuous trading result holds regardless of the other properties of the security.

The theoretical literature on the optimal clearing frequency of markets is relatively sparse. To the best of our knowledge, Garbade and Silber (1979) were the first to show that market quality for an average investor is maximized for intermediate clearing frequencies, i.e., most markets should neither operate continuously nor be cleared very infrequently. Later on, studies were less concerned with determining the optimal speed of markets, but rather compared continuous and periodic market clearings in general (e.g., Madhavan (1992) and Budish et al. (2013)).

More recently, Farmer and Skouras (2012), Budish et al. (2013), and Cochrane (2013) have proposed periodic market clearings as a method to mitigate an “arms

race” for speed among liquidity providers and the eventual over-investment in technology that results.¹ Notably, these papers do not include an analysis of the optimal time between clearings, nor do they allow welfare improvements due to higher speed.

Finally, our paper is related to the literature on the relationship between the liquidity of an asset and its correlation with the overall market. In a recent empirical study, Chan et al. (2013) show that the liquidity of a security increases with the fraction of volatility due to systematic risk, exactly as predicted in our model. Furthermore, they find that improvement in liquidity following the addition of a stock to the S&P 500 Index is directly related to the stocks increase in correlation with the market. Baruch and Saar (2009) and Gerig and Michayluk (2014) both model the relationship between the liquidity of a security and its correlation with other securities. In their models, as in our model, liquidity providers can form a better estimate of prices when observing order flow from correlated assets which improves overall market quality.

II Baseline Model

As in Garbade and Silber (1979), we consider a single security that is traded by public investors in a market with periodic clearings. (In later sections, we consider the addition of liquidity providers and also a second security.) The time interval between clearings is τ , and ultimately, we will be interested in determining the optimal τ from an average investor’s perspective.²

Between clearings, investors (indexed in each interval by i) arrive at a constant rate ω and submit excess demand schedules to the market. These demand schedules are unobservable to other investors and remain in the market until the next market clearing. At each clearing, the transaction price is set to the value that clears the market, i.e., to the value that produces zero aggregate excess demand. The excess

¹The core idea behind socially wasteful investment in speed was originally discussed in Hirshleifer (1971) and Stiglitz (1989).

²Note that we attempt to keep our notation as consistent as possible with Garbade and Silber’s original paper.

demand schedule of the i th investor is a linearly increasing function of the reservation price of the investor, r_i , and a linear decreasing function of the clearing price, p ,

$$D(p) = a(r_i - p), \tag{1}$$

where a is a positive constant assumed the same for all investors.³ Note that the i th investor will be a net seller of the security if $r_i < p$ and will be a net buyer if $r_i > p$.

Between any two clearings, a total number $K = \omega\tau$ investors will submit excess demand schedules to the market. The market clearing price is the unique price that sets aggregate excess demand to zero,

$$0 = \sum_{i=1}^K a(r_i - p). \tag{2}$$

Rearranging the equation reveals that the clearing price is the average reservation price of the arriving investors,

$$p = \sum_{i=1}^K r_i / K. \tag{3}$$

We assume there exists an unobservable equilibrium price for the security, m_t , at all times and that the reservation price of investor i is normally distributed around the prevailing equilibrium price, $m_{t-1+i/\tau}$ (which we denote m_i for short), at the instant the investor decides to trade,⁴

$$r_i = m_i + g_i, \tag{4}$$

$$g_i \sim N(0, \sigma^2), \tag{5}$$

³A linear demand function arises when agents optimize a quadratic utility function subject to their budget constraint.

⁴In Garbade and Silber (1979), the investor decides to trade at time $t - 1/2$ but has a reservation price that is normally distributed around the future equilibrium price at time t . We have chosen a different setup (which we believe is more natural) where the reservation price of an investor is normally distributed around the instantaneous equilibrium price at the time he/she decides to trade. This departure means that our equations will use the average equilibrium price over the interval τ rather than the instantaneous equilibrium price as in Garbade and Silber (1979).

where g_i is assumed to be uncorrelated across investors. We denote by \bar{r}_t the average reservation price of the investors at market clearing t (which is the market clearing price when the market does not contain liquidity providers),

$$\bar{r}_t = \sum_{i=1}^K (m_i + g_i)/K. \quad (6)$$

We denote by \bar{m}_t the average equilibrium price over the interval, $\bar{m}_t = \sum_i m_i/K$, and we denote by f_t the average of g_i , i.e., $f_t = \sum_i g_i/K$. Note that,

$$\bar{r}_t = \bar{m}_t + f_t, \quad (7)$$

$$f_t \sim N(0, \sigma^2/(\omega\tau)). \quad (8)$$

We assume that the instantaneous equilibrium price m_t evolves as a driftless Brownian motion with variance $(3/2)\psi^2$, i.e., $m_t = \sqrt{3/2}\psi B_t$ (the prefactor $\sqrt{3/2}$ is used for convenience and its purpose will become apparent in the following equation). Therefore, the average equilibrium price for investors at clearing t evolves according to the following equation,

$$\bar{m}_t = \bar{m}_{t-1} + e_t, \quad (9)$$

$$e_t \sim N(0, \tau\psi^2), \quad (10)$$

where we have used the result that the variance of the difference between two consecutive averaged points (each over an interval τ) of a standard Brownian motion is,

$$\text{Var} \left[(1/\tau) \int_{\tau}^{2\tau} B_t dt - (1/\tau) \int_0^{\tau} B_t dt \right] = (2/3)\tau. \quad (11)$$

We assume that e_t is serially uncorrelated and also uncorrelated with g_i and therefore f_t .

A Liquidity

As in Garbade and Silber (1979), we measure market quality with an inverse liquidity metric called *liquidity risk*. Liquidity risk is defined as the variance of the difference between the equilibrium value of the security when an investor arrives at the market, m_i , and the transaction price ultimately realized for the investor's trade, in this case \bar{r}_t .⁵ The liquidity risk for investor i in a market without liquidity providers is therefore,

$$V_P = \text{Var}[\bar{r}_t - m_i], \quad (12)$$

$$= \text{Var}[(\bar{r}_t - \bar{m}_t) + (\bar{m}_t - m_i)], \quad (13)$$

$$= \text{Var}[\bar{r}_t - \bar{m}_t] + \text{Var}[\bar{m}_t - m_i], \quad (14)$$

where the two expressions in parentheses separate because there is no covariance between them. The variance of the first term, $\text{Var}[\bar{r}_t - \bar{m}_t]$, is just the variance of f_t . For the second term, the variance depends on the arrival time of the investor. If the investor arrives at a point in time that is a fraction ϕ of the total interval τ from the previous clearing, then the variance of the second term will be,

$$\text{Var}[\bar{m}_t - m_i] = \text{Var} \left[\left(\int_0^{\phi\tau} \sqrt{3/2} \psi B_t dt + \int_0^{\tau-\phi\tau} \sqrt{3/2} \psi B_t dt \right) / \tau \right], \quad (15)$$

$$= (1/2) [\phi^3 + (1 - \phi)^3] \tau \psi^2. \quad (16)$$

If the investor arrives at the beginning or end of the interval ($\phi = 0$ or $\phi = 1$), then the variance is at its maximum value, $(1/2)\tau\psi^2$, and if the investor arrives in the middle of the interval ($\phi = 1/2$), the variance is at its minimum value, $(1/8)\tau\psi^2$. The final equation for liquidity risk in a market of public investors is therefore,

$$V_P = \sigma^2/(\omega\tau) + (1/2) [\phi^3 + (1 - \phi)^3] \tau \psi^2. \quad (17)$$

⁵Grossman and Miller (1988) use a very similar definition of liquidity risk.

If we assume that the timing of an investor’s trading decision is uncorrelated with the timing of market clearings, we can average over all ϕ in the interval $[0, 1]$, which gives $\int_0^1(\phi^3 + (1 - \phi)^3) = 1/2$. Liquidity risk for the average investor is therefore,⁶

$$V_P = \sigma^2/(\omega\tau) + \tau\psi^2/4. \quad (18)$$

Notice that liquidity risk is increasing in the volatility of the security, increasing in the variance of investor reservation prices, and decreasing in the frequency of investor arrival. The effect of the clearing frequency ($1/\tau$) on liquidity risk is nonlinear. When market clearings are frequent, this decreases the difference between the clearing price and the average equilibrium price of the security, but it also increases the difference between the average equilibrium price of the security and the specific equilibrium price used as a reference by the investor. There is a “Goldilocks” value for τ that optimizes the tradeoff between these two effects, and we determine this value below.

The optimal trading interval τ_P^* from an investor’s perspective is just the value of τ that minimizes liquidity risk. This value can be found by taking the derivative of liquidity risk with respect to τ and setting to zero,

$$\tau_P^* = 2 \frac{\sigma/\omega^{1/2}}{\psi}. \quad (19)$$

The minimum value of liquidity risk, $V_P^* = V_P(\tau_P^*)$, is,

$$V_P^* = (\sigma/\omega^{1/2}) \psi. \quad (20)$$

In Fig. 1, we plot liquidity risk as a function of the time between market clearings, τ , when $\psi = 1$, $\sigma = 1$, and $\omega = 10$. We also show the optimal point (V_P^*, τ_P^*) .

⁶Our equation for liquidity risk is slightly different from that of Garbade and Silber (1979) because of the modifications we have made to their setup.

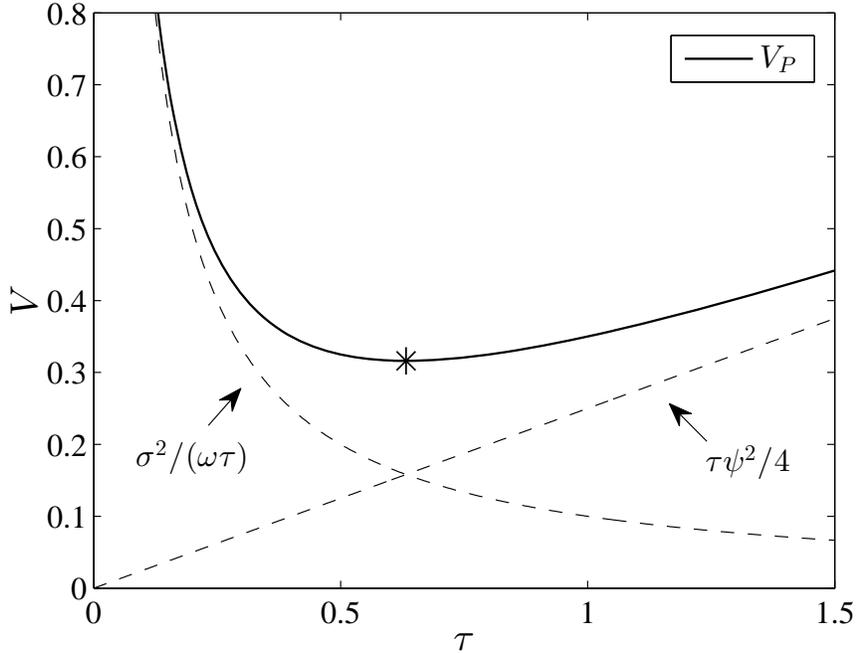


Figure 1: Liquidity risk, V_P , as a function of the time between market clearings, τ , in a public market without a liquidity provider. Parameters used in the plot are $\psi = 1$, $\sigma = 1$, and $\omega = 10$. The optimal point (V_P^*, τ_P^*) is shown with an asterisk. Also shown are the components of liquidity risk $\sigma^2/(\omega\tau)$ and $\tau\psi^2/4$.

III Model with a Liquidity Provider

As discussed in Garbade and Silber (1979), enterprising individuals can devise a better estimate for the equilibrium price than is contained in the market clearing price r_t and can profit by buying and selling according to this estimate. In effect, these speculators act as liquidity providers in the market.

Here, we assume that a single competitive and risk neutral liquidity provider exists, that she observes the aggregate excess demand of the market directly before the market is cleared, and that she submits an excess demand schedule at each market clearing such that the clearing price always equals her estimate of the equilibrium price. Many of the seminal market microstructure papers published after Garbade and Silber (1979) (such as Kyle (1985) and Glosten and Milgrom (1985)) assume the

same type of competitive, risk neutral liquidity provider. However, in these other models, the benefit of the liquidity provider cannot be analyzed, whereas it can in Garbade and Silber's framework. Below we show that the liquidity provider reduces the minimum liquidity risk of public investors by a factor of 1.5. In the next section, we show that when the liquidity provider is further enabled so that she observes the price of the market security, liquidity risk is reduced even further.

A Liquidity

The liquidity provider observes the order flow of the public investors and therefore \bar{r}_t . This information allows her to form an estimate of the average equilibrium price over the interval, which we denote by \hat{m}_t . Because she is competitive and risk neutral, she submits a demand schedule that forces the clearing price to this value. Therefore, in the equation for liquidity risk, the clearing price is \hat{m}_t instead of \bar{r}_t .

The model with a liquidity provider is a special case of the model presented in the next section. Here, we just present results for liquidity risk and leave details of the derivation to the next section and the Appendix.

$$V_L = \text{Var}[(\hat{m}_t - \bar{m}_t) + (\bar{m}_t - m_i)], \quad (21)$$

$$= \text{Var}[\hat{m}_t - \bar{m}_t] + \text{Var}[\bar{m}_t - m_i] + 2 \text{Cov}[\hat{m}_t - \bar{m}_t, \bar{m}_t - m_i], \quad (22)$$

$$= \frac{2(\phi_1 + 2\phi_2)\tau\psi^2 + 2(\phi_1 - 2\phi_2)\tau\psi^2\sqrt{1 + \frac{4\sigma^2/\omega}{\tau^2\psi^2}} + 4\sigma^2/(\omega\tau)}{2\left(1 + \sqrt{1 + \frac{4\sigma^2/\omega}{\tau^2\psi^2}}\right)}, \quad (23)$$

where,

$$\phi_1 \equiv (1/2) [\phi^3 + (1 - \phi)^3], \quad (24)$$

$$\phi_2 \equiv (1/4) [\phi^3 + 2(1 - \phi)^3 + 3(1 - \phi)\phi^2]. \quad (25)$$

As before, assuming that the investor's arrival time is not correlated with the timing

of market clearings, then liquidity risk is the expectation over ϕ ,

$$V_L = \frac{(1/2)\tau\psi^2 \left(1 - \sqrt{1 + \frac{4\sigma^2/\omega}{\tau^2\psi^2}}\right) + 4\sigma^2/(\omega\tau)}{2 \left(1 + \sqrt{1 + \frac{4\sigma^2/\omega}{\tau^2\psi^2}}\right)}. \quad (26)$$

A plot of $V_L(\tau)$ is shown later in Fig. 4. The optimal trading interval τ_L^* is,

$$\tau_L^* = \left(\frac{2}{\sqrt{3}}\right) \frac{\sigma/\omega^{1/2}}{\psi}, \quad (27)$$

$$= \left(\frac{1}{\sqrt{3}}\right) \tau_P^*, \quad (28)$$

and the minimum value of liquidity risk is,

$$V_L^* = \left(\frac{7}{6\sqrt{3}}\right) (\sigma/\omega^{1/2}) \psi. \quad (29)$$

Notice that with the liquidity provider, the optimal clearing frequency ($1/\tau_L^*$) increases by a factor of $\sqrt{3} \approx 1.7$ from the public market case (regardless of the other parameters). In addition, the liquidity provider reduces liquidity risk by a factor of $6\sqrt{3}/7 \approx 1.5$, again regardless of the values of other parameters in the model.

IV Model with a Liquidity Provider and Market Information

In general, for a market of N securities, the average reservation price of the different securities at market clearing t can be written,

$$\bar{\mathbf{r}}_t = \bar{\mathbf{m}}_t + \mathbf{f}_t, \quad (30)$$

$$\mathbf{f}_t \sim N(0, \Sigma), \quad (31)$$

and the average equilibrium price over the market clearing interval can be written,

$$\bar{\mathbf{m}}_t = \bar{\mathbf{m}}_{t-1} + \mathbf{e}_t, \quad (32)$$

$$\mathbf{e}_t \sim N(0, \Psi), \quad (33)$$

where $\bar{\mathbf{r}}$, $\bar{\mathbf{m}}$, $\bar{\mathbf{f}}$, and $\bar{\mathbf{e}}$ are $N \times 1$ vectors and Σ and Ψ are $N \times N$ matrices.

For a market of relatively few securities, it is not too difficult to calculate estimates of $\bar{\mathbf{m}}_t$ (denoted $\hat{\mathbf{m}}_t$) and to determine liquidity risk when Σ and Ψ are fully specified. The process involves numerically solving the appropriate discrete time algebraic Riccati equation (see the Appendix) and then using this solution in straightforward equations. Analytic results, however, are often extremely messy – even for just two securities.

In order to present analytic results, we treat the model with a liquidity provider in a market with many assets as a special case of a two security market where the second security is the “market security”,

$$\bar{\mathbf{r}}_t = \begin{pmatrix} \bar{r}_t \\ \bar{r}_{M,t} \end{pmatrix} \quad \bar{\mathbf{m}}_t = \begin{pmatrix} \bar{m}_t \\ \bar{m}_{M,t} \end{pmatrix} \quad (34)$$

$$\mathbf{f}_t = \begin{pmatrix} f_t \\ f_{M,t} \end{pmatrix} \quad \Sigma = \begin{pmatrix} \sigma^2/(\omega\tau) & \varrho\sigma\sigma_M/(\sqrt{\omega\omega_M\tau}) \\ \varrho\sigma\sigma_M/(\sqrt{\omega\omega_M\tau}) & \sigma_M^2/(\omega_M\tau) \end{pmatrix} \quad (35)$$

$$\mathbf{e}_t = \begin{pmatrix} e_t \\ e_{M,t} \end{pmatrix} \quad \Psi = \begin{pmatrix} \tau\psi^2 & \rho\tau\psi\psi_M \\ \rho\tau\psi\psi_M & \tau\psi_M^2 \end{pmatrix}, \quad (36)$$

where ϱ is the correlation of the difference between reservation prices and equilibrium prices across the two assets, and ρ is the correlation of equilibrium price changes across the two assets. We make an idealized assumption that order flow for the market security is so frequent that $\omega_M \gg 1$ and,

$$\Sigma \approx \begin{pmatrix} \sigma^2/(\omega\tau) & 0 \\ 0 & 0 \end{pmatrix} \quad (37)$$

The liquidity provider, therefore, has noiseless information about the average equilibrium price of the market security at each clearing.

A Liquidity

Liquidity risk is,

$$V_M = \text{Var}[(\hat{m}_t - \bar{m}_t) + (\bar{m}_t - m_i)], \quad (38)$$

$$= \text{Var}[\hat{m}_t - \bar{m}_t] + \text{Var}[\bar{m}_t - m_i] + 2 \text{Cov}[\hat{m}_t - \bar{m}_t, \bar{m}_t - m_i], \quad (39)$$

$$= \mathbf{S}_{(1,1)} + \phi_1 \tau \psi^2 + 2(\mathbf{G}_{(1,1)} - 1)\phi_2 \tau \psi^2 + 2\mathbf{G}_{(1,2)}\phi_2 \rho \tau \psi \psi_M, \quad (40)$$

where $\mathbf{S}_{(1,1)}$, $\mathbf{G}_{(1,1)}$, and $\mathbf{G}_{(1,2)}$ are the respective elements of the matrices used in the Kalman filter when solving for \hat{m}_t . A derivation of this equation is given in the Appendix.

Solving the Riccati equation and plugging into Eq.40 (see the Appendix),

$$V_M = \frac{2(\phi_1 + 2\phi_2\Theta)\tau\psi^2 + 2(\phi_1 - 2\phi_2\Theta)\tau\psi^2\sqrt{1 + \frac{4\sigma^2/\omega}{\Theta\tau^2\psi^2}} + 4\sigma^2/(\omega\tau)}{2\left(1 + \sqrt{1 + \frac{4\sigma^2/\omega}{\Theta\tau^2\psi^2}}\right)}, \quad (41)$$

where $\Theta \equiv 1 - \rho^2$. Again, assuming that the investor's arrival time is not correlated with the timing of market clearings, then liquidity risk is the expectation over ϕ ,

$$V_M = \frac{(1/2 + \Theta)\tau\psi^2 + (1/2 - \Theta)\tau\psi^2\sqrt{1 + \frac{4\sigma^2/\omega}{\Theta\tau^2\psi^2}} + 4\sigma^2/(\omega\tau)}{2\left(1 + \sqrt{1 + \frac{4\sigma^2/\omega}{\Theta\tau^2\psi^2}}\right)} \quad (42)$$

In Fig. 2, we show liquidity risk, V_M , as a function of the time between market clearings, τ , when $\psi = 1$, $\sigma = 1$, $\omega = 10$, and with $\Theta = 0$ to $\Theta = 1$ in increments of 0.1. Liquidity risk decreases as the correlation of the asset with the market increases (i.e., as Θ decreases). When the asset is perfectly correlated with the market, $\Theta = 0$, liquidity risk is simply the line $\tau\psi^2/4$ and there is no risk when markets clear continuously, $\tau = 0$. When the asset is uncorrelated with the market, $\Theta = 1$, liquidity

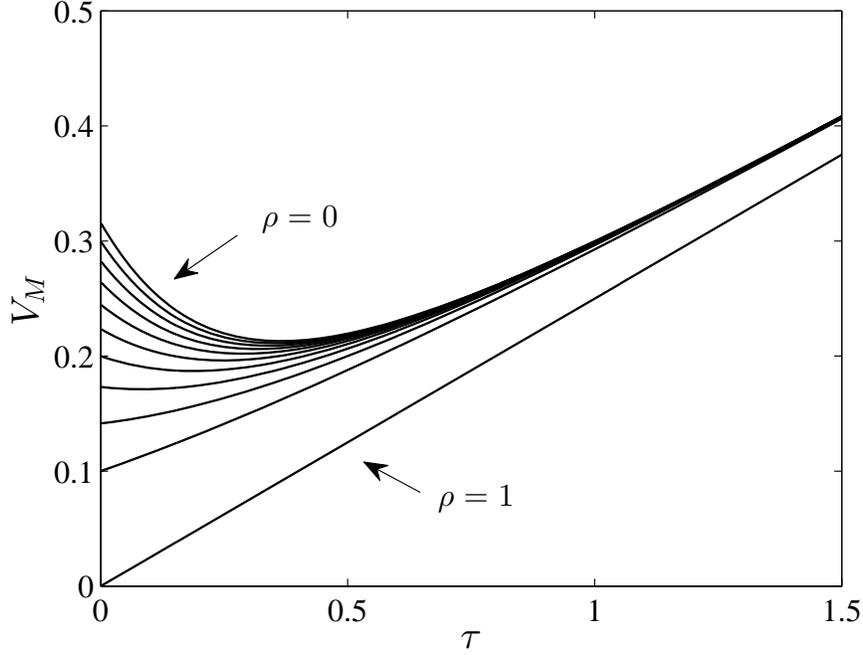


Figure 2: Liquidity risk, V_M , as a function of the time between market clearings, τ , in a market with a liquidity provider and market information. Curves are shown for parameters $\psi = 1$, $\sigma = 1$, $\omega = 10$, and with $\rho = 0$ to $\rho = 1$ ($\Theta = 1$ to $\Theta = 0$ in increments of 0.1).

risk is the same as if the market security was absent, $V_M = V_L$.

The optimal trading interval τ_M^* is,

$$\tau_M^* = \left(\frac{2h_1(\Theta)}{\sqrt{3}} \right) \frac{\sigma/\omega^{1/2}}{\psi}, \quad (43)$$

$$= \left(\frac{h_1(\Theta)}{\sqrt{3}} \right) \tau_P^*, \quad (44)$$

where,

$$h_1(\Theta) = \frac{\sqrt{1 - 32\Theta + 12\Theta^2 + (1 + 6\Theta)\sqrt{1 + 20\Theta + 4\Theta^2}}}{4\Theta}. \quad (45)$$

This equation goes to zero at the critical value $\Theta^c = 1/4$, i.e., when $\rho = \sqrt{3/4} \approx 0.87$. From then on, it is optimal for markets to clear continuously. In Fig. 3, we plot the optimal interval, τ^* , as a function of correlation with the market, ρ , for the three

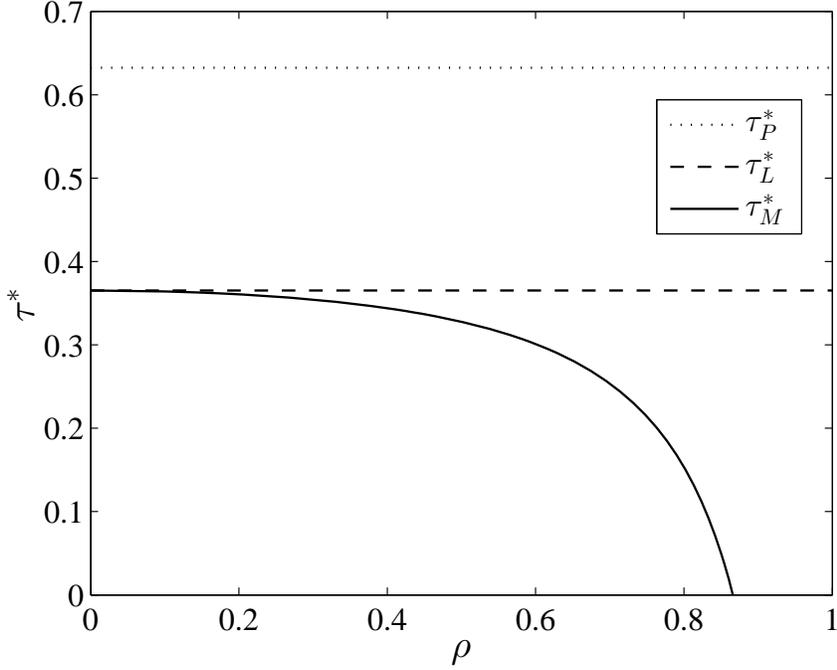


Figure 3: Speed vs. correlation for the three models studied in the text. Parameters used in the plot are $\psi = 1$, $\sigma = 1$, and $\omega = 10$.

models. Notice how $\tau_M^* = 0$ at the critical value $\rho = \sqrt{3/4}$.

For $\Theta > \Theta^c$, the minimum liquidity risk is,

$$V_M^{*+} = h_2(\Theta)/\omega^{1/2} + h_3(\Theta) (\sigma/\omega^{1/2}) \psi, \quad (46)$$

where $h_2(\Theta)$ and $h_3(\Theta)$ are rather complicated functions. For $\Theta \leq \Theta^c$, liquidity risk is minimized when markets clear continuously, i.e., when $\tau = 0$. The equation for liquidity risk when $\Theta \leq \Theta^c$ is,

$$V_M^{*-} = \sqrt{\Theta} (\sigma/\omega^{1/2}) \psi. \quad (47)$$

In Fig. 4, we compare liquidity risk for the three models studied in the text. Parameters used in the plot are $\psi = 1$, $\sigma = 1$, $\omega = 10$, and $\Theta = 0.3$. We also show the optimal points (V_P^*, τ_P^*) , (V_L^*, τ_L^*) , and (V_M^*, τ_M^*) . Notice how liquidity risk

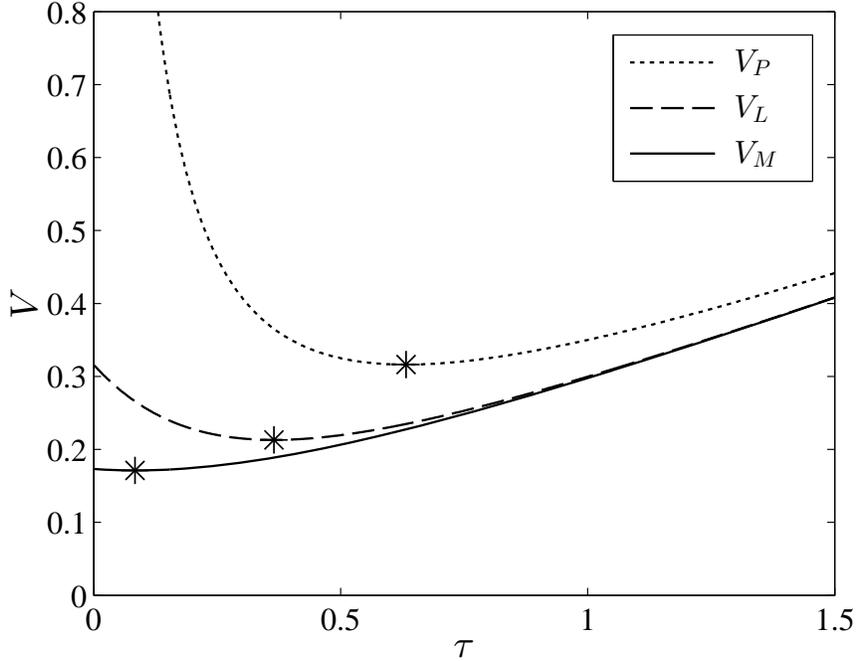


Figure 4: A comparison of liquidity risk, V , for the three models studied in the text. Parameters used in the plot are $\psi = 1$, $\sigma = 1$, $\omega = 10$, and $\Theta = 0.3$ ($\rho \approx 0.84$). The optimal points (V_P^*, τ_P^*) , (V_L^*, τ_L^*) , and (V_M^*, τ_M^*) are shown with asterisks.

decreases with the addition of the liquidity provider and reduces even further when the market security is added.

V Estimating the Optimal Trading Interval

Because the model's inputs (σ , ψ , Θ , and ω) are statistical properties of order flow and returns, we can use rough estimates of these parameters to determine the optimal clearing frequency of a typical U.S. stock. In this section, we do just that.

To standardize the calculation, we always use units of seconds in our estimates. As a consequence, the final estimate of the optimal trading interval is given in seconds.

- For the standard deviation of the security's value, we use $\psi = 0.0001$, which corresponds to an annualized volatility of approximately 25%.

- For the standard deviation of reservation prices, we use a typical percent spread, $\sigma = 0.0003$, which is equivalent to a \$0.01 quoted spread for a \$33 stock.
- For the correlation of the security with the market, we use $\rho = 0.75$ so that $\Theta = 0.4375$.
- For the order arrival rate, we use two different values based on reported Tape A/B quotation updates. There are approximately 15,000 quote changes per second for Tape A/B securities during trading hours. During peak times, this increases dramatically, to approximately 300,000 quote changes per second (see www.utpplan.com for both estimates). Because approximately 3,000 securities are reported on Tape A/B, we use the following estimates for the order arrival rate, $\omega = 5$ and $\omega_{peak} = 100$.

In the model with liquidity providers and the market security,

$$\tau^* = \left(\frac{2h_1(\Theta)}{\sqrt{3}} \right) \frac{\sigma/\omega^{1/2}}{\psi}. \quad (48)$$

Putting everything into the equation, we have

$$\tau^* \approx 0.9 \text{ seconds}, \quad (49)$$

and,

$$\tau_{peak}^* \approx 0.2 \text{ seconds}. \quad (50)$$

Our estimates suggest that a typical U.S. stock should trade at intervals of 0.2 to 0.9 seconds. Of course we do not wish to over-interpret this result. Individuals place limit orders in the market instead of price schedules, the market security is not really infinitely liquid, securities are correlated to many other securities in addition to the market security, and liquidity providers (rather than investors) represent a large fraction of the orders we use in our estimate of the order arrival rate. All of these points may require us to tweak the final value and we would place error bands of up

to an order of magnitude around this estimate.

VI Conclusions

U.S. markets have undergone considerable changes over the last two decades. Whereas before, trading was human mediated and quite slow (taking over half a minute for a market order to execute), it is now electronic, automated, and extremely fast (limited mainly by the speed of light). Market quality metrics have improved considerably as market speeds have increased, but it is unclear if the current milli- and microsecond environment is really necessary.

This paper attempts to determine the optimal speed of trading in financial markets. In our model of periodic market clearings, the optimal trading interval for a security depends on three factors: (1) the volatility of the security, (2) the intensity of trading in the security, and (3) the correlation of the security's value with other securities. All other things equal, a security should be traded more quickly if it is volatile, has intense trading, and is highly correlated with other securities.

When plugging in rough estimates of the model parameters for a typical U.S. stock, we calculate an optimal trading interval of 0.2 to 0.9 seconds. Delaying markets longer than these intervals is likely to harm market quality. However, in light of these estimates, for many securities it is hard to justify the extreme speeds at which U.S. markets operate.

REFERENCES

- Baruch, S. and G. Saar, 2009, Asset returns and listing choice of firms, *Review of Financial Studies*, 22, 2239-2274.
- Budish, E., P. Cramton and J. Shim, 2013, The High-Frequency Trading Arms Race: Frequent Batch Auctions as a Market Design Response, Working paper.
- Chan, K., A. Hameed and W. Kang, 2013, Stock price synchronicity and liquidity, *Journal of Financial Markets*, 16(3), 414-438.

Cochrane, J. H., 2013, Finance: Function Matters, Not Size, *Journal of Economic Perspectives*, 27(2), 29-50.

Farmer, J.D. and S. Skouras, 2012, Review of the Benefits of a Continuous Market vs. Randomised Stop Auctions and of Alternative Priority Rules (Policy Options 7 and 12). *UK Governments Foresight Project, The Future of Computer Trading in Financial Markets*, Economic Impact Assessment EIA11.

Garbade, K.D. and W.L. Silber, 1979, Structural organization of secondary markets: Clearing frequency, dealer activity and liquidity risk, *Journal of Finance*, 34(3), 577—593.

Gerig A. and D. Michayluk, 2014, Automated Liquidity Provision, Working paper.

Glosten, L.R. and P.R. Milgrom, 1985, Bid, ask and transaction prices in a specialist market with heterogeneously informed traders, *Journal of Financial Economics*, 14, 71-100.

Grossman, S.J. and M.H. Miller, 1988, Liquidity and Market Structure, *Journal of Finance*, 43(3).

Hirshleifer, J., 1971, The Private and Social Value of Information and the Reward to Incentive Activity, *American Economic Review*, 61(4), 561–574.

Kyle, A.S., 1985, Continuous auctions and insider trading, *Econometrica* 53, 1315-1335.

Madhavan, A., 1992, Trading Mechanisms in Securities Markets, *Journal of Finance*, 47(2), 601–641.

Stiglitz, J.E., 1989, Using Tax Policy to Curb Speculative Short-Term Trading, *Journal of Financial Services Research*, 3, 101–115.

APPENDIX

The Kalman Filter

The following is a straightforward application of the Kalman filter for the estimation of $\bar{\mathbf{m}}_t$ using contemporaneous and lagged values of $\bar{\mathbf{r}}_t$ (see Meinhold and Singpurwalla (1983)). The *observation equation* is,

$$\bar{\mathbf{r}}_t = \bar{\mathbf{m}}_t + \mathbf{f}_t, \quad (51)$$

$$\mathbf{f}_t \sim N(0, \Sigma). \quad (52)$$

and the *system equation* is,

$$\bar{\mathbf{m}}_t = \bar{\mathbf{m}}_{t-1} + \mathbf{e}_t, \quad (53)$$

$$\mathbf{e}_t \sim N(0, \Psi). \quad (54)$$

Denote by $\hat{\mathbf{m}}_t$ the estimate of $\bar{\mathbf{m}}_t$ based on $\{\bar{\mathbf{r}}_t, \bar{\mathbf{r}}_{t-1}, \bar{\mathbf{r}}_{t-2}, \dots\}$. It can be shown that,

$$P(\bar{\mathbf{m}}_t | \bar{\mathbf{r}}_t, \bar{\mathbf{r}}_{t-1}, \dots) \sim N(\hat{\mathbf{m}}_{t-1} + \mathbf{G}_t[\bar{\mathbf{r}}_t - \hat{\mathbf{m}}_{t-1}], \mathbf{S}_t), \quad (55)$$

$$P(\bar{\mathbf{m}}_{t+1} | \bar{\mathbf{r}}_t, \bar{\mathbf{r}}_{t-1}, \dots) \sim N(\hat{\mathbf{m}}_t, \mathbf{R}_{t+1}). \quad (56)$$

where \mathbf{G}_t is known as the *Kalman gain* and,

$$\mathbf{G}_t = \mathbf{R}_t(\mathbf{R}_t + \Sigma)^{-1}, \quad (57)$$

$$\mathbf{R}_{t+1} = \mathbf{S}_t + \Psi, \quad (58)$$

$$\mathbf{S}_t = \mathbf{R}_t - \mathbf{G}_t \mathbf{R}_t. \quad (59)$$

The best estimate of $\bar{\mathbf{m}}_t$ based on $\{\bar{\mathbf{r}}_t, \bar{\mathbf{r}}_{t-1}, \bar{\mathbf{r}}_{t-2}, \dots\}$ is just the mean of the distribution $P(\bar{\mathbf{m}}_t | \bar{\mathbf{r}}_t, \bar{\mathbf{r}}_{t-1}, \dots)$,

$$\hat{\mathbf{m}}_t = \hat{\mathbf{m}}_{t-1} + \mathbf{G}_t(\bar{\mathbf{r}}_t - \hat{\mathbf{m}}_{t-1}). \quad (60)$$

The estimation variance is

$$\text{Var}[\hat{\mathbf{m}}_t - \bar{\mathbf{m}}_t] = \mathbf{S}_t. \quad (61)$$

In general, the above equations are solved iteratively, starting at time zero. Here, we search for convergence of the estimation variance to a limiting value, i.e., we search for a solution when $\mathbf{R}_{t+1} = \mathbf{R}_t$. Rearranging the above equations and setting $\mathbf{R} = \mathbf{R}_{t+1} = \mathbf{R}_t$ produces the following equation,

$$\mathbf{R}(\mathbf{R} + \mathbf{\Sigma})^{-1}\mathbf{R} - \mathbf{\Psi} = 0, \quad (62)$$

which is a version of the discrete time algebraic Riccati equation. The conditions required for a solution to exist are discussed in Anderson and Moore (2005). Note that when \mathbf{R} has reached its steady state, that \mathbf{G} and \mathbf{S} will also be steady. Once \mathbf{R} is determined, then \mathbf{G} and \mathbf{S} can be calculated as follows,

$$\mathbf{G} = \mathbf{\Psi}\mathbf{R}^{-1}, \quad (63)$$

$$\mathbf{S} = \mathbf{R} - \mathbf{\Psi}. \quad (64)$$

Solving the Riccati Equation

In the model with a liquidity provider who does not have access to market information, all variables in the Kalman filter are scalars. Furthermore,

$$\mathbf{\Sigma} = \sigma^2/(\omega\tau), \quad (65)$$

$$\mathbf{\Psi} = \tau\psi^2. \quad (66)$$

The Riccati equation is therefore,

$$R^2/(R + \sigma^2/(\omega\tau)) - \tau\psi^2 = 0, \quad (67)$$

Solving for R and the rest of the variables in the Kalman filter,

$$R = (1/2) \left[\tau\psi^2 + \sqrt{\tau^2\psi^4 + 4\psi^2\sigma^2/\omega} \right], \quad (68)$$

$$G = \frac{2\tau\psi^2}{\tau\psi^2 + \sqrt{\tau^2\psi^4 + 4\psi^2\sigma^2/\omega}}, \quad (69)$$

$$S = (1/2) \left[\sqrt{\tau^2\psi^4 + 4\psi^2\sigma^2/\omega} - \tau\psi^2 \right], \quad (70)$$

In the model with a liquidity provider who has access to market information, we have,

$$\Sigma = \begin{pmatrix} \sigma^2/(\omega\tau) & 0 \\ 0 & 0 \end{pmatrix} \quad \Psi = \begin{pmatrix} \tau\psi^2 & \rho\tau\psi\psi_M \\ \rho\tau\psi\psi_M & \tau\psi_M^2 \end{pmatrix}. \quad (71)$$

Solving the Riccati equation,

$$\mathbf{R} = \begin{pmatrix} (1/2) \left[(2 - \Theta)\tau\psi^2 + \Theta\tau\psi^2 \sqrt{1 + \frac{4\sigma^2/\omega}{\Theta\tau^2\psi^2}} \right] & \rho\tau\psi\psi_M \\ \rho\tau\psi\psi_M & \tau\psi_M^2 \end{pmatrix}, \quad (72)$$

$$\mathbf{G} = \begin{pmatrix} \frac{2}{-1 + \sqrt{1 + \frac{4\sigma^2/\omega}{\Theta\tau^2\psi^2}}} & \frac{(-1 + \sqrt{1 + \frac{4\sigma^2/\omega}{\Theta\tau^2\psi^2}}) \rho\tau\psi\psi_M}{\left(1 + \sqrt{1 + \frac{4\sigma^2/\omega}{\Theta\tau^2\psi^2}}\right) \tau\psi_M^2} \\ 0 & 1 \end{pmatrix}, \quad (73)$$

$$\mathbf{S} = \begin{pmatrix} (1/2) \left[\Theta\tau\psi^2 \left(-1 + \sqrt{1 + \frac{4\sigma^2/\omega}{\Theta\tau^2\psi^2}} \right) \right] & 0 \\ 0 & 0 \end{pmatrix} \quad (74)$$

where $\Theta \equiv 1 - \rho^2$. Note that when the security is uncorrelated with the market, i.e., $\Theta = 1$, that the elements $\mathbf{R}_{(1,1)}$, $\mathbf{G}_{(1,1)}$, and $\mathbf{S}_{(1,1)}$ all reduce to the values found in the case when the liquidity provider has no market information (Eqs. 68-70).

Liquidity Risk

The equation for the liquidity risk of an investor trading the security when a liquidity provider is present is,

$$V_{L,M} = \text{Var}[(\hat{m}_t - \bar{m}_t) + (\bar{m}_t - m_i)], \quad (75)$$

$$= \text{Var}[\hat{m}_t - \bar{m}_t] + \text{Var}[\bar{m}_t - m_i] + 2 \text{Cov}[\hat{m}_t - \bar{m}_t, \bar{m}_t - m_i]. \quad (76)$$

We will start with the first term, $\text{Var}[(\hat{m}_t - \bar{m}_t)]$. The estimation variance of $\bar{\mathbf{m}}_t$ is just \mathbf{S} (see Eq.61). For the security, the variance is reported at position (1, 1),

$$\text{Var}[\hat{m}_t - \bar{m}_t] = \mathbf{S}_{(1,1)}. \quad (77)$$

The second term is derived in the text (Eq. 16),

$$\text{Var}[\bar{m}_t - m_i] = (1/2) [\phi^3 + (1 - \phi)^3] \tau \psi^2, \quad (78)$$

$$= \phi_1 \tau \psi^2. \quad (79)$$

where $\phi_1 \equiv (1/2) [\phi^3 + (1 - \phi)^3]$.

The third term, $2\text{Cov}[\hat{m}_t - \bar{m}_t, \bar{m}_t - m_i]$, can be derived as follows. Subtracting $\bar{\mathbf{m}}_t$ from both sides of Eq. 60 and rearranging,

$$\hat{\mathbf{m}}_t - \bar{\mathbf{m}}_t = (\mathbf{I} - \mathbf{G}_t)(\hat{\mathbf{m}}_{t-1} - \bar{\mathbf{m}}_{t-1}) + \mathbf{G}_t(\bar{\mathbf{r}}_t - \bar{\mathbf{m}}_t) + (\mathbf{G}_t - \mathbf{I})(\bar{\mathbf{m}}_t - \bar{\mathbf{m}}_{t-1}), \quad (80)$$

where \mathbf{I} is the identity matrix. The elements in the vectors $(\mathbf{I} - \mathbf{G}_t)(\hat{\mathbf{m}}_{t-1} - \bar{\mathbf{m}}_{t-1})$ and $\mathbf{G}_t(\bar{\mathbf{r}}_t - \bar{\mathbf{m}}_t)$ are uncorrelated with $(\bar{m}_t - m_i)$ so we can disregard them. In the last vector, $(\mathbf{G}_t - \mathbf{I})(\bar{\mathbf{m}}_t - \bar{\mathbf{m}}_{t-1})$, the relevant contribution to $\hat{m}_t - \bar{m}_t$ is the first element,

$$(\mathbf{G}_{(1,1)} - 1)(\bar{m}_t - \bar{m}_{t-1}) + \mathbf{G}_{(1,2)}(\bar{m}_{M,t} - \bar{m}_{M,t-1}). \quad (81)$$

The covariance of the random terms in this equation with $(\bar{m}_t - m_i)$ are,

$$\text{Cov}[\bar{m}_t - \bar{m}_{t-1}, \bar{m}_t - m_i] = \phi_2 \tau \psi^2, \quad (82)$$

$$\text{Cov}[\bar{m}_{M,t} - \bar{m}_{M,t-1}, \bar{m}_t - m_i] = \phi_2 \rho \tau \psi \psi_M. \quad (83)$$

where $\phi_2 \equiv (1/4) [\phi^3 + 2(1 - \phi)^3 + 3(1 - \phi)\phi^2]$. The structure of ϕ_2 can be derived by noting the covariance of the difference of averaged points of a Brownian motion with the difference of an averaged point and a particular point of the same Brownian motion. The result is left for the reader to verify.

Putting everything together, we have,

$$V_{L,M} = \text{Var}[(\hat{m}_t - \bar{m}_t) + (\bar{m}_t - m_i)], \quad (84)$$

$$= \text{Var}[\hat{m}_t - \bar{m}_t] + \text{Var}[\bar{m}_t - m_i] + 2 \text{Cov}[\hat{m}_t - \bar{m}_t, \bar{m}_t - m_i], \quad (85)$$

$$= \mathbf{S}_{(1,1)} + \phi_1 \tau \psi^2 + 2(\mathbf{G}_{(1,1)} - 1) \phi_2 \tau \psi^2 + 2\mathbf{G}_{(1,2)} \phi_2 \rho \tau \psi \psi_M, \quad (86)$$