

# Smooth Trading with Overconfidence and Market Power

ALBERT S. KYLE

*University of Maryland*

ANNA A. OBIZHAEVA

*New Economic School*

and

YAJUN WANG

*University of Maryland*

*First version received October 2013; Editorial decision February 2017; Accepted March 2017 (Eds.)*

We describe a symmetric continuous-time model of trading among relatively overconfident, oligopolistic informed traders with exponential utility. Traders agree to disagree about the precisions of their continuous flows of Gaussian private information. The price depends on a trader's inventory (permanent price impact) and the derivative of a trader's inventory (temporary price impact). More disagreement makes the market more liquid; without enough disagreement, there is no trade. Target inventories mean-revert at the same rate as private signals. Actual inventories smoothly adjust towards target inventories at an endogenous rate which increases with disagreement. Faster-than-equilibrium trading generates "flash crashes" by increasing temporary price impact. A "Keynesian beauty contest" dampens price fluctuations.

*Key words:* Market microstructure, Price impact, Liquidity, Transaction costs, Double auctions, Information aggregation, Rational expectations, Agreement to disagree, Imperfect competition, Keynesian beauty contest, Overconfidence, Strategic trading, Dynamic trading, Flash crash.

*JEL Codes:* D8, D43, D47, G02, G14

## 1. INTRODUCTION

When large traders in financial markets seek to profit from perishable private information, they face a fundamental tradeoff. On the one hand, they want to trade slowly, to reduce their own temporary price impact costs resulting from adverse selection. On the other hand, they want to trade quickly, before the permanent price impact of competitors trading on similar information makes profit opportunities go away. We illustrate this tradeoff using a stationary model of continuous trading among oligopolistic traders who agree to disagree about the precisions of private signals. The equilibrium with smooth trading reveals important insights about dynamic properties of inventories, prices, and liquidity.

The model combines the following assumptions: (1) There is one type of trader, a strategic informed trader; there are no noise traders or market makers. (2) Each trader has a flow of

private information about the fundamental value; the noise in their signals is uncorrelated. (3) Traders are relatively overconfident in that each trader believes his private information is more precise than other traders believe it to be. (4) Each trader applies Bayes law correctly; in doing so, he infers from prices the economically relevant aggregation of other traders' information. (5) Traders trade strategically, correctly taking into account how the permanent and temporary price impact of their trades affects prices. (6) Random variables are jointly normally distributed. (7) Traders are symmetric in that they have the same additive exponential utility functions and symmetrically different beliefs about the information structure. (8) All state variables have stationary distributions.

Disagreement about the precision of private signals motivates trade. This differs from the models of Vayanos (1999) and Du and Zhu (2017), who motivate trade in a common prior setting by shocks to inventories and to private values, respectively.

The one-period version of our model is an equilibrium in demand curves. An equilibrium with linear trading strategies and positive trading volume exists if and only if each trader believes that his signal is slightly more than twice as accurate as other traders' signals. The equilibrium has a simple closed-form solution. As disagreement falls, liquidity dries up and trade vanishes.

The continuous-time model implements a continuous auction in which traders continuously submit demand schedules. An "almost-closed-form" steady-state equilibrium is characterized by six endogenous parameters which solve a set of six polynomial equations. Numerical calculations indicate that the same existence condition holds in the continuous-time model as in the one-period model.

### 1.1. *Inventories*

Our stationary model provides a realistic description of trading by large asset managers who exploit private information about securities. In the equilibrium, inventories follow a partial adjustment process with coefficients implied by the model's deep parameters. Each trader calculates a target inventory based on how his own estimate of the long-term dividend growth rate differs from the estimates of other traders. We prove analytically that the half-life of traders' target inventories matches the half-life of private signals; both decay at a rate equal to the sum of the natural mean reversion rate of dividend growth and the total precision of all information flowing into the market.

Since the market offers no instantaneous liquidity for block trades, each trader "shreds orders" and only partially adjusts his inventory in the direction of a target inventory, so that actual inventories are differentiable or "smooth" functions of time.<sup>1</sup> We obtain additional robust results numerically. The endogenous speed with which actual inventories move towards target inventories is faster when signals decay faster and when there is more disagreement, which makes markets more liquid. Contrary to the common intuition that high trading volume results from a focus on short-term quarterly earnings announcements, all trading volume is informative about long-term value.

We show analytically that when traders' beliefs are "correct on average", a more liquid market tends to be associated with a lower autocorrelation of actual inventories but a higher contemporaneous correlation of actual inventories with target inventories. Hasbrouck and Sofianos (1993), Madhavan and Smidt (1993), and Hendershott and Menkveld (2014) find that intermediaries' inventories adjust rapidly towards time-varying targets and tend to have higher autocorrelations

1. The market clears in time derivatives of inventories. Our informal use of the term "smooth trading" is different from the mathematical usage, which implies derivatives of all orders exist. Since the first derivatives of traders' inventories follow diffusions, higher order derivatives do not exist.

and lower mean-reversion rates in smaller and less-frequently-traded stocks, about which less information is likely to be available. Even though our model has no separate category of intermediaries, its implications are consistent with these findings.

## 1.2. *Liquidity*

Our model generates a clean distinction between endogenous permanent and temporary price impact. From a trader's perspective, the level of prices is a linear function of his level of inventories, the derivative of his inventories, and other traders' expectations of fundamental value. Trading costs therefore depend on two liquidity parameters. First, a permanent price impact parameter, denoted  $\lambda$  as in Kyle (1985), measures the price impact of a change in the level of inventories. Second, a temporary price impact parameter, denoted  $\kappa$ , measures the price impact of a change in the derivative of inventories. While permanent price impact is commonly used in microstructure models, temporary price impact only appears in settings where all traders, including noise traders, smooth out their trading. The temporary component makes trading a given quantity over a shorter horizon more expensive than trading the same quantity over a longer horizon; the market offers no instantaneous liquidity for block trades.

Black (1971) describes liquidity using the concepts of tightness, depth, and resiliency. In our continuous-time model, the market has no instantaneous depth, tightness is related to temporary price impact, and resiliency depends on the aggregate rate of information production. These concepts of liquidity play out differently from Kyle (1985), in which the equilibrium would break if noise traders—like the informed trader—were also allowed to smooth their trading; when all traders smooth their trading, the nature of liquidity changes significantly.

Since the market clears both in inventories and in time derivatives of inventories, continuous time makes the distinction between permanent and temporary price impact intuitively and mathematically clear. In the discrete-time setup of Vayanos (1999), an analogous distinction between permanent and temporary price impact is obtained in the limit as the interval between rounds of trading goes to zero.

The speed with which actual inventories move towards target inventories results from a tradeoff between temporary price impact costs and the speed with which signals decay. We show numerically that increasing disagreement makes markets more liquid and increases the speed of trading. The smooth trading model therefore realistically predicts that high-volume markets will be highly liquid.

Our use of the terms “temporary” and “permanent” price impact differs from that of empirical researchers who think of temporary impact as short-term negative autocorrelation in returns arising from dealer spreads (‘bid-ask bounce’) and permanent impact as persistent (martingale) price changes arising from private information being impounded into market prices. As a result of traders' optimizing behaviour, higher trading costs show up as more gradual changes in inventories, not as more short-term mean reversion in prices. In principle, price impact can be inferred from abnormally fast “out-of-equilibrium” execution of a bet, which leads to a price spike resembling a “flash crash”.

Our price impact model, derived *endogenously* from equilibrium trading, is similar to empirical, practitioner-oriented transaction-cost models exogenously assumed by Grinold and Kahn (1994), Almgren and Chriss (2000), and Obizhaeva and Wang (2013). Our model provides a theoretical explanation for robust empirical findings that faster trading increases transaction costs by increasing temporary price changes, as documented by Holthausen *et al.* (1990), Chan and Lakonishok (1995), Keim and Madhavan (1997), and Dufour and Engle (2000). Brunnermeier and Pedersen (2005), Carlin *et al.* (2007), and Longstaff (2001) examine the

economic implications of fast trading given *exogenously* specified price impact functions depending on the speed of trading.

### 1.3. *Prices*

Even though traders adjust inventories slowly, prices immediately reflect all of the information in the market, both public and private. Each trader infers the average valuation of other traders from the price.

There is a “beauty contest”, like Keynes (1936), because traders forecast how the expectations of other traders will evolve in the future, and their trading to take advantage of these forecasts influences prices. We obtain numerically the interesting result that prices are dampened due to this beauty contest. The growth-rate component of prices is a weighted average of the growth-rate expectations of each trader; “dampening” means that the weights sum to a constant less than one. Here is the intuition: When prices are high and a trader believes that the high prices reflect fundamental value, he believes that other traders, who overweight their signals, will revise their forecasts down so that it is profitable to sell ahead of such revisions in the short run. Dampened price fluctuations lead to momentum (positive autocorrelation) in returns; see Section 4.3 for a more detailed explanation. Dampening is more pronounced when disagreement is larger and markets are more liquid. This explains the otherwise puzzling empirical finding of Lee and Swaminathan (2000), Moskowitz *et al.* (2012), and Cremers and Pareek (2015) that momentum is more pronounced in high-volume and liquid securities.

### 1.4. *Alternative models*

We also characterize equilibrium in an otherwise similar model of perfect competition like Kyle and Lin (2001). With perfect competition, traders adjust holdings to target inventories infinitely fast; markets are more liquid. Consistent with the intuition that low trading costs amplify the economic importance of the dampening effect, perfect competition leads to more pronounced dampening than imperfect competition.

We also examine an otherwise similar model with privately observed shocks to private values and a common prior. Analytical tractability requires assuming that shocks to private values mean revert at the same rate as private signals. This model has properties analogous to our preferred model of overconfidence in all respects except that price dampening goes away. Prices are equal to an average of traders’ private valuations, adjusted for private values. Price dampening does not occur in the model of Du and Zhu (2017), which has non-stationary private values; the model of Vayanos (1999), which has endowment shocks; the model of Banerjee and Kremer (2010), which has myopic traders; or rational expectations models such as Wang (1993), Wang (1994), and He and Wang (1995), in which noise affects the weights on signals but the weights on valuations sum to one. We infer that price dampening in the Keynesian beauty contest results from a combination of overconfidence and substantial market liquidity, not from noise trading or private values with a common prior.

Harsanyi (1976) conjectures that a model without a common prior can be mapped into an isomorphic model with a common prior, therefore making models with different priors unnecessary. Obtaining price dampening with a common prior would likely require complicated *ad hoc* assumptions with externalities related to auto- and cross-correlations of private values. Disagreement generates both trading volume and price dampening while satisfying Ockham’s razor.

This article is structured as follows. Section 2 presents a one-period model. Section 3 presents the continuous-time model. Section 4 examines properties of the smooth-trading equilibrium.

Section 5 concludes. Proofs are in Appendix A. Appendix B presents a similar model of competitive trading. Supplementary Appendix C presents a similar model in which private values and a common prior replace overconfidence.

## 2. ONE-PERIOD MODEL

The one-period model has a simple closed-form solution illustrating the interaction between overconfidence and market power.

A risky asset with random liquidation value  $v \sim N(0, 1/\tau_v)$  is traded for a safe numeraire asset. It is common knowledge that the asset is in zero net supply. Trader  $n$  is endowed with a privately observed inventory  $S_n$  with  $\sum_{n=1}^N S_n = 0$ . While initial inventories play no significant role in this one-period model, they help map results into the continuous-time model. Traders observe signals about the normalized liquidation value  $\tau_v^{1/2} v \sim N(0, 1)$ . All traders observe a public signal  $i_0 := \tau_0^{1/2} (\tau_v^{1/2} v) + e_0$  with  $e_0 \sim N(0, 1)$ . Each trader  $n$  observes a private signal  $i_n := \tau_n^{1/2} (\tau_v^{1/2} v) + e_n$  with  $e_n \sim N(0, 1)$ . The asset payoff  $v$ , the public signal error  $e_0$ , and  $N$  private signal errors  $e_1, \dots, e_N$  are independently distributed.

Traders agree about the precision of the public signal  $\tau_0$  and agree to disagree about the precisions of private signals  $\tau_n$ . Each trader is relatively overconfident, believing his own signal to have a high precision  $\tau_n = \tau_H$  and other traders' signals to have low precisions  $\tau_m = \tau_L$  for  $m \neq n$ , with  $\tau_H > \tau_L \geq 0$ . Each trader believes other traders are like noise traders who overtrade on their information. There are no explicit noise traders or market makers. The model is like Treynor (1995), who discusses “transactors acting on information which they believe has not yet been fully discounted in the market price but which in fact has”. Similarly, Black (1986) defines noise trading as “trading on noise as if it were information”.

Each trader submits a demand schedule  $X_n(p) := X_n(i_0, i_n, S_n, p)$  to a single-price auction. An auctioneer clears the market at price  $p := p[X_1, \dots, X_N]$ . Trader  $n$ 's terminal wealth is

$$W_n := v(S_n + X_n(p)) - pX_n(p). \quad (1)$$

Each trader  $n$  maximizes the same expected exponential utility function of wealth  $\mathbb{E}^n[-e^{-AW_n}]$  using his own beliefs to calculate the expectation.

An *equilibrium* is a set of trading strategies  $X_1, \dots, X_N$  such that each trader's strategy maximizes his expected utility, taking as given the trading strategies of other traders. Except for the assumption that traders do not share a common prior, this is equivalent to a Bayesian Nash equilibrium. As imperfect competitors, traders take into account how the price  $p$  depends on the quantities they trade.

### 2.1. Linear strategies and Bayesian updating

Let  $i_{-n} := \frac{1}{N-1} \sum_{m=1, m \neq n}^N i_m$  denote the average of other traders' signals. When trader  $n$  conjectures that other traders submit symmetric linear demand schedules

$$X_m(i_0, i_m, S_m, p) = \alpha i_0 + \beta i_m - \gamma p - \delta S_m, \quad m = 1, \dots, N, \quad m \neq n, \quad (2)$$

he infers from the market-clearing condition

$$x_n + \sum_{\substack{m=1 \\ m \neq n}}^N (\alpha i_0 + \beta i_m - \gamma p - \delta S_m) = 0 \quad (3)$$

that his residual supply schedule  $P(x_n)$  is a function of his quantity  $x_n$  given by

$$P(x_n) = \frac{\alpha}{\gamma} i_0 + \frac{\beta}{\gamma} i_{-n} + \frac{\delta}{(N-1)\gamma} S_n + \frac{1}{(N-1)\gamma} x_n. \quad (4)$$

Since trader  $n$  observes the public signal  $i_0$ , his own inventory  $S_n$ , and the quantity he trades himself  $x_n$ , he can infer the average of other traders' signals  $i_{-n}$  from observing the intercept of his residual supply schedule.

Let  $\mathbb{E}^n[\dots]$  and  $\mathbb{V}\text{ar}^n[\dots]$  denote trader  $n$ 's expectation and variance operators conditional on all signals  $i_0, i_1, \dots, i_N$ . Define "total precision"  $\tau$  by

$$\tau := (\mathbb{V}\text{ar}^n[v])^{-1} = \tau_v (1 + \tau_0 + \tau_H + (N-1)\tau_L). \quad (5)$$

The projection theorem for jointly normally distributed random variables implies

$$\mathbb{E}^n[v] = \frac{\tau_v^{1/2}}{\tau} \left( \tau_0^{1/2} i_0 + \tau_H^{1/2} i_n + (N-1)\tau_L^{1/2} i_{-n} \right). \quad (6)$$

## 2.2. Utility maximization with market power

Conditional on all information, trader  $n$ 's terminal wealth  $W_n$  is a normally distributed random variable with mean and variance given by

$$\mathbb{E}^n[W_n] = \mathbb{E}^n[v](S_n + x_n) - P(x_n)x_n, \quad \mathbb{V}\text{ar}^n[W_n] = (S_n + x_n)^2 \mathbb{V}\text{ar}^n[v]. \quad (7)$$

Normal distributions imply that expected utility is given by

$$\mathbb{E}^n[-e^{-A W_n}] = -\exp\left(-A \mathbb{E}^n[W_n] + \frac{1}{2} A^2 \mathbb{V}\text{ar}^n[W_n]\right). \quad (8)$$

Maximizing this function is equivalent to maximizing  $\mathbb{E}^n[W_n] - \frac{1}{2} A \mathbb{V}\text{ar}^n[W_n]$ . Plugging equations (5), (6), and (7) into equation (8), trader  $n$  solves the maximization problem

$$\max_{x_n} \left[ \frac{\tau_v^{1/2}}{\tau} \left( \tau_0^{1/2} i_0 + \tau_H^{1/2} i_n + (N-1)\tau_L^{1/2} i_{-n} \right) (S_n + x_n) - P(x_n)x_n - \frac{A}{2\tau} (S_n + x_n)^2 \right]. \quad (9)$$

Oligopolistic trader  $n$  exercises market power by taking into account how his quantity  $x_n$  affects the price  $P(x_n)$  on his residual supply schedule (4).

## 2.3. Equilibrium with linear demand schedules

There always exists a no-trade equilibrium in which each trader submits a no-trade schedule  $X_n(\cdot) \equiv 0$  and the auctioneer cannot establish a meaningful price.

An equilibrium with trade may also exist. Appendix A.1 proves the following theorem using the "no-regret" approach: Each trader observes his residual linear supply schedule, infers the average of other traders' signals from its intercept, picks the optimal quantity  $x_n$ , and implements this choice with a demand schedule  $x_n = X_n(i_0, i_n, S_n, p)$ , without observing the residual supply schedule itself.

Let  $\tau_H/\tau_L$  measure "disagreement". Define the exogenous quantity  $\Delta_H$  by

$$\Delta_H := \frac{\tau_H^{1/2}}{\tau_L^{1/2}} - 2 - \frac{2}{(N-2)}. \quad (10)$$

**Theorem 1. (Characterization of Equilibrium in the One-Period Model with Overconfidence and Imperfect Competition).** *There exists a unique symmetric equilibrium with linear trading strategies and nonzero trade if and only if the second-order condition  $\Delta_H > 0$  holds. The equilibrium satisfies the following:*

(1) *Trader  $n$  trades the quantity  $x_n^*$  given by*

$$x_n^* = \frac{(N-2)\tau_L^{1/2}\Delta_H}{AN} \tau_v^{1/2} (i_n - i_{-n}) - \delta S_n, \quad (11)$$

*where the inventory adjustment factor  $\delta$  is*

$$0 < \delta = \frac{(N-2)\tau_H^{1/2} - 2(N-1)\tau_L^{1/2}}{(N-1)(\tau_H^{1/2} - \tau_L^{1/2})} < 1. \quad (12)$$

(2) *The price  $p^*$  is the average of traders' valuations:*

$$p^* = \frac{1}{N} \sum_{n=1}^N \mathbb{E}^n[v] = \frac{\tau_v^{1/2}}{\tau} \left( \tau_0^{1/2} i_0 + \frac{\tau_H^{1/2} + (N-1)\tau_L^{1/2}}{N} \sum_{n=1}^N i_n \right). \quad (13)$$

(3) *The parameters  $\alpha > 0$ ,  $\beta > 0$ , and  $\gamma > 0$ , defining the linear trading strategies in equation (2), have unique closed-form solutions defined in (A6).*

For an equilibrium with positive trading volume to exist, there must be enough disagreement so that  $\Delta_H > 0$ . This requires  $N \geq 3$  and requires  $\tau_H^{1/2}$  to be sufficiently more than twice as large as  $\tau_L^{1/2}$ . Each trader trades in the direction of his private signal  $i_n$ , trades against the average of other traders' signals  $i_{-n}$ , and hedges a fraction  $\delta$  of his initial inventory. Trading volume increases in disagreement and decreases in risk aversion. Equation (13) implies that the equilibrium price is a weighted average of traders' valuations with weights summing to one. As shown in Section 3, the weights sum to less than one in a continuous-time model.

#### 2.4. Equilibrium properties

As in Kyle (1989) and Rostek and Weretka (2012), each trader exercises market power by shading the quantity traded relative to the quantity a perfect competitor would trade. Define a trader's "target inventory"  $S_n^{TI}$  as the inventory such that he would not want to trade ( $x_n^* = 0$ ). From equation (11), it is equal to

$$S_n^{TI} = \frac{1}{A} \left( 1 - \frac{1}{N} \right) \tau_v^{1/2} (\tau_H^{1/2} - \tau_L^{1/2}) (i_n - i_{-n}). \quad (14)$$

Then trader  $n$ 's optimal quantity traded can be written

$$x_n^* = \delta (S_n^{TI} - S_n). \quad (15)$$

The parameter  $\delta$ , defined in equation (12), is the fraction by which traders adjust positions towards target levels. As a function of disagreement  $\tau_H/\tau_L$ ,  $\delta$  increases monotonically from a lower bound of zero when the existence condition  $\tau_H^{1/2}/\tau_L^{1/2} - 2 - 2/(N-2) > 0$  is barely satisfied towards an



upper bound of  $(N-2)/(N-1)$  as  $\tau_H^{1/2}/\tau_L^{1/2} \rightarrow \infty$ . If there is not enough disagreement to sustain an equilibrium with trade, each trader would want to shade his bid more than the others, and this breaks the equilibrium.

Consider an otherwise equivalent one-period model with perfect competition. Appendix B.1 proves the following result.

**Theorem 2. (Characterization of Equilibrium in the One-Period Model with Overconfidence and Perfect Competition).** *Assume  $\tau_H > \tau_L$ . Then there exists a unique symmetric equilibrium with linear trading strategies and nonzero trade, which has the following properties:*

- (1) *Trader  $n$  chooses the quantity  $x_n^* = S_n^{TI} - S_n$  (equation (15) with  $\delta = 1$ ).*
- (2) *The price  $p^*$  is the same as with imperfect competition (equation (13)).*

The existence condition for the competitive equilibrium ( $\tau_H > \tau_L$ ) is less restrictive than with imperfect competition ( $\Delta_H > 0$ ), even in the limit  $N \rightarrow \infty$ , because an imperfectly competitive trader remains a monopolist over his private signal.

From the perspective of trader  $n$ , equation (4) implies that with imperfect competition, price impact can be written as a function of both  $x_n$  and  $S_n$ ,

$$P(x_n, S_n) := p_{0,n} + \lambda S_n + \kappa x_n, \quad (16)$$

where  $p_{0,n}$  is a linear combination of random variables  $i_0$  and  $i_{-n}$ , and equations (A5) and (A6) imply that constants  $\lambda$  and  $\kappa$  are given by

$$\lambda := \frac{\delta}{(N-1)\gamma} = \frac{A}{\tau} \frac{\tau_H^{1/2} + (N-1)\tau_L^{1/2}}{(N-1)(\tau_H^{1/2} - \tau_L^{1/2})}, \quad (17)$$

$$\kappa := \frac{\lambda}{\delta} = \frac{1}{(N-1)\gamma} = \frac{A}{\tau} \frac{\tau_H^{1/2} + (N-1)\tau_L^{1/2}}{(N-2)\tau_L^{1/2} \Delta_H}. \quad (18)$$

The price impact parameters  $\lambda$  and  $\kappa$  increase in risk aversion  $A$  and decrease in disagreement  $\tau_H/\tau_L$ ; these results are consistent with the continuous-time model. In the continuous-time model, the first component  $\lambda S_n$  measures permanent price impact as in Kyle (1985). The second component  $\kappa x_n$  measures temporary price impact determined by the speed of trading, with  $x_n$  replaced by the derivative of the trader's inventory  $dS_n/dt$ . We next discuss the continuous-time model.

### 3. CONTINUOUS-TIME MODEL

There are  $N$  risk-averse oligopolistic traders who trade at price  $P(t)$  a risky asset in zero net supply against a risk-free asset which earns constant risk-free rate  $r > 0$ .

The risky asset pays out dividends at continuous rate  $D(t)$ . Dividends follow a stochastic process with mean-reverting stochastic growth rate  $G^*(t)$ , constant instantaneous volatility  $\sigma_D > 0$ , and constant rate of mean reversion  $\alpha_D > 0$ :

$$dD(t) := -\alpha_D D(t)dt + G^*(t)dt + \sigma_D dB_D(t). \quad (19)$$

The dividend  $D(t)$  is publicly observable, but the growth rate  $G^*(t)$  is not observed by any trader. The growth rate  $G^*(t)$  follows an AR-1 process with mean reversion  $\alpha_G$  and volatility  $\sigma_G$ :

$$dG^*(t) := -\alpha_G G^*(t)dt + \sigma_G dB_G(t). \quad (20)$$



Each trader  $n$  observes a continuous stream of private information  $I_n(t)$  defined by the stochastic process

$$dI_n(t) := \tau_n^{1/2} \frac{G^*(t)}{\sigma_G \Omega^{1/2}} dt + dB_n(t), \quad n = 1, \dots, N. \quad (21)$$

Since its drift is proportional to  $G^*(t)$ , each increment  $dI_n(t)$  in the process  $I_n(t)$  is a noisy observation of  $G^*(t)$ . The denominator  $\sigma_G \Omega^{1/2}$  scales  $G^*(t)$  so that its conditional variance is one; this simplifies the intuitive interpretation of the model. The precision parameter  $\tau_n$  measures the informativeness of the signal  $dI_n(t)$  as a signal-to-noise ratio describing how fast new information flows into the market. The parameter  $\Omega$  measures the steady-state error variance of the trader's estimate of  $G^*(t)$  in units of time; it is defined algebraically below (see equation (25)).<sup>2</sup>

As in the one-period model, each trader is certain that his own private information  $I_n(t)$  has high precision  $\tau_n = \tau_H$  and the other traders' private information has low precision  $\tau_m = \tau_L$  for  $m \neq n$ , with  $\tau_H > \tau_L \geq 0$ . Traders do not update their dogmatic beliefs about  $\tau_H$  and  $\tau_L$  over time; for plausible parameter values, it would take a long time for a trader to learn that his beliefs are incorrect. Since relatively overconfident traders agree to disagree about the precisions of their private information, they do not share a common prior even though their beliefs are common knowledge.<sup>3</sup> Agreement to disagree is a simple assumption with realistic implications: it can naturally break no-trade results and generate trading volume.

Each trader's information set at time  $t$ , denoted  $\mathcal{F}_n(t)$ , consists of the histories of (1) the dividend process  $D(s)$ , (2) the trader's own private information  $I_n(s)$ , and (3) the market price  $P(s)$ ,  $s \in (-\infty, t]$ . All traders process information rationally; they apply Bayes law correctly given their possibly incorrect beliefs.

Let  $S_n(t)$  denote the inventory of trader  $n$  at time  $t$ . Zero net supply implies  $\sum_{n=1}^N S_n(t) = 0$ .

We only consider "smooth trading" equilibria in which inventories  $S_n(t)$  are differentiable functions of time. Trading strategies and market clearing are specified using rates of trading, not shares traded. Trader  $n$ 's trading strategy  $X_n$  is a mapping from his information set  $\mathcal{F}_n(t)$  at time  $t$  into a flow-demand schedule which defines the derivative of his inventory  $x_n(t) := dS_n(t)/dt$  ('trading intensity') as a function of the market-clearing price  $P(t)$  with  $x_n(t) = X_n(t, P(t); \mathcal{F}_n(t))$ . An auctioneer continuously calculates the market-clearing price  $P(t) := P[X_1, \dots, X_N](t)$  such that the market-clearing condition  $\sum_{n=1}^N x_n(t) = 0$  is satisfied. Each trader explicitly takes into account the effect of his trading intensity on market prices.

Each trader has the same time preference parameter  $\rho$  and the same time-additively-separable exponential utility function  $U(c_n(s)) := -e^{-A c_n(s)}$  with constant-absolute-risk-aversion parameter  $A$ . Trader  $n$ 's consumption strategy  $C_n$  defines a consumption rate  $c_n(t) := C_n(t; \mathcal{F}_n(t))$  for all  $t > -\infty$ . Let  $\mathbb{E}_t^n[\dots]$  denote the conditional expectations operator  $\mathbb{E}[\dots | \mathcal{F}_n(t)]$  based on trader  $n$ 's beliefs.

Define an *equilibrium* as a set of trading strategies  $X_1^*, \dots, X_N^*$  and consumption strategies  $C_1^*, \dots, C_N^*$  such that, for  $n = 1, \dots, N$ , trader  $n$ 's optimal consumption and trading strategies  $X_n = X_n^*$  and  $C_n = C_n^*$  solve his maximization problem taking as given the optimal strategies of the other

2. Since the innovation variance of  $dI_n(t)$  can be estimated arbitrarily precisely by observing past information continuously, it is common knowledge that the innovation variance is one. Scaling the innovation variance of  $I_n(t)$  in equation (21) to make it equal to one is therefore a normalization without loss of generality. See footnote 11 for further discussion.

3. We call this belief structure "relative overconfidence" to distinguish it from a belief structure with "absolute overconfidence" in which traders believe the precisions of their information is greater than empirically true precisions. Empirically true precisions do not affect the equilibrium strategies but do affect empirical predictions about asset returns (see Section 4.5).

traders. For all dates  $t > -\infty$ , the optimal strategies  $X_n^*$  and  $C_n^*$  solve trader  $n$ 's maximization problem

$$\max_{[C_n, X_n]} \mathbb{E}_t^n \left[ \int_{s=t}^{\infty} e^{-\rho(s-t)} U(c_n(s)) ds \right], \quad (22)$$

where inventories satisfy  $dS_n(t) = x_n(t)dt$  and money holdings  $M_n(t)$  satisfy

$$dM_n(t) = (rM_n(t) + S_n(t)D(t) - c_n(t) - P(t)x_n(t))dt. \quad (23)$$

When solving the maximization problem, trader  $n$  takes as given the trading strategies  $X_m$ ,  $m \neq n$ , for the other  $N - 1$  traders; he exercises market power by taking into account how his own strategy affects equilibrium prices  $P(t)$  and future trading opportunities. Except for the assumption that traders do not share a common prior (since  $\tau_H \neq \tau_L$ ), the equilibrium is a perfect Bayesian equilibrium.

We show that with enough disagreement—if  $\tau_H$  is sufficiently larger than  $\tau_L$ —there will be trade based on private information. The degree of disagreement  $\tau_H/\tau_L$  affects the equilibrium prices and quantities traded. Without overconfidence—in a model of rational expectations with a common prior—there would be no trade.

It is important to distinguish between the common prior assumption (which we do not make) and the traditional economists' assumption of rationality as consistently applying Bayes law when maximizing expected utility with respect to some probability distribution (which we do make). Morris (1995) eloquently discusses why “dropping the common prior assumption from otherwise rational behavior” is an important research agenda.

The equilibrium has smooth trading and temporary price impact. Indeed, infinitely fast portfolio updating towards target inventories cannot be an equilibrium, and temporary price impact is intuitively necessary to prevent this possibility. If there were no temporary price impact—and the price were only an increasing function of the level of a trader's inventory as in most models—then a trader would reduce price impact costs by moving continuously but very quickly along his residual demand schedule. This could not be a symmetric equilibrium, however, because the counterparties would require compensation, in the form of temporary price impact costs, to compensate for losses from being “picked off” by the discriminating monopolist. To reduce transaction costs, each trader would try to slow his trading relative to others, and the equilibrium would break. With temporary price impact, infinitely fast trading is infinitely expensive because the price is an unboundedly increasing function of the derivative of a trader's inventory.

The continuous equilibrium of Kyle (1985) is conceptually different. While the informed trader optimally smooths out his trading so that his inventory is a continuous function of time, the noise traders are assumed to trade suboptimally. In response to a shock to desired inventories  $\Delta U$ , the noise traders immediately trade the quantity  $\Delta U$  all at once, incurring price impact cost  $\lambda \Delta U$ . If the noise traders were instead to trade smoothly and move quickly but continuously along their residual demand schedule at rate  $\Delta U/\Delta t$  over some small time interval  $\Delta t$ , then they would incur approximately only one-half the price impact cost,  $\frac{1}{2}\lambda \Delta U$ .<sup>4</sup> Such optimized smooth trading by noise traders would break the equilibrium of Kyle (1985), because the market makers on the other side of this smooth trading would suffer losses, significantly changing the nature of liquidity.

4. In models with “impatient” noise traders—such as Chau and Vayanos (2008), Foster and Viswanathan (1994), Caldentey and Stacchetti (2010), and Holden and Subrahmanyam (1992)—a discrete-time setting is needed to prevent optimizing traders from trading infinitely fast. Back *et al.* (2000) are able to implement the discrete-time model of Foster and Viswanathan (1996) in continuous time, because declining permanent price impact over time deters infinitely aggressive trading immediately after trading begins.

### 3.1. Bayesian updating by traders in the model

Traders use the history of private information and the history of the dividend process  $D(t)$  to forecast unobserved dividend growth rate  $G^*(t)$ . To simplify Kalman filtering formulas, the information content of the dividend  $D(t)$  can be expressed in a form analogous to the notation for private information in equation (21). Define  $dI_0(t) := (\alpha_D D(t)dt + dB(t))/\sigma_D$  and  $dB_0 := dB_D$ . Then the public information  $I_0(t)$  in the divided stream (19) can be written

$$dI_0(t) := \tau_0^{1/2} \frac{G^*(t)}{\sigma_G \Omega^{1/2}} dt + dB_0(t), \quad \text{where} \quad \tau_0 := \frac{\Omega \sigma_G^2}{\sigma_D^2}. \quad (24)$$

The process  $I_0(t)$  is informationally equivalent to the dividend process  $D(t)$ . The quantity  $\tau_0$  measures the precision of the dividend process.

Consider next how traders update their estimates of the unobserved growth rate. In a symmetric equilibrium, each trader infers from prices a sufficient statistic for other traders' private information. Thus, all traders update estimates of the unobserved growth rate  $G^*(t)$  as if fully informed about all information  $I_0(s) \equiv D(s)$ ,  $I_1(s), \dots, I_N(s)$ ,  $s \in (-\infty, t]$ , including the private information of other traders. Let  $G_n(t) := \mathbb{E}_t^n[G^*(t)]$  denote trader  $n$ 's estimate of the growth rate conditional on all information. The superscript  $n$  indicates that conditional distributions of growth rates are calculated by trader  $n$ , who believes that his own information has high precision  $\tau_H$  and other traders' information has low precision  $\tau_L$ . The subscript  $t$  denotes conditioning on the history of all information at date  $t$ . Similarly, let  $\text{Var}_t^n[G^*(t)]$  denote trader  $n$ 's conditional variance at date  $t$ .

Appendix A.2 presents Stratonovich–Kalman–Bucy filtering formulas for calculating estimates of  $G^*(t)$  from information of arbitrary precision  $\tau_0, \tau_1, \dots, \tau_N$ .

Equations (A8) and (A9) imply that, for the beliefs of any trader  $n$ , total precision  $\tau$  and non-time-varying scaled error variance  $\Omega$  are given by

$$\tau := \tau_0 + \tau_H + (N-1)\tau_L, \quad \Omega^{-1} := \left( \frac{\text{Var}_t^n[G^*(t)]}{\sigma_G^2} \right)^{-1} = 2\alpha_G + \tau. \quad (25)$$

Although traders agree to disagree about whose information has high precision, it is common knowledge that they use the same values of  $\tau$  and  $\Omega$ .

From the history of each raw information process  $I_n(s)$ ,  $s \in (-\infty, t]$ , define a “signal”  $H_n(t)$ ,  $n = 0, \dots, N$ , by plugging  $\tau$  and  $\Omega$  into equation (A13). The resulting exponentially weighted average of past innovations, given by

$$H_n(t) := \int_{u=-\infty}^t e^{-(\alpha_G + \tau)(t-u)} dI_n(u), \quad n = 0, 1, \dots, N, \quad (26)$$

is a sufficient statistic for the information in the history of  $I_n(s)$ . Equation (26) implies that more distant information  $dI_n(t)$  receive exponentially lower weight since (1) past signals contain information about the past growth rate which mean-reverts to zero at rate  $\alpha_G$  and (2) new information is generated at a rate  $\tau$ . Let  $H_{-n}(t)$  denote the average of the other traders' signals:

$$H_{-n}(t) := \frac{1}{N-1} \sum_{\substack{m=1 \\ m \neq n}}^N H_m(t). \quad (27)$$

For trader  $n$ 's beliefs  $\tau_H$  and  $\tau_L$ , equation (A15) implies that his estimate of the growth rate  $G_n(t)$  is a linear combination of  $H_0(t)$ ,  $H_n(t)$ , and  $H_{-n}(t)$  given by

$$G_n(t) := \sigma_G \Omega^{1/2} \left( \tau_0^{1/2} H_0(t) + \tau_H^{1/2} H_n(t) + \tau_L^{1/2} (N-1) H_{-n}(t) \right). \quad (28)$$

This equation has a simple intuition. All traders place the same weight  $\tau_0^{1/2}$  on the dividend-information signal  $H_0(t)$ . Each trader assigns a larger weight  $\tau_H^{1/2}$  to his own signal and a lower weight  $\tau_L^{1/2}$  to each of the other  $N-1$  traders' signals. Since equation (28) describes a steady state in which traders agree about the constant value of  $\Omega$ , the weights on the  $H$ -variables do not vary over time or across traders.

As discussed next, trader  $n$ 's optimal trading strategy depends on both the average of other traders' estimates of  $G^*(t)$ , defined as  $G_{-n}(t) := \frac{1}{N-1} \sum_{m=1, m \neq n}^N G_m(t)$ , and his own beliefs about the dynamic statistical relationship between  $G^*(t)$  and the sufficient statistics  $H_0(t)$ ,  $H_n(t)$ , and  $H_{-n}(t)$ .

### 3.2. Linear conjectured strategies

We seek a symmetric equilibrium in which traders use simple Markovian linear strategies. To reduce the number of state variables, it is convenient to replace the three state variables  $H_0(t)$ ,  $H_n(t)$ ,  $H_{-n}(t)$  with two composite state variables  $\hat{H}_n$  and  $\hat{H}_{-n}$  defined using a constant  $\hat{a}$  by

$$\hat{H}_n(t) := H_n(t) + \hat{a} H_0(t), \quad \hat{H}_{-n}(t) := H_{-n}(t) + \hat{a} H_0(t), \quad \hat{a} := \frac{\tau_0^{1/2}}{\tau_H^{1/2} + (N-1)\tau_L^{1/2}}. \quad (29)$$

Trader  $n$  conjectures that four constant “ $\gamma$ -parameters”— $\gamma_D$ ,  $\gamma_H$ ,  $\gamma_S$ , and  $\gamma_P$ —define symmetric linear demand schedules for other traders  $m$ ,  $m \neq n$ , given by

$$x_m(t) = \frac{dS_m(t)}{dt} = \gamma_D D(t) + \gamma_H \hat{H}_m(t) - \gamma_S S_m(t) - \gamma_P P(t). \quad (30)$$

Market clearing implies that trader  $n$ 's flow-demand  $x_n(t) = dS_n(t)/dt$  satisfies

$$x_n(t) + \sum_{\substack{m=1 \\ m \neq n}}^N \left( \gamma_D D(t) + \gamma_H \hat{H}_m(t) - \gamma_S S_m(t) - \gamma_P P(t) \right) = 0. \quad (31)$$

Using zero net supply  $\sum_{m=1}^N S_m(t) = 0$ , this equation can be solved for  $P(t)$  as a function of  $x_n(t)$  to obtain trader  $n$ 's conjectured price impact function

$$P(x_n(t)) = \frac{\gamma_D}{\gamma_P} D(t) + \frac{\gamma_H}{\gamma_P} \hat{H}_{-n}(t) + \frac{\gamma_S}{(N-1)\gamma_P} S_n(t) + \frac{1}{(N-1)\gamma_P} x_n(t). \quad (32)$$

Equation (32) is analogous to equation (4) from the one-period model, with the quantity traded  $x_n(t)$  interpreted as the time derivative of inventories (or trading intensity). The intercept of the residual supply schedule depends on dividends  $D(t)$  and the signals of other traders  $\hat{H}_{-n}(t)$ .

We call the term linear in  $S_n(t)$  “permanent impact” and the term linear in  $x_n(t)$  “temporary impact”. Analogous to the one-period model (equations (17) and (18)), equation (32) defines coefficients of permanent impact  $\lambda$  and temporary impact  $\kappa$ :

$$\lambda := \frac{\gamma_S}{(N-1)\gamma_p}, \quad \kappa := \frac{1}{(N-1)\gamma_p}. \quad (33)$$

We refer to the inverse of temporary price impact  $1/\kappa$  as “market liquidity”.

Imperfect competition requires trader  $n$  to take into account both his permanent and temporary price impact in choosing how fast to change his inventory. Trader  $n$  exercises monopoly power in choosing how fast to demand liquidity from other traders to profit from information. He also exercises monopoly power in choosing how fast to provide liquidity to the other  $N-1$  traders who, according to trader  $n$ ’s beliefs, trade with overconfidence and therefore make supplying liquidity to them profitable. Intuitively, the symmetry of equilibrium trading strategies requires traders to believe they are being adequately compensated for both supplying and demanding liquidity in a manner consistent with market clearing.

### 3.3. Equilibrium with linear trading strategies

Define a steady-state equilibrium with symmetric, linear flow-strategies as a Bayesian perfect equilibrium in which (1) traders maximize expected utility by choosing symmetric flow-strategies of the form (30) with constant  $\gamma$ -parameters as functions of time and (2) inventories have non-stochastic, finite variances which do not vary over time. The Bayesian perfect equilibrium concept requires strategies to be dynamically consistent. In our model, prices, inventories, and expected returns have stationary distributions; in Vayanos (1999) and Du and Zhu (2017) in contrast, these variables are non-stationary.

Appendix A.3 characterizes equilibrium using the “no-regret” approach in the same way as the proof of Theorem 1 for the one-period model. Trader  $n$  solves for his optimal consumption and trading strategy by plugging the price impact function (32) into his dynamic optimization problem. He infers the value of  $H_{-n}(t)$  by observing his residual flow-supply schedule, picks the optimal point on this schedule, and implements it with a linear demand schedule. Linear conjectured strategies for other traders  $m \neq n$  make the optimization problem quadratic in  $x_n(t)$ ; thus, the optimal flow-demand  $x_n^*(t)$  is the solution to a linear equation. This linear solution generates higher profits than any non-linear demand schedule.

The proof in Appendix A.3 conjectures an exponential value function whose exponent is a specific quadratic function of the state variables  $M_n(t)$ ,  $D(t)$ ,  $\hat{H}_n(t)$ ,  $\hat{H}_{-n}(t)$ , and  $S_n(t)$ , defined in terms of nine “ $\psi$ -parameters”; obtains first-order necessary conditions from the Hamilton–Jacobi–Bellman equation; equates coefficients in the conjectured linear solution; and then combines the resulting nine  $\psi$ -equations with four  $\gamma$ -equations, imposing symmetry on the solution. The proof shows that these thirteen equations can be reduced to six polynomial equations (A57)–(A62) in six unknowns. A solution determines the nine  $\psi$ -parameters defining the value function in equation (A37) and the four  $\gamma$ -parameters defining trading strategies in equation (30). The thirteen endogenous parameters are functions of the ten exogenous parameters  $r$ ,  $\rho$ ,  $A$ ,  $\alpha_D$ ,  $\sigma_D$ ,  $\alpha_G$ ,  $\sigma_G$ ,  $N$ ,  $\tau_H$ , and  $\tau_L$  (in terms of which the quasi-exogenous parameters  $\tau_0$ ,  $\tau$ ,  $\Omega$ , and  $\hat{a}$  are also defined).

There always exists a no-trade equilibrium  $X_n \equiv 0$ , with no well-defined price.

**Theorem 3. (Characterization of Equilibrium in the Continuous-Time Model with Overconfidence and Imperfect Competition).** *There exists a steady-state, Bayesian-perfect equilibrium with symmetric, linear flow-strategies and positive trading volume if and only if*

the six polynomial equations (A57)–(A62) have a solution satisfying the second-order condition  $\gamma_P > 0$  and the stationarity condition  $\gamma_S > 0$ . Such an equilibrium has the following properties:

- (1) There is an endogenously determined constant  $C_L > 0$ , defined in equation (A49), such that trader  $n$ 's optimal flow-strategy  $x_n^*(t)$  makes time-differentiable inventories  $S_n(t)$  change at rate

$$x_n^*(t) = \frac{dS_n^*(t)}{dt} = \gamma_S \left( C_L \left( \hat{H}_n(t) - \hat{H}_{-n}(t) \right) - S_n^*(t) \right). \quad (34)$$

- (2) There is an endogenously determined constant  $C_G > 0$ , defined in equation (A49), such that the equilibrium price is

$$P^*(t) = \frac{D(t)}{r + \alpha_D} + C_G \frac{\bar{G}(t)}{(r + \alpha_D)(r + \alpha_G)}, \quad (35)$$

where  $\bar{G}(t) := \frac{1}{N} \sum_{n=1}^N G_n(t)$  denotes the average expected growth rate.

Equations (34) and (35) are similar to equations (11) and (13) in the one-period model. Use equation (34) to define trader  $n$ 's “target inventory”  $S_n^{TI}(t)$  as the inventory level such that trader  $n$  does not trade ( $x_n(t) = 0$ ):

$$S_n^{TI}(t) = C_L \left( \hat{H}_n(t) - \hat{H}_{-n}(t) \right). \quad (36)$$

Trader  $n$  targets a long position if his own signal  $\hat{H}_n(t)$  is greater than the average signal of other traders  $\hat{H}_{-n}(t)$  and a short position if it is less. The proportionality constant  $C_L$  measures the sensitivity of target inventories to the difference. Trader  $n$ 's optimal flow-strategy  $x_n^*(t)$  can be written

$$x_n^*(t) = \frac{dS_n(t)}{dt} = \gamma_S \left( S_n^{TI}(t) - S_n(t) \right). \quad (37)$$

Equation (37) defines a partial adjustment strategy similar to the one in the partial equilibrium models of Garleanu and Pedersen (2013, 2016). The parameter  $\gamma_S$  measures the “speed of trade” as the rate at which inventories adjust towards target levels. Trading volume is finite. Section 4.1 provides a more detailed analysis.

The price in equation (35) immediately reveals the average of all signals, responding instantaneously to innovations in each trader's private information. This occurs even though each trader intentionally slows down trading to reduce price impact resulting from adverse selection. If  $C_G$  were equal to one, the price in equation (35) would equal the average of traders' risk-neutral buy-and-hold valuations, consistent with Gordon's growth formula and the one-period model. As discussed in Section 4.3, a Keynesian beauty contest makes the multiplier  $C_G$  less than one.

It is an analytical result that risk aversion affects quantities, not prices:

**Theorem 4. (Comparative Statics for Risk Aversion).** *If risk aversion  $A$  is scaled by a factor of  $F$  to  $A/F$ , then  $C_L$  changes to  $C_L F$ ,  $\lambda$  changes to  $\lambda/F$ ,  $\kappa$  changes to  $\kappa/F$ ,  $S_n^{TI}(t)$  changes to  $S_n^{TI}(t)F$ , but  $\gamma_S$  and  $C_G$  remain the same.*

When risk tolerance  $1/A$  scales up by factor  $F > 1$ , Theorem 4 says that traders scale up target inventories proportionally in response to proportional reductions in temporary and permanent price impact  $\lambda$  and  $\kappa$ . The speed of trade  $\gamma_S$  remains the same. With infinite risk aversion, each trader's target inventory  $S_n^{TI}(t)$  drops to zero.



### 3.4. An existence condition

Obtaining an analytical solution for the equilibrium in Theorem 3 requires solving the six polynomial equations (A57)–(A62). While these equations have no obvious analytical solution, they can be solved numerically. Extensive numerical calculations lead us to conjecture that the existence condition for the continuous-time model is exactly the same as the existence condition for the one-period model:

**Conjecture 1. (Existence Condition).** *A steady-state, Bayesian-perfect equilibrium with symmetric, linear flow-strategies exists if and only if*

$$\Delta_H := \frac{\tau_H^{1/2}}{\tau_L^{1/2}} - 2 - \frac{2}{N-2} > 0. \quad (38)$$

We have examined numerical solutions to the six equations (A57)–(A62) for many exogenous parameter values. When existence condition (38) is satisfied, the numerical algorithm always finds precisely one solution satisfying the second-order condition  $\gamma_P > 0$ , and this solution also satisfies the stationarity condition  $\gamma_S > 0$ . When existence condition (38) is reversed, the numerical algorithm sometimes finds solutions satisfying the second-order condition  $\gamma_P > 0$ , but these solutions do not satisfy the stationarity condition  $\gamma_S > 0$ . The closed-form solution derived in Section 4.4 for vanishing liquidity ( $\Delta_H \rightarrow 0$ ) is consistent with conjecture (38).

Similar to the one-period model, we expect equilibrium with trade to exist only if there is enough disagreement. Here is some intuition. Suppose the market price of an asset would be \$90 if trader  $n$  does not trade. Suppose further that trader  $n$  values additional units of the asset at \$100. To optimally exploit his market power, trader  $n$  has an incentive to buy at a rate such that short-term price impact moves the price about half-way between these two values, to \$95. To be willing to take the other side of such smooth trades of their competitors, traders must believe that their competitors' signals are only about half as precise as their competitors believe them to be. This intuition is consistent with the existence condition  $\Delta_H > 0$ , which is equivalent to  $\tau_H^{1/2}/2 > \tau_L^{1/2}(1 + 1/(N-2))$ . In this context, “half as precise” means  $\tau_H^{1/2}/2 \approx \tau_L^{1/2}$ ; the term  $1/(N-2)$  is due to market power.

### 3.5. A competitive model as benchmark

To understand how imperfect competition affects the equilibrium, Appendix B.2 characterizes the equilibrium of an otherwise equivalent model in which the assumption of perfect competition replaces imperfect competition.

Conceptually, the model with perfect competition differs from the model with imperfect competition in two ways. First, when traders construct their strategies  $(c_n(t), S_n(t))$ , they do not take into account the effect of their trades on prices, and this simplifies their wealth dynamics (B8). Second, since it is not necessary for a trader to consider separately money holdings  $M_n$  and a stock holdings  $S_n$  in the case of perfect competition, the value function conjectured in (B12) is a quadratic exponential function of only three state variables, wealth  $W_n$  and two information variables  $\hat{H}_n$  and  $\hat{H}_{-n}$ ; this reduces the number of parameters in the value function. The results are summarized in the following theorem.

**Theorem 5. (Characterization of Equilibrium in the Continuous-Time Model with Overconfidence and Perfect Competition).** *Assume  $\tau_H > \tau_L$ . There exists a steady-state, Bayesian-perfect equilibrium with symmetric, linear strategies with positive trading volume if and only if the*



three polynomial equations (B26)–(B28) have a solution. Such an equilibrium has the following properties:

- (1) There is an endogenously determined constant  $C_L > 0$ , defined in equation (B19), such that trader  $n$ 's optimal inventories  $S_n^*(t)$  are

$$S_n^*(t) = C_L \left( \hat{H}_n(t) - \hat{H}_{-n}(t) \right). \quad (39)$$

- (2) There is an endogenously determined constant  $C_G > 0$ , defined in equation (B17), such that the equilibrium price is

$$P^*(t) = \frac{D(t)}{r + \alpha_D} + C_G \frac{\bar{G}(t)}{(r + \alpha_D)(r + \alpha_G)}, \quad (40)$$

where  $\bar{G}(t) := \frac{1}{N} \sum_{n=1}^N G_n(t)$  denotes the average expected growth rate.

The existence condition for the equilibrium with imperfect competition ( $\Delta_H > 0$ ) is more restrictive than the existence condition for the competitive equilibrium ( $\tau_H > \tau_L$ ), even when  $N$  goes to infinity, because perfectly competitive traders do not exercise monopoly power over their private signals and trade more aggressively.

Equations (39) and (40) are similar to corresponding equations (34) and (35) in Theorem 3, but the values of  $C_L$  and  $C_G$  are different. As discussed in Section 4.3, competition enhances price dampening, making the value of  $C_G$  smaller than with imperfect competition. Most importantly, price-taking competitors do not smooth their trading. Each trader immediately adjusts actual inventories to target levels (as if  $\gamma_S \rightarrow \infty$ ); since target inventories are diffusions, trading volume is infinite.

### 3.6. An analogous model with private values

Instead of motivating trade using a model based on agreement to disagree with no common prior, trade can be motivated by private values with a common prior. Consider an alternative model of imperfect competition identical to our disagreement model except for two important differences: (1) Instead of agreeing to disagree, all private signals have the same precision  $\tau_I$ . (2) In addition to the common cash dividend  $D(t)$ , each trader receives an orthogonal private value or convenience yield  $\pi_J H_n^J(t)$  which follows an AR-1 process. This structure is common knowledge; traders share a common prior. To keep the number of state variables the same, assume that the exogenous mean-reversion rate of the convenience yield is the same as the endogenous mean-reversion rate of private signals. Supplementary Appendix C examines such a model in detail.

All of the equations describing the disagreement model (Section 3 and Appendix A) map nicely into corresponding equations describing the private-values model (Supplementary Appendix C). Noise from private values lowers the precision of the estimate of other traders' signals inferred from prices. To make the models as similar as possible, the parameter  $\tau_I$  can be chosen to equal the parameter  $\tau_H$ , and the level of innovation variance in shocks to private values can be chosen so that the endogenous lower precision inferred from prices is equal to  $\tau_L$ .

A comparison of the two models highlights a subtle dynamic structure of beliefs in the model with disagreement. In the model with private values, traders agree that they have different valuations in the present, and they furthermore agree that these different valuations will mean revert towards the same unconditional common mean consistent with a common prior. In the

model with disagreement, traders also agree that they have different valuations in the present, but—in contrast to the model with private values—they furthermore agree to disagree about the stochastic process their different current valuations will follow in the future. Specifically, each trader believes that the other traders' valuations will converge on average to his own valuation in the long run but deviate in the short run; because they have different beliefs about valuation dynamics as a result of not sharing a common prior, they disagree in the present about how their expectations will differ in the future. Algebraically, this effect shows up in equations (C33) and (C34) in Supplementary Appendix C; the discussion following these equations clarifies the intuition further.

As discussed in detail in Section 4.3, this disagreement about valuation dynamics leads to a Keynesian beauty contest with dampening of prices ( $0 < C_G < 1$ ). With private values, it can be shown analytically that no such dampening occurs ( $C_G = 1$ ). The private-values model is simpler than the model with disagreement because traders disagree about the present only; they do not disagree about the future. Similarly, both Vayanos (1999) and Du and Zhu (2017) obtain no price dampening in models where inventories or private values follow random walks.

To generate a Keynesian beauty contest from an isomorphic model of private values with a common prior, it would be necessary to make complicated, unnatural assumptions about exogenous cross-sectional and time series correlations of private values and private information to mimic artificially the natural dynamics of Bayes law with agreement to disagree. Ockham's razor supports modelling a Keynesian beauty contest using disagreement, not a common prior.

#### 4. IMPLICATIONS OF THE CONTINUOUS-TIME MODEL

This section presents implications of the continuous-time model for (1) trading strategies, (2) market liquidity, and (3) prices. For notational simplicity, the superscript “\*” on equilibrium prices and strategies is suppressed.

##### 4.1. Trading strategies: a partial adjustment process

Traders trade smoothly. Trader  $n$  follows a partial adjustment strategy (equation (37)), and his inventory  $S_n(t)$  gradually converges towards its target level  $S_n^{TI}(t)$  at rate  $\gamma_S$  (equation (36)). Sample paths for target inventories  $S_n^{TI}(t)$  and trading intensity  $x_n(t)$  are diffusions (of order  $dt^{1/2}$ ). Sample paths for actual inventories  $S_n(t)$  are not diffusions but rather differentiable functions of time (of order  $dt$ ).

The integral representation of the inventory dynamics in equations (36) and (37),

$$S_n(t+s) = e^{-\gamma_S s} \left( S_n(t) + \int_{u=t}^{t+s} e^{-\gamma_S(t-u)} \gamma_S C_L (\hat{H}_n(u) - \hat{H}_{-n}(u)) du \right), \quad (41)$$

shows that traders add to existing inventories based on current differences in signals  $\hat{H}_n(t) - \hat{H}_{-n}(t)$  and liquidate their existing inventories, accumulated based on past signal differences, at rate  $\gamma_S$ . Even if signals  $\hat{H}_n(t)$  and  $\hat{H}_{-n}(t)$  were to remain constant over some period of time and the price did not change, trader  $n$  would continue to trade based on the level of his past disagreement with the market.

Although prices adjust instantaneously, quantities adjust slowly. As soon as trader  $n$  adjusts his trading intensity  $x_n(t)$  after getting new information, the price instantaneously moves to a new equilibrium level, even though he has not yet traded a single share.

The smooth trading model captures in a realistic manner the inventory behaviour of asset managers who use public and private information to forecast stock returns. When information

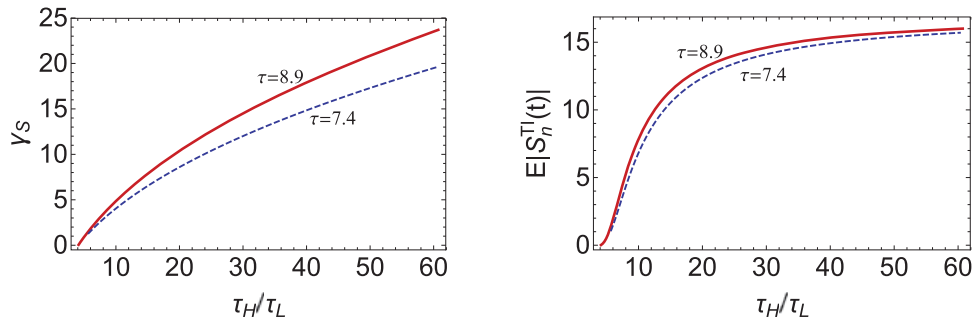


FIGURE 1

Values of  $\gamma_S$  and  $\mathbb{E}[|S_n^{TI}(t)|]$  as functions of  $\tau_H/\tau_L$  for  $\tau=7.4$  and  $\tau=8.9$ .

changes, an asset manager updates signals, obtains a new estimate of the asset's value, recalculates his target inventory, and adjusts trading to move inventories in the direction of the new target. Since moving large blocks over short periods of time is expensive, an asset manager builds positions gradually, trading off price impact against the speed with which signals decay.

Trader  $n$  believes that the information variables  $\hat{H}_n(t)$  and  $\hat{H}_{-n}(t)$  in equation (36) follow a bivariate vector auto-regression process. Traders disagree about drift rates. Trader  $n$  believes that  $\hat{H}_n(t) - \hat{H}_{-n}(t)$  mean-reverts at rate  $\alpha_G + \tau$  but also drifts in a direction proportional to  $(\tau_H^{1/2} - \tau_L^{1/2})G_n(t)$  (see equation (A36)).

Intuition suggests that more disagreement will make markets more liquid, and this additional liquidity will be associated with more rapid adjustment of actual inventories towards target levels. Our numerical results support this intuition.

Figure 1 shows how the speed of trading and level of inventories change when disagreement  $\tau_H/\tau_L$  changes.<sup>5</sup> Intuitively, as disagreement increases, it becomes less costly for a trader to trade towards the target inventory more rapidly because other trades are more willing to provide liquidity. Therefore, the expected size of target inventories  $\mathbb{E}[|S_n^{TI}(t)|]$  increases (right panel) and the speed of inventory adjustment  $\gamma_S$  also increases (left panel). The speed of inventory convergence to target levels also increases when the decay rate of signals  $\alpha_G + \tau$  increases. Intuitively, when a signal decays faster, a trader trades faster.<sup>6</sup>

Figure 2 shows how the speed of trading and level of inventories change when the number of traders  $N$  changes. The speed of inventory adjustment  $\gamma_S$  increases steadily with  $N$  (left panel) since more competition makes trading less costly. Expected target inventories  $\mathbb{E}[|S_n^{TI}(t)|]$  increase towards a constant level when  $N$  is large (right panel) since risk aversion limits the maximum size of inventories when more competition makes trading costs fall.<sup>7</sup>

Figure 3 presents three simulated paths for target inventories (dashed lines) and actual inventories (solid lines).<sup>8</sup> When disagreement  $\tau_H/\tau_L$  is larger—and the market is more liquid as discussed in Section 4.2—actual inventories closely track target inventories, as in Figure 3a.

5. Throughout this article, to conduct comparative statics analysis for changing the degree of disagreement, we change  $\tau_H/\tau_L$  by changing  $\tau_H$  and  $\tau_L$  in opposite directions so that the value of total precision  $\tau$  remains constant. When we change the signal decay rate  $\alpha_G + \tau$ , we change the total precision  $\tau$  by increasing  $\tau_H$  and  $\tau_L$  proportionally while holding  $\alpha_G$  constant. When we change the number of traders, we set  $\tau_L = 0$  and therefore fix total precision at  $\tau_H$ .

6. Numerical calculations in Figure 1, Figure 4, and Figure 8a are based on exogenous parameter values  $\tau=7.4$  (or  $\tau=8.9$ ),  $r=0.01$ ,  $A=1$ ,  $\alpha_D=0.1$ ,  $\alpha_G=0.02$ ,  $\sigma_D=0.5$ ,  $\sigma_G=0.1$ ,  $\tau_0=\Omega\sigma_G^2/\sigma_D^2=0.0054$ , and  $N=100$ .

7. Numerical calculations in Figure 2, Figure 5, and panel (b) of Figure 8 are based on the exogenous parameter values  $\tau_L=0$ ,  $\tau=1.4$ ,  $r=0.01$ ,  $A=1$ ,  $\alpha_D=0.1$ ,  $\alpha_G=0.02$ ,  $\sigma_D=0.5$ ,  $\sigma_G=0.1$ , and  $\tau_0=\Omega\sigma_G^2/\sigma_D^2=0.0279$ .

8. The paths are generated using equations (36), (37), (41), (A34), and (A35), which describe the dynamics of  $H_n(t)$ ,  $H_{-n}(t)$ ,  $S_n(t)$ , and  $S_n^{TI}(t)$ . Numerical calculations in Figure 3 are based on the exogenous parameter values  $\alpha_D=0.1$ ,

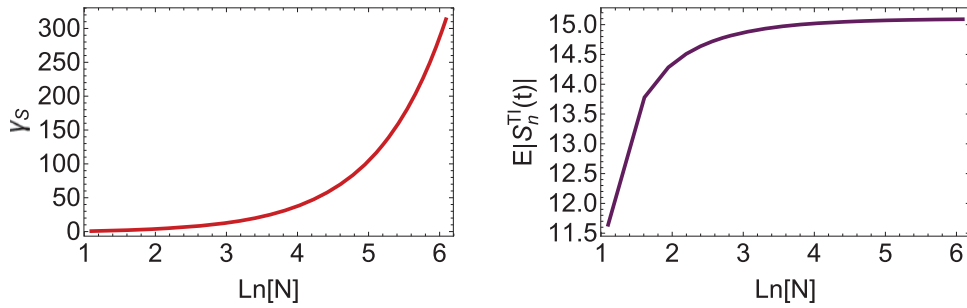


FIGURE 2

Values of  $\gamma_S$  and  $\mathbb{E}[|S_n^{TI}(t)|]$  as functions of  $\ln(N)$  for  $\tau = 1.4$  and  $\tau_L = 0$ .

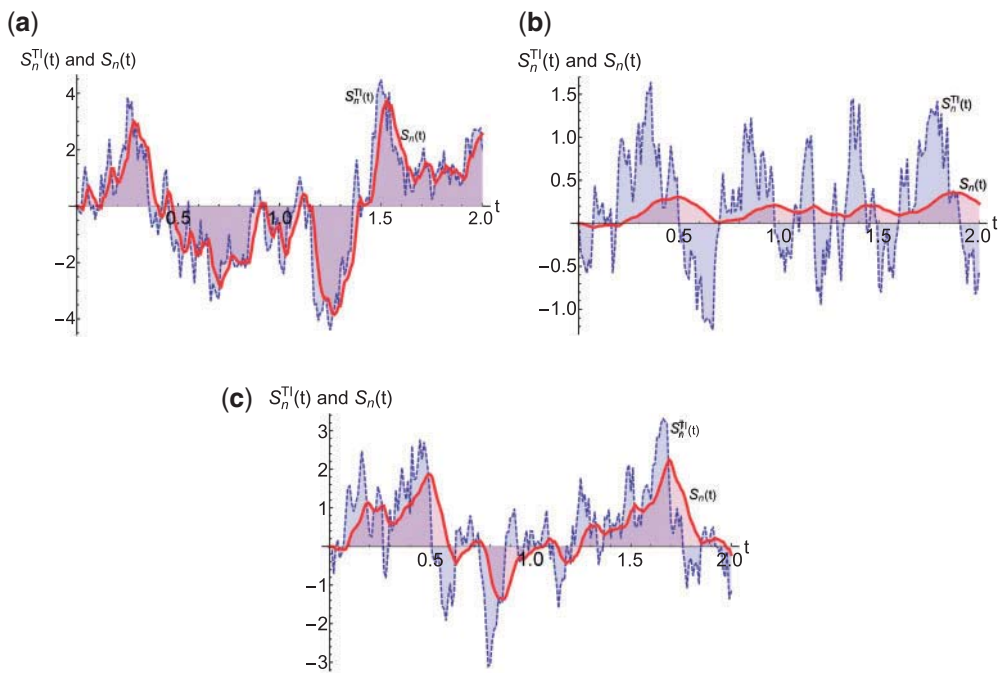


FIGURE 3

Simulated paths of  $S_n^{TI}(t)$  (Dashed) and  $S_n(t)$  (Solid). (a) with high disagreement; (b) with low disagreement; (c) with low decay rate and high disagreement.

When disagreement  $\tau_H/\tau_L$  is smaller—and the market is less liquid—actual inventories deviate significantly from target inventories since traders restrict their speed of trading, as in Figure 3b. To illustrate that the speed at which traders' inventories converge to target levels also depends on the decay rate of their signals, Figure 3c plots actual and target inventories using the same level of disagreement as in Figure 3a but a lower decay rate  $\alpha_G + \tau$  (by varying  $\tau$ ); actual inventories

$\alpha_G = 0.02$ ,  $\sigma_D = 0.5$ ,  $\sigma_G = 0.1$ , and  $N = 100$ , with  $\tau = 8.9$  and  $\tau_0 = \Omega\sigma_G^2/\sigma_D^2 = 0.0045$  in both Figure 3a and b;  $\tau_H = 4.46$  and  $\tau_L = 0.045$  in Figure 3a;  $\tau_H = 0.5$  and  $\tau_L = 0.085$  in Figure 3b; and  $\tau = 3.15$ ,  $\tau_H = 1.56$ ,  $\tau_L = 0.016$ , and  $\tau_0 = 0.0126$  in Figure 3c.

track target inventories less closely than in panel (a), in line with Figure 1. Note that target and actual inventories would coincide in the competitive model.

Our model explains how asset managers try to outperform benchmarks by trading securities they perceive to be undervalued or overvalued. Stationary, mean-reverting target inventories and perceived expected returns are endogenous consequences of the simultaneous solutions to optimization problems based on public and private information flow, total precision of information in the market, disagreement among traders, and traders' risk-bearing capacity. If actively traded stocks have faster information flow (larger  $\alpha_G + \tau$ ), then our model predicts more rapidly mean-reverting target inventories in more active markets. Our model also predicts that the speed of inventory adjustment  $\gamma_S$  tends to increase with faster signal decay (increasing  $\alpha_G + \tau$ ) or more disagreement (increasing  $\tau_H/\tau_L$ ); markets with high trading volume are therefore more liquid.

Empirical evidence on long- and short-term trading is consistent with partial adjustment towards fluctuating target inventories. One of our main contributions is the empirical hypothesis that long-term trading results from slow information flow and high trading costs in low-volume markets while short-term trading results from fast information flow and low trading costs in high-volume markets. Using granular proprietary databases, Puckett and Yan (2011) and Chakrabarty *et al.* (2013) find that institutional investors indeed engage in intensive short-term trading, while institutional holdings reported to the SEC on Form 13F suggest long-term strategies with complicated patterns.

#### 4.2. Temporary and permanent price impact

The concepts of temporary and permanent price impact are important for the practical management of transaction costs. Our model links endogenous temporary and permanent impact to deep structural parameters such as the precision of information flow and the magnitude of disagreement.

We use the terms “temporary” and “permanent” price impact differently from the empirical market microstructure literature, which usually describes time series properties of market prices. Temporary price impact is often associated with negative first-order autocorrelation in price changes (bid-ask bounce), and permanent price impact is associated with persistent correlations between price changes and order flow. Instead, like sophisticated traders in the asset management industry, we think of temporary and permanent price impacts as components of transaction costs over which traders explicitly optimize when constructing trading strategies. Traders correctly understand that faster execution leads to larger temporary price impact but has no additional effect on permanent price impact.

Combining equations (32) and (33), we can write the price as

$$P(S_n(t), x_n(t)) := p_{0,n}(t) + \lambda S_n(t) + \kappa x_n(t). \quad (42)$$

The intercept  $p_{0,n}(t) = \frac{\gamma_D}{\gamma_P} D(t) + \frac{\gamma_H}{\gamma_P} \hat{H}_{-n}(t)$  defines what the price would be, as a weighted average of other traders' signals, if trader  $n$  had no inventories and did not trade. Permanent impact is linear in the trader's own inventory level  $S_n(t)$ . Temporary impact is linear in the time derivative of his inventory  $x_n(t)$  (trading intensity). The permanent price impact parameter  $\lambda$  and the temporary price impact parameter  $\kappa$  are defined in equation (33). A trader correctly believes that the price changes when either the level of his inventory or the intensity of his trading changes. If trader  $n$  suddenly stops trading, then the price immediately reverses by  $\kappa x_n(t)$  as his temporary price impact disappears with his permanent price impact  $\lambda S_n(t)$  remaining intact.

Figures 4 and 5 show that both permanent depth  $1/\lambda$  and temporary depth  $1/\kappa$  increase as disagreement  $\tau_H/\tau_L$  increases, as total precision increases, or as the number of traders  $N$  increases. These numerical results are consistent with the intuition that more disagreement, a faster signal

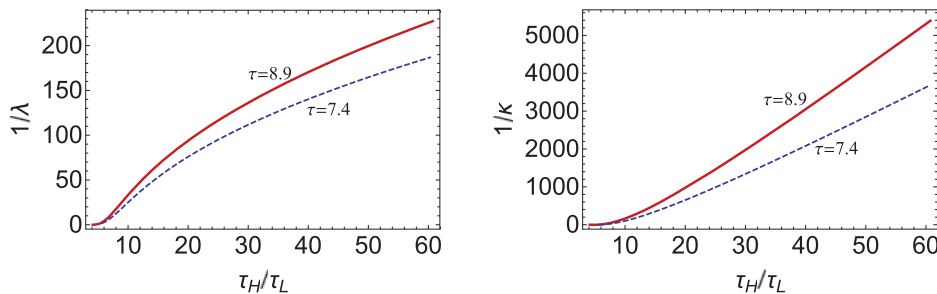


FIGURE 4

Values of  $1/\lambda$  and  $1/\kappa$  as functions of  $\tau_H/\tau_L$  for  $\tau = 7.4$  and  $\tau = 8.9$ .

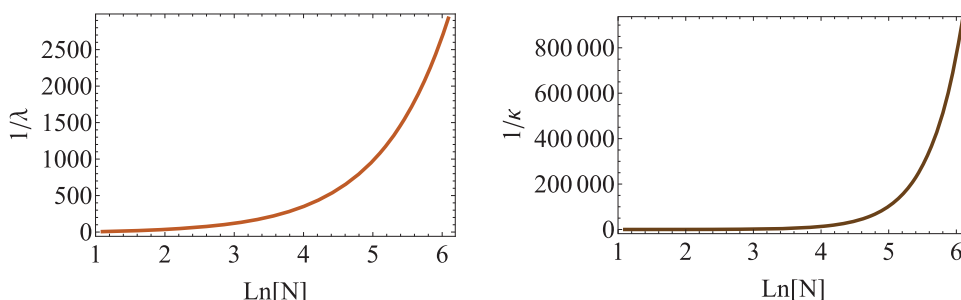


FIGURE 5

Values of  $1/\lambda$  and  $1/\kappa$  as functions of  $\ln(N)$  for  $\tau = 1.4$  and  $\tau_L = 0$ .

decay rate, or a greater number of traders decreases transaction costs by making traders more willing to provide liquidity to one another.

Our price impact model differs sharply from models with linear permanent price impact, but no temporary price impact. As an illustration, consider the continuous-time model of Kyle (1985), in which the informed trader correctly conjectures that the price is given by  $P(t) = P(0) + \lambda(\sigma_U B_U(t) + S_n(t))$ , where  $\sigma_U B_U(t)$  is the inventory of noise traders (Brownian motion) and  $S_n(t)$  is the inventory of the informed trader. This formula is similar to equation (42), except there is no temporary price impact term  $\kappa$ . The informed trader optimizes the permanent impact of his trades; there is no temporary impact as long as his inventory is a continuous function of time. For example, to buy  $Q$  shares over a fixed period of time  $T$ , he walks up the demand schedule, gradually pushing the price up by  $\lambda Q$ . He expects to incur a price impact cost  $\frac{1}{2}\lambda Q$  per share regardless of the speed of his trading.

In our model, in contrast, temporary price impact makes trading costs depend on the speed of trading. Suppose a trader buys  $Q$  shares at a constant rate over time interval  $T$ . For simplicity, assume  $P(t) = \hat{H}_n(t) = \hat{H}_{-n}(t) = 0$ . The average execution price is  $(\frac{1}{2}\lambda + \kappa/T)Q$ , obtained by integrating over equation (42). The first term  $\frac{1}{2}\lambda Q$  is the permanent price impact cost and the second term  $(\kappa/T)Q$  is the additional temporary price impact cost proportional to the execution speed  $Q/T$ . When the trader initiates order execution, the price immediately jumps from zero to  $\kappa Q/T$ , then gradually rises to  $(\lambda + \kappa/T)Q$ , and finally drops back to  $\lambda Q$  at time  $T$  when he stops buying. By varying the speed of his trading over time, the trader affects the temporary-impact component of transaction costs.



There are few other equilibrium models in which temporary price impact results from all traders optimizing price impact over time. Vayanos (1999) and Du and Zhu (2017) derive endogenous transaction costs from models in which trade is motivated by inventory shocks or shocks to private values, respectively. In these two models, permanent and temporary components are difficult to isolate because these papers are set in discrete time.<sup>9</sup>

Asset management practitioners recognize the importance of managing both permanent and temporary price impact costs. The practitioner-oriented model of Almgren and Chriss (2000) is essentially a non-linear generalization of our equation (42); one difference is that our intercept  $p_{0,n}(t)$  changes over time in a manner trader  $n$  believes he can predict, whereas the intercept in the Almgren–Chriss model is an exogenous random walk. Using an alternative model of price resilience in which temporary price impact is exogenously assumed to decay gradually at an exponential rate, Obizhaeva and Wang (2013) derive an optimal way to manage temporary price impact costs.

Since our model is symmetric across traders, market clearing implies that quantities traded are uncorrelated with the price process. The equilibrium price process resembles a Brownian motion. It is therefore impossible to infer price impact from correlations between prices and equilibrium quantities (at any lag). As Black (1982) points out, we can hope to learn about price impact by studying mistakes that traders make or from performing experiments.

To illustrate how suboptimal execution relates to permanent and temporary price impact, consider the following off-equilibrium scenario. Suppose trader  $n$  silently decides to deviate from his equilibrium strategy by trading towards his target inventory at some rate  $\bar{\gamma}_S$ , which is arbitrarily faster or slower than the equilibrium rate  $\gamma_S$ . To fix ideas, suppose he thinks about implementing the strategy

$$\bar{x}_n(t) = \bar{\gamma}_S \left( S_n^{TI}(t) - \bar{S}_n(t) \right) \quad (43)$$

at each point  $t > 0$ . When  $\bar{\gamma}_S = \gamma_S$ , this equation coincides with the equilibrium strategy in equation (37); when  $\bar{\gamma}_S > \gamma_S$ , the trader moves to his target inventory  $S_n^{TI}(t)$  more aggressively; and when  $\bar{\gamma}_S < \gamma_S$ , the trader is more patient. After date  $t = 0$ , the off-equilibrium inventory level  $\bar{S}_n(t)$  is given by

$$\bar{S}_n(t) = e^{-\bar{\gamma}_S t} \left( S_n(0) + \int_{u=0}^t e^{\bar{\gamma}_S u} \bar{\gamma}_S C_L (\hat{H}_n(u) - \hat{H}_{-n}(u)) du \right). \quad (44)$$

For simplicity, suppose  $D(0) = 0$  and trader  $n$  holds a positive target inventory at  $t = 0$ , with other traders' signals at their long-term mean  $\hat{H}_{-n}(0) = 0$ , implying

$$S_n(0) = S_n^{TI}(0) = C_L \hat{H}_n(0) > 0. \quad (45)$$

Then, from equation (32) and (33), we get

$$P(0) = \frac{\gamma_S}{(N-1)\gamma_P} S_n(0) > 0. \quad (46)$$

9. Distinguishing temporary from permanent price impact is more complicated in discrete-time models. If time intervals between rounds of trading were to become infinitely short, then price impact would equal the product of an infinitely large price impact coefficient and an infinitely small quantity traded. Since the market clears in both the time derivative of inventories and the level of inventories, our continuous-time model gracefully deals with this “infinity-times-zero” problem and crystalizes how the speed of trading affects the equilibrium by making a clean distinction between temporary and permanent price impact. One component, which we call permanent price impact, is linear in the level of inventories (stocks). The other component, which we call temporary price impact, is linear in the time derivative of inventories (flows).



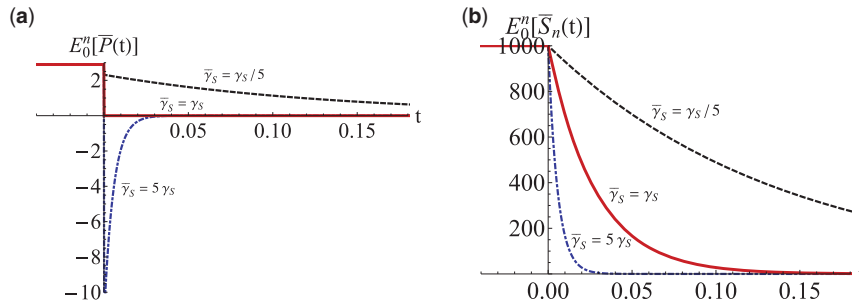


FIGURE 6

The dynamics of expected prices and inventories. (a) The dynamics of expected prices; (b) The dynamics of inventories.

Next, assume that at time  $t=0^+$ , trader  $n$ 's sufficient statistic  $\hat{H}_n(0)$  suddenly drops to zero, reducing both his target inventory and the price to zero. Since  $\hat{H}_n(0^+) = \hat{H}_{-n}(0^+) = 0$ , the price is  $\mathbb{E}_0^n[P(t)] = 0$ . Then, equations (44) and (42) imply that trader  $n$ 's expectation of future inventories and prices—the impulse-response functions from the perspective of trader  $n$ —are given by

$$\mathbb{E}_0^n[\bar{S}_n(t)] = e^{-\bar{\gamma}_S t} S_n(0), \quad (47)$$

$$\mathbb{E}_0^n[\bar{P}(t)] = -\frac{\bar{\gamma}_S - \gamma_S}{(N-1)\gamma_P} e^{-\bar{\gamma}_S t} S_n(0). \quad (48)$$

In Figure 6a shows expected paths of future prices based on equation (48), and Figure 6b shows paths of future inventories based on equation (47). As shown by the solid red lines, if trader  $n$  liquidates his inventory at an equilibrium rate  $\bar{\gamma}_S = \gamma_S$ , then the price immediately drops to zero, but the trader continues to trade out of his inventories over time. Since his equilibrium trading is “expected”, it has no additional effect on prices after time  $0^+$ ; the initial temporary price impact gradually turns into permanent impact at a pace that keeps price changes relatively unpredictable.

Figure 6 also illustrates two off-equilibrium cases.<sup>10</sup> When trader  $n$  sells at a rate five times slower than the equilibrium rate,  $\bar{\gamma}_S = \gamma_S/5$ , the immediate price drop is only 1/5 as large as in equilibrium. The slow rate is not optimal because the higher profits on the early trades at initially better prices are more than offset by lower profits on later trades, when information is being incorporated into prices through the trading of others.

When trader  $n$  sells at a rate five times faster than the equilibrium rate,  $\bar{\gamma}_S = 5\gamma_S$ , the price is expected to drop sharply initially, by five times as much as in equilibrium. Speeding up execution exacerbates temporary price impact initially and elevates transaction costs overall. As the price comes back, the price path exhibits a distinct V-shaped pattern.

*The Flash Crash.* The price response from trading too fast matches the price patterns observed during the flash crash of 6 May 2010. On that day, the E-mini S&P 500 futures price plunged by 5% over a 13-minute period, triggered a 5-second trading halt, and then rose by 6% over the next

10. We assume  $S_n(0) = 1,000$  shares,  $\tau = 9.95$  with  $\tau_0 = 0.004$ ,  $\tau_L = 0.05$ , and  $\tau_H = 5.00$ , implying equilibrium price  $P(0) = 2.896$  and equilibrium  $\gamma_S = 35.8$ . The other exogenous parameter assumptions are  $r = 0.01$ ,  $A = 1$ ,  $\alpha_D = 0.1$ ,  $\alpha_G = 0.02$ ,  $\sigma_D = 0.5$ ,  $\sigma_G = 0.1$ ,  $N = 100$ , and  $D(0^+) = 0$ .

23 minutes. The Staffs of the CFTC and SEC (2010a,b) report that the flash crash was triggered by an automated execution algorithm that sold S&P 500 E-mini futures worth approximately \$4 billion. Kyle and Obizhaeva (2016) note that market microstructure invariance would imply a price impact of less than 1% and attribute the difference between predicted and realized price dynamics to unusually fast execution of the sales. Indeed, the entire \$4 billion quantity was executed over a 36-minute period, even though orders of similar magnitude would normally be executed over several hours. Our model does not explain why a trader chose to trade large quantities so quickly, but it does predict how market prices would respond to a gigantic order, executed much faster than the market expects orders of such size to be executed.

Our explanation for flash crashes is different from explanations based on models with permanent linear price impact but no temporary price impact. For example, sharp price changes may occur in the continuous-time model of Kyle (1985) in response to large trades by noise traders. The size of price declines depends only on quantities sold, not on the speed of selling. Unlike the rapid price recovery after the flash crash, such price declines are corrected only slowly as the informed trader pushes the price back to fundamental value.

Dugast and Foucault (2014) suggest that V-shaped price patterns may occur in response to false signals. The flash crash pattern in Figure 6 occurs because the market falsely infers from unexpectedly rapid selling that other traders receive negative signals. Duffie (2010) suggests that flash crashes may occur because capital moves slowly. Our model shows why traders endogenously choose to move their capital slowly due to adverse selection. Menkveld and Yueshen (2016) conclude that the flash crash could not have been caused by one large sell order because most of the selling took place after the market had crashed and while prices were recovering. This pattern is reasonably consistent with our model, which predicts that the selling itself occurs *after* prices crash and while the market recovers. It is not necessary for this selling to trigger additional selling by others. Of course, flash crashes do not happen in equilibrium in our model. A rare event, perhaps unintended, the flash crash was like an experiment from which something about price impact can be learned.

#### 4.3. *Prices: a Keynesian beauty contest*

Define trader  $n$ 's estimate of the fundamental value of the risky asset  $F_n(t)$  as the expected present value of all future dividends based on all information, discounted at the risk-free rate  $r$  and calculated using the beliefs of trader  $n$ . Gordon's growth formula implies that  $F_n(t)$  is a function of trader  $n$ 's expected growth rate  $G_n(t)$ :

$$F_n(t) := \frac{D(t)}{r + \alpha_D} + \frac{G_n(t)}{(r + \alpha_D)(r + \alpha_G)}. \quad (49)$$

Since the risky asset is in zero net supply, intuition suggests that the price is the average estimate  $\frac{1}{N} \sum_n F_n(t)$  obtained by replacing  $G_n(t)$  with the average  $\bar{G}(t)$  in equation (49). This intuition is consistent with the one-period model. Surprisingly, in the continuous-time model, this intuition turns out to be wrong!

A comparison of equations (35) and (49) reveals that the price equals the average estimate of fundamental value  $\frac{1}{N} \sum_n F_n(t)$  if and only if  $C_G = 1$  in equation (35). Since we always find  $C_G < 1$  in numerical calculations, we conjecture that  $0 < C_G < 1$  in any equilibrium with trade. Even if all  $N$  traders unanimously agree on the same expected growth rate  $G_n(t) = \bar{G}(t)$ , the equilibrium price reflects a dampened implied growth rate  $C_G \bar{G}(t)$ , not  $\bar{G}(t)$  itself.

Why does price dampening occur? Since our one-period model implies  $C_G = 1$  and our continuous-time model implies  $C_G \rightarrow 1$  in the limit as liquidity vanishes (see Section 4.4), price

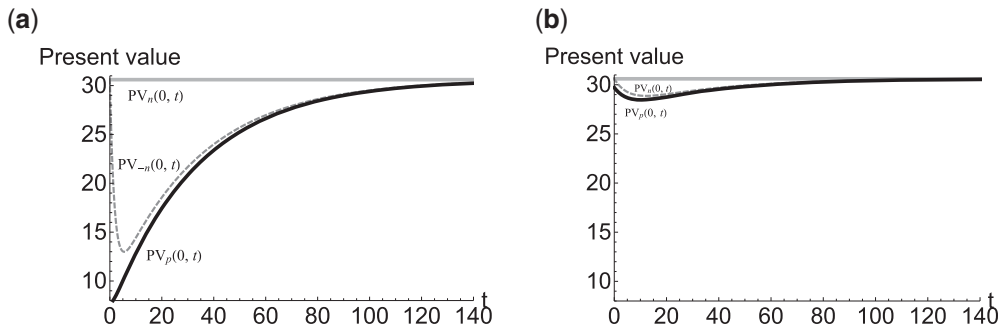


FIGURE 7

Present value of liquidating at date  $t$  using a trader's own valuation ( $PV_n(0, t)$ ), other traders' valuation ( $PV_{-n}(0, t)$ ), and the market price ( $PV_p(0, t)$ ). (a) with high disagreement; (b) with low disagreement.

dampening must involve multiple rounds of trading.<sup>11</sup> It is not based on imperfect competition because similar dampening occurs in our competitive model. It does not occur when disagreement is replaced by private values with a common prior, as in our private-values model. It also does not occur in noisy rational expectations models, such as Wang (1993), Wang (1994), and He and Wang (1995). To summarize, price dampening may result from combining different beliefs with dynamic trading in a liquid market—but not from imperfect competition, illiquidity, or asymmetric information with a common prior.

Figure 7 illustrates the intuition behind the dampening effect.<sup>12</sup> To simplify exposition, assume that the valuations  $F_n(0) > 0$  of all  $N$  traders coincide at  $t = 0$ , implying  $G_n(0) = G_{-n}(0) = \bar{G}(0) > 0$  for all  $n$ . For negative values, the figure would be symmetric.

Each panel of Figure 7 depicts graphs of three different expected present value calculations as functions of a liquidation date  $t$ ; these calculations are conditioned on information at time 0. Each function represents the present value to trader  $n$  resulting from collecting dividends on one share of stock between dates 0 and  $t$ , depositing the dividends into a money market account, selling the asset based on an assumed valuation at date  $t$ , and then discounting the resulting cash flows back to date 0 at the risk-free rate. The three graphs correspond to a different valuation assumptions at date  $t$ . All calculations are based on trader  $n$ 's beliefs by assumption, and they are identical for all traders by symmetry. Derivations and analytical proofs are provided in Appendix A.7 in equations (A81)–(A91).

First, the horizontal light solid line, denoted  $PV_n(0, t)$ , is based on the assumption that trader  $n$  liquidates the asset at date  $t$  at his own estimate of its fundamental value  $F_n(t)$ . Since trader  $n$  applies Bayes law correctly given his beliefs, the martingale property of his valuation (law of iterated expectations) makes the present value  $PV_n(0, t)$  a constant function for  $t \geq 0$ ; its graph is a horizontal line.

11. Consider rescaling the private information (21) as a scaled growth rate plus noise  $\tau_n^{-1/2} dB_n(t)$ , as in (A95), so that trader  $n$  observes  $\tau_n^{-1/2} dI_n(t)$  rather than  $dI_n(t)$ . This changes the equilibrium because traders disagree about whether to use  $\tau_H$  or  $\tau_L$  to convert one scaling into the other. We solved the equilibrium under this alternative setup. While most results remain qualitatively the same, the dampening effect disappears ( $C_G = 1$ ). The revised dynamics of  $\hat{H}_n(t)$  and  $\hat{H}_{-n}(t)$  are presented in (A96) and (A97) in Appendix A.8. Since a trader can estimate the diffusion variance with high accuracy by observing  $dI_n(t)$  over short intervals in continuous time, it violates minimal rationality for one trader to assume that another trader observes the diffusion variance of his information incorrectly.

12. Numerical calculations are based on the parameter values  $r = 0.01$ ,  $A = 1$ ,  $\alpha_D = 0.1$ ,  $\alpha_G = 0.02$ ,  $\sigma_D = 0.5$ ,  $\sigma_G = 0.01$ ,  $N = 10$ ,  $G_n(0) = 0.08$ , and  $D(0) = 0.7$ , with  $\tau_H = 0.524$ ,  $\tau_L = 0.0009$ , and  $\tau_0 = 0.0007$  in Panel (a), and  $\tau_H = 0.054$ ,  $\tau_L = 0.0098$ , and  $\tau_0 = 0.0022$  in Panel (b).

Second, the light dashed curve, denoted  $PV_{-n}(0, t)$ , is based on the assumption that trader  $n$  liquidates the asset at date  $t$  at his estimate of the average valuation of the other  $N-1$  traders,  $\frac{1}{N-1} \sum_{m=1, m \neq n}^N F_m(t)$ . The graph first falls below the horizontal line in the short run and then rises asymptotically back towards it in the long run. Even though all traders have the same estimates of fundamental value at date 0, disagreement about signal precision makes trader  $n$  believe that the other  $N-1$  traders' estimates of the growth rate will mean revert to zero at rate  $\alpha_G + (\tau_H^{1/2} - \tau_L^{1/2})^2$ , which is faster than the mean reversion rate  $\alpha_G$  he assumes for his own forecast. As a result of the higher mean-reversion rate, trader  $n$  expects that  $PV_{-n}(0, t)$  will fall in the short run. Since trader  $n$  also believes that his own initial present value calculation is correct, he expects that  $PV_{-n}(0, t)$  will rise back to his own estimate of the fundamental value in the long run.

Third, the dark solid curve, denoted  $PV_p(0, t)$ , is based on the assumption that trader  $n$  liquidates the asset at date  $t$  at the equilibrium price  $P(t)$ . Consistent with the equilibrium result  $0 < C_G < 1$ , the initial price  $P(0)$  is lower than the consensus estimate of fundamental value  $F_n(0)$ , even though all traders agree about this value and agree about how it will evolve in the future. The dampening effect nevertheless arises due to interactions among expectations of traders. If prices were equal to the consensus fundamental valuation  $F_n(0)$ , then all traders would want to hold short positions because all of them would expect prices to fall below their estimates of fundamental value in the short run as the others temporarily became more bearish. This explains why the price  $P(0)$  is dampened relative to the average valuation of fundamentals, and yet this is consistent with each trader having a target inventory of zero at  $t=0$ .

As Figure 7 illustrates and Appendix A.7 proves, only two patterns are possible for  $PV_p(0, t)$ . If  $C_G$  is less than some threshold  $\hat{C}_G := (1 + (1 - 1/N)(\tau_H^{1/2} - \tau_L^{1/2})^2 / (r + \alpha_G))^{-1}$ , then trader  $n$  expects  $PV_p(0, t)$  to increase monotonically over time, as in Figure 7a. If  $C_G$  is greater than the threshold  $\hat{C}_G$ , then he expects  $PV_p(0, t)$  first to decrease over a time interval  $\hat{t}$ , defined in equation (A93), before monotonically increasing over time, as in Figure 7b. The complicated dynamics of price-based present value  $PV_p(0, t)$  can be attributed to two factors. First, it tracks the average of  $PV_n(0, t)$  and  $PV_{-n}(0, t)$ , where  $PV_n(0, t)$  remains constant and  $PV_{-n}(0, t)$  falls in the short run and then rises back in the long run, as discussed above. Second, there is an additional effect related to the magnitude of  $C_G$ . To summarize, each trader in the smooth-trading model believes that equilibrium prices deviate from average estimates of fundamental values and do not have a martingale property.

The above discussion shows that our model captures precisely the intuition of the beauty contest of Keynes (1936):<sup>13</sup>

'For most of these persons are, in fact, largely concerned, not with making superior long-term forecasts of the probable yield on an investment over its whole life, but with foreseeing changes in the conventional basis of valuation a short time ahead of the general public.

13. "... Professional investment may be likened to those newspaper competitions in which the competitors have to pick out the six prettiest faces from a hundred photographs, the prize being awarded to the competitor whose choice most nearly corresponds to the average preferences of the competitors as a whole; so that each competitor has to pick, not those faces which he himself finds prettiest, but those which he thinks likeliest to catch the fancy of the other competitors, all of whom are looking at the problem from the same point of view. It is not a case of choosing those which, to the best of one's judgment, are really the prettiest, nor even those which average opinion genuinely thinks the prettiest. We have reached the third degree where we devote our intelligences to anticipating what average opinion expects the average opinion to be. And there are some, I believe, who practise the fourth, fifth and higher degrees." Our model implicitly assumes that all traders anticipate the expectations of other traders for arbitrarily higher degrees. Han and Kyle (2017) examine a one-period model in which, instead of agreeing to disagree, traders disagree about higher order beliefs.

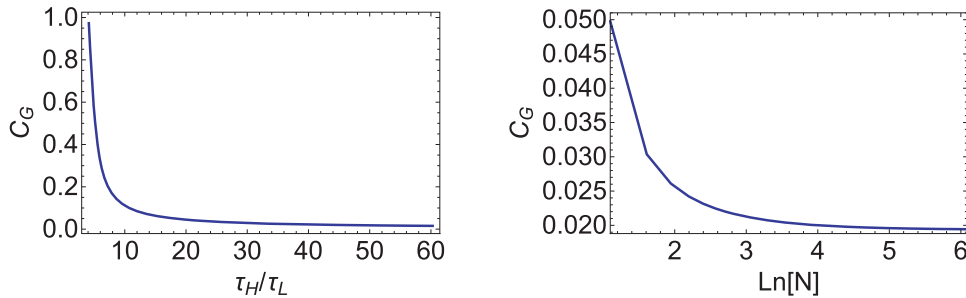


FIGURE 8

The parameter  $C_G$  as a function of disagreement (a)  $(\tau_H/\tau_L)$  and (b)  $\ln(N)$ .

They are concerned not with what an investment is really worth to a man who buys it “for keeps,” but with what the market will value it at, under the influence of mass psychology, three months or a year hence.’

As in Keynes (1936), traders in our model use trading strategies which respond to short-term price dynamics. As Keynes puts it, “it is not sensible to pay 25 for an investment of which you believe the prospective yield to justify value of 30, if you also believe that the market will value it at 20 three months hence”.

As a result of his belief that financial markets are dominated by short-term speculation rather than long-term enterprise, Keynes thought that financial markets are not too different from a casino and exhibit excessive volatility. In contrast to Keynes’ intuition, short-term trading dynamics dampens price volatility in our model relative to the volatility of fundamentals. Furthermore, prices in our model are not “noisy”; the levels of current prices and dividends are sufficient statistics for inferring the average “true” valuations of traders.

The Keynesian beauty contest requires liquidity. Consider the intuition of two special cases studied in detail in Section 4.4. When the market is very illiquid (with the degree of disagreement  $\tau_H/\tau_L$  close to the existence boundary  $\Delta_H = 0$ ), it is costly for traders to implement short-term strategies due to high temporary price impact costs, the profit opportunities based on the beauty contest are therefore too costly to exploit, the dampening effect goes away, and thus  $C_G \rightarrow 1$ . When the market is very liquid (due to large disagreement  $\tau_H/\tau_L$ ), short-term strategies are cheap, traders trade aggressively against one another’s perceived mistakes, the dampening effect is substantial, and  $C_G$  decreases with the degree of disagreement. For example, the price reduction relative to fundamentals is greater in the left panel than the right panel of Figure 7 because the degree of disagreement is larger and therefore the market is more liquid. Figure 8 shows that  $C_G$  declines as disagreement  $(\tau_H/\tau_L)$  increases and the number of traders  $(\ln(N))$  increases.

The above analysis suggests the empirical hypothesis that there is more price dampening and therefore more time series momentum in more liquid markets (see Section 4.5). Indeed, Lee and Swaminathan (2000) document that price momentum is more pronounced among high-volume stocks. Moskowitz *et al.* (2012) find significant time series momentum in equity index, currency, commodity, and bond futures markets; they show, in line with our predictions, that more liquid contracts tend to exhibit greater momentum profits. Cremers and Pareek (2015) report that momentum profits increase with decreasing stock duration, a measure of how quickly institutions turn over their positions. Our results may help to explain recent growth in assets of managed futures funds which implement trend-following strategies in liquid markets.

#### 4.4. *Special cases*

There are two special cases with closed-form solutions.

*Case 1. Vanishing Liquidity.* A closed-form solution exists for the limiting case when market liquidity vanishes (when  $\Delta_H \rightarrow 0$  implies  $\kappa \rightarrow \infty$  and  $\gamma_P \rightarrow 0$ ).

**Corollary 1.** *Assume  $\gamma_P = 0$ . Then the six equations characterizing equilibrium (A57)–(A62) have a solution if and only if  $\Delta_H = 0$ . This solution has a closed form, presented in equations (A70)–(A71), which implies  $\gamma_S = \gamma_H = \gamma_D = 0$  and  $C_G = 1$ .*

When disagreement decreases, traders offer steeper flow-supply schedules, which make each trader trade smaller quantities based on his own private information. The variable  $\gamma_P$  becomes zero at exactly the point where there is not enough disagreement to support trade ( $\Delta_H = 0$  in conjecture 1).

Our numerical results suggest that the converse is also true: when the existence condition  $\Delta_H > 0$  fails to be satisfied, market liquidity falls to zero. Since the solution to the six equations (A57)–(A62) is continuous in the exogenous parameters, this suggests that when  $\Delta_H$  is a small positive number, there will be an equilibrium with low liquidity and modest trade. As the values of the exogenous parameters  $\tau_H$ ,  $\tau_L$ , and  $N$  change so that  $\Delta_H := \tau_H^{1/2}/\tau_L^{1/2} - 2 - 2/(N-2) \rightarrow 0$ , then liquidity will vanish, the value of trading on private information will vanish, trading volume will fall to zero, and price dampening will disappear. Trading volume and price dampening disappear when  $\tau_H = \tau_L$  in the model with perfect competition, as shown in Appendix B.2.

*Case 2. Many Noise Traders.* A closed-form solution exists when traders believe other traders have no information ( $\tau_L = 0$ ) and the number of traders  $N$  is large. In the spirit of Black (1986) and Treynor (1995), the limit  $\tau_L \rightarrow 0$  implies each trader believes that all other traders trade on noise as if it were information. An equilibrium exists because disagreement is large ( $\Delta_H \rightarrow \infty$ ). For tractability, we assume that  $\tau_0$  is close to zero. Increasing the number of traders  $N$  increases competition. In the limit as  $N \rightarrow \infty$ , Appendix A.6 shows that equations (A57)–(A62) have a closed form solution, presented in equations (A73)–(A77), in which the parameters  $\gamma_P$  and  $\gamma_S$  are proportional to  $N$ . Markets become infinitely liquid. Permanent impact satisfies  $\lambda \rightarrow 0$ , temporary impact satisfies  $\kappa \rightarrow 0$ , the speed of inventory adjustment satisfies  $\gamma_S \rightarrow \infty$ , and price dampening is substantial, with  $C_G$  satisfying  $\lim_{N \rightarrow \infty} C_G = (r + \alpha_G)/(r + \alpha_G + \tau) < 1$ .

These two closed-form solutions further support the conclusion that increasing disagreement makes markets more liquid and price dampening more pronounced.

#### 4.5. *Empirically correct beliefs*

Our results on trading strategies, price impact, and price levels depend on traders' different subjective priors about signal precision, not on the objective, true, or "empirically correct" priors. As long as each trader believes that his own private information has high precision  $\tau_H$  and the other traders' private information has low precision  $\tau_L$  with  $\tau_H > \tau_L \geq 0$ , our results do not depend on whether each trader assumes the empirically correct precision for his signal but underestimates the precision of others' signals or each trader overestimates the empirically correct precision for his own signal but assumes the empirically correct precision of others' signals. In contrast, understanding the empirically correct dynamic properties of prices and quantities, measured by a non-trader such as an economist or econometrician, requires knowing the true values of the parameters.



Consider three symmetric alternative specifications for empirically correct probabilities. In the first specification, the empirically correct precision of all traders' signals is  $\tau_H$ ; each trader assumes the correct precision for his own signal but underestimates the precision of others' signals. In the second specification, the empirically correct precision of all traders' signals is  $\tau_L$ ; each trader assumes the correct precision of others' signals but overestimates the precision for his own signal. In the third specification, the empirically correct precisions are such that traders are "correct on average" in the sense that each trader's precision is  $\tau_I = \frac{1}{N}(\tau_H + (N-1)\tau_L)$ . Eyster *et al.* (2015) refer to the first case as "dismissiveness" and to the second case as "overconfidence". For simplicity, assume traders' beliefs about the parameters  $N$ ,  $\alpha_G$ ,  $\sigma_G$ ,  $\alpha_D$ , and  $\sigma_D$  are correct. *Quantities.* The following theorem holds for all three cases.

**Theorem 6. (Empirical Implications for Inventories).** *When the empirically correct precision of all signals has the same value, target inventories  $S_n^{TI}(t)$  and actual inventories  $S_n(t)$  follow a linear bivariate process (defined in equation (A99) in Appendix A.9). The autocorrelation function for actual inventories and the correlation between actual inventories and target inventories are*

$$\text{Corr}[S_n(t), S_n(t + \Delta t)] = \frac{(\alpha_G + \tau)e^{-\gamma_S \Delta t} - \gamma_S e^{-(\alpha_G + \tau)\Delta t}}{\alpha_G + \tau - \gamma_S}, \quad (50)$$

$$\text{Corr}[S_n(t), S_n^{TI}(t)] = \left( \frac{\gamma_S}{\alpha_G + \tau + \gamma_S} \right)^{1/2}. \quad (51)$$

The bivariate inventory process is fully characterized by the three parameters  $\alpha_G + \tau$ ,  $\gamma_S$ , and  $C_L$ . Target inventories  $S_n^{TI}(t)$  follow a univariate AR-1 process. Symmetry implies that both  $S_n(t)$  and  $S_n^{TI}(t)$  are distributed independently from prices and have correlation  $-1/(N-1)$  across traders. Since the model is Gaussian, the auto-correlation function completely describes the statistical properties of inventories. Equation (51) implies that the correlation between the actual inventory and target inventory increases when markets become more liquid ( $\gamma_S$  increases).

Consistent with the empirical evidence discussed in the introduction, the auto-correlation of actual inventories tends to be smaller in more liquid markets: Equation (50) implies that  $\text{Corr}[S_n(t), S_n(t + \Delta t)]$  decreases in  $\gamma_S$  if  $\alpha_G + \tau < \gamma_S$  or if  $\alpha_G + \tau > \gamma_S$  and  $\Delta t > 1/(\alpha_G + \tau - \gamma_S)$ . In more liquid markets, actual inventories are further away from past actual inventories and closer to target inventories.

We have emphasized that our model provides a link between the half-life of information in prices and the half-life of traders' target inventories related to  $\alpha_G + \tau$  and  $\gamma_S$ , respectively. The auto-correlation function (50) has the following interesting symmetry: The auto-correlation function does not change if the values of  $\alpha_G + \tau$  and  $\gamma_S$  are interchanged. The symmetry property of  $\gamma_S$  and  $\alpha_G + \tau$  implies that the inventory process in a model with rapidly mean-reverting target inventories and slow convergence of actual inventories to target inventories (large  $\alpha_G + \tau$  and small  $\gamma_S$ ) is observationally equivalent to the inventory process in a different model with slowly mean-reverting target inventories and fast convergence of actual inventories to target inventories (small  $\alpha_G + \tau$  and large  $\gamma_S$ ). In both cases, actual inventories will change slowly and appear to be almost non-stationary. Empirically, actual inventories  $S_n(t)$  are more likely to be observable than target inventories  $S_n^{TI}(t)$ . If it is known whether  $\alpha_G + \tau$  is greater or less than  $\gamma_S$ , then the values of both  $\alpha_G + \tau$  and  $\gamma_S$  can be inferred from the autocorrelation of actual inventories  $S_n(t)$  in equation (50).

*Prices.* While the price in equation (40) does not depend on empirically correct precisions, expected returns depend on both traders' precisions and empirically correct precisions. In the first specification with "dismissiveness", the true total precision is bigger than the total precision



assumed by traders, so the econometrician will assign lower weights than traders to the distant information when constructing his own signals. In the second specification with “overconfidence”, the true total precision is lower than the total precision assumed by traders, so the econometrician will assign bigger weights to the distant information. These two cases are complicated to analyse because it is necessary to track the econometrician’s signals separately since they are different from signals constructed by traders. In the third specification where traders are correct on average, the total precision attributed to information by traders coincides with the true precision, and the econometrician will agree with traders on how to construct signals. In Appendix A.10, we show numerically then that the auto-covariance of returns is positive for a large range of parameter values and prove analytically that it is positive for the limiting case with  $\tau_L = 0$  discussed in Section 4.4. This implies time series momentum in the sense that higher returns in the past tend to be followed by higher returns in the future. We also show numerically that the auto-correlation tends to increase with disagreement, implying that the momentum is more pronounced in more liquid markets. Kyle *et al.* (2016) provide a detailed analysis of return predictability in a competitive setting with a general specification for empirically correct beliefs.

## 5. CONCLUSION: IMPLICATIONS FOR PRACTICE

We have described a steady-state model of continuous trading in which relatively overconfident traders have market power. This model provides a framework for thinking about how the dynamics of trading affect market liquidity, transaction costs, and market prices. It helps to analyse temporary and permanent price impact. The model provides a realistic framework for understanding how inventory and price dynamics are affected by permanent and temporary price impact when large traders optimize their trading to beat the market.

In their influential book for quantitative asset managers, Grinold and Kahn (1994) describe a partial-equilibrium trading model with decaying private information, risk aversion, and temporary price impact. They pose as an important open research question (p. 580) how to set up a proper trading model with a finite half-life for signals, risk aversion, and transaction costs with components of tightness, depth, and resiliency. Our equilibrium model not only solves an appropriate optimization problem for all asset managers simultaneously but also derives endogenously a realistic transaction cost model with stationary dynamics for inventories, prices, and expected returns.

Our model of smooth order flow implements ideas about market liquidity described informally by Black (1995). Black envisioned a future frictionless market for exchanges as “an equilibrium in which traders use indexed limit orders at different levels of urgency but do not use market orders or conventional limit orders”. In that equilibrium, there is no conventional liquidity available for market orders and conventional limit orders. Orders are executed gradually and move the price by an amount increasing in the level of urgency. Order-shredding algorithms have been incorporating the idea of urgency for years. In popular algorithms based on VWAP (Volume Weighted Average Price), a trader chooses a target number of shares to trade, a time frame (say one day), and a participation rate (say 5% of volume); the higher the participation rate, the greater the trader’s impatience.

The idea that securities markets offer a flow equilibrium rather than a stock equilibrium may seem far-fetched at first glance. Yet, recent trends in the way liquidity is supplied and demanded in electronic markets are consistent with optimal trading strategies in our model. The reduction in tick size to one cent in 2001 reduced size available at the best bid and offer. Our model predicts vanishingly small market depth to be available at a given point in time; instead, market depth is made available only over time. A positive tick size probably increases instantaneous depth relative to the predictions of our model.

In the future, centralized exchanges may change order matching rules to implement limit orders conforming to the intuition of our model. For example, exchanges might consolidate order flow into frequent batch auctions, say once per second, consistent with Budish *et al.* (2015). Alternatively, a time parameter could be added to limit orders, allowing smooth trading strategies to be implemented using simple message types without heavy message traffic. For example, a trader who might in today's market place a limit order to buy 10,000 shares at a price of \$40 per share might instead enter an order to purchase one share per second at a price of \$40 or better for the next 10,000 seconds. Such new order types would allow traders to implement smooth trading strategies without generating the heavy message traffic associated with submitting and canceling thousands of conventional limit orders.

## APPENDIX

### A. PROOFS

#### A.1. Proof of Theorem 1

Under the tentative assumption that trader  $n$  knows the value of  $i_{-n}$ , plug equation (4) into equation (9) and use the first-order condition to find his optimal demand:

$$x_n = \frac{\frac{\tau_v^{1/2}}{\tau} \left( \tau_0^{1/2} i_0 + \tau_H^{1/2} i_n + (N-1) \tau_L^{1/2} i_{-n} \right) - \left( \frac{\alpha}{\gamma} i_0 + \frac{\beta}{\gamma} i_{-n} \right) - \left( \frac{\delta}{(N-1)\gamma} + \frac{A}{\tau} \right) S_n}{\frac{2}{(N-1)\gamma} + \frac{A}{\tau}}. \quad (\text{A1})$$

In the numerator of this equation, the first term is trader  $n$ 's expectation of the liquidation value, the second term is the market-clearing price when trader  $n$  trades a quantity of zero and has no inventory, and the last term is the adjustment for existing inventory. In the denominator, the first and second terms reflect how trader  $n$  restricts the quantity traded due to market power and risk aversion, respectively.

As in Kyle (1989), even though trader  $n$  does not observe  $i_{-n}$  explicitly, he is still able to implement this optimal strategy with a demand schedule which implicitly infers  $i_{-n}$  from the market-clearing price.

Define the constant

$$C := \frac{1}{(N-1)\gamma} + \frac{A}{\tau} + \frac{\tau_L^{1/2} \tau_v^{1/2}}{\tau \beta}. \quad (\text{A2})$$

Solving for  $i_{-n}$  instead of  $p$  in the market-clearing condition (3), substituting this solution into equation (A1) above, and then solving for  $x_n$ , yields a demand schedule  $X_n(i_0, i_n, S_n, p)$  for trader  $n$  as a function of price  $p$ :

$$\begin{aligned} X_n(i_0, i_n, S_n, p) = & \frac{1}{C} \left( \frac{\tau_v^{1/2}}{\tau} \left( \tau_0^{1/2} - (N-1) \tau_L^{1/2} \frac{\alpha}{\beta} \right) i_0 + \frac{\tau_H^{1/2}}{\tau} \tau_v^{1/2} i_n \right. \\ & \left. + \left( \frac{(N-1) \tau_L^{1/2} \gamma \tau_v^{1/2}}{\tau \beta} - 1 \right) p - \left( \frac{\delta \tau_L^{1/2} \tau_v^{1/2}}{\tau \beta} + \frac{A}{\tau} \right) S_n \right). \end{aligned} \quad (\text{A3})$$

In a symmetric linear equilibrium, the strategy chosen by trader  $n$  must be the same as the linear strategy (2) conjectured for the other traders. Equating the corresponding coefficients of the variables  $i_0$ ,  $i_n$ ,  $p$ , and  $S_n$  yields a system of four equations in terms of the four unknowns  $\alpha$ ,  $\beta$ ,  $\gamma$ , and  $\delta$ :

$$\alpha = \frac{\tau_v^{1/2}}{C} \left( \frac{\tau_0^{1/2}}{\tau} - \frac{(N-1) \tau_L^{1/2}}{\tau} \frac{\alpha}{\beta} \right), \quad \beta = \frac{\tau_v^{1/2}}{C} \frac{\tau_H^{1/2}}{\tau}, \quad (\text{A4})$$

$$\gamma = -\frac{1}{C} \left( \frac{(N-1) \tau_L^{1/2}}{\tau} \frac{\gamma}{\beta} \tau_v^{1/2} - 1 \right), \quad \delta = \frac{1}{C} \left( \frac{\tau_L^{1/2}}{\tau} \frac{\delta}{\beta} \tau_v^{1/2} + \frac{A}{\tau} \right). \quad (\text{A5})$$

The unique solution is

$$\beta = \frac{(N-2) \tau_H^{1/2} - 2(N-1) \tau_L^{1/2}}{A(N-1)} \tau_v^{1/2}, \quad (\text{A6})$$

$$\alpha = \frac{\tau_0^{1/2}}{\tau_H^{1/2} + (N-1) \tau_L^{1/2}} \beta, \quad \gamma = \frac{\tau}{\tau_H^{1/2} + (N-1) \tau_L^{1/2}} \frac{\beta}{\tau_v^{1/2}}, \quad \delta = \frac{A}{\tau_H^{1/2} - \tau_L^{1/2}} \frac{\beta}{\tau_v^{1/2}}.$$

Substituting (A6) into (A1) yields trader  $n$ 's optimal demand (11). Substituting (11) into (4) yields the equilibrium price (13).

The second-order condition has the correct sign if and only if  $\frac{2}{(N-1)\gamma} + \frac{A}{\tau} > 0$ . Given the definition  $\Delta_H := \tau_H^{1/2}/\tau_L^{1/2} - 2 - 2/(N-2)$ , this is equivalent to

$$\frac{A}{\tau} \frac{N}{N-2} \frac{\tau_H^{1/2}}{\tau_L^{1/2}} \frac{1}{\Delta_H} > 0. \quad (\text{A7})$$

Therefore, assuming  $N > 2$ , the second-order condition holds if and only if  $\Delta_H > 0$ .<sup>14</sup>

## A.2. Bayesian updating with signals of arbitrary precision

This section derives signal processing formulas for arbitrary “generic” beliefs  $\bar{\tau}_0, \bar{\tau}_1, \dots, \bar{\tau}_N$  about signal precisions.

Define  $G(t) = \mathbb{E}_t[G^*(t)]$ , where the subscript  $t$  denotes conditioning on the history of the signals  $I_0(s), \dots, I_N(s)$  for  $s \in (-\infty, t]$ . Without loss of generality, let  $\bar{\Omega}$  denote the error variance  $\bar{\Omega} := \text{Var}[(G^*(t) - G(t))/\sigma_G]$ . Assume a steady state in which  $\bar{\Omega}$  is a constant which does not depend on time. Like a squared Sharpe ratio,  $\bar{\Omega}$  measures the error variance in units of time. For example, if time is measured in years,  $\bar{\Omega} = 4$  means that the estimate of  $G^*(t)$  is “behind” the true value of  $G^*(t)$  by an amount equivalent to four years of volatility unfolding at rate  $\sigma_G$ . There are simple and intuitive formulas for information processing:

**Lemma 1.** *Given generic beliefs  $\bar{\tau}_1, \dots, \bar{\tau}_N$ , let  $\bar{\tau}$  denote the sum of precisions*

$$\bar{\tau} := \bar{\tau}_0 + \sum_{n=1}^N \bar{\tau}_n. \quad (\text{A8})$$

Then  $\bar{\Omega}$  and  $dG(t)$  satisfy

$$\bar{\Omega}^{-1} := \text{Var}^{-1} \left[ \frac{G^*(t) - G(t)}{\sigma_G} \right] = 2\alpha_G + \bar{\tau}, \quad (\text{A9})$$

$$dG(t) = -(\alpha_G + \bar{\tau}) G(t) dt + \sigma_G \bar{\Omega}^{1/2} \sum_{n=0}^N \bar{\tau}_n^{1/2} dI_n. \quad (\text{A10})$$

*Proof.* Apply the Stratonovich–Kalman–Bucy filter to the filtering problem summarized by equation (20) for signals and by equations (21) and (24) for observations. This yields the filtering estimate defined by the Itô differential equation

$$dG(t) = -\alpha_G G(t) dt + \sum_{n=0}^N \sigma_G^2 \bar{\Omega} \frac{\bar{\tau}_n^{1/2}}{\sigma_G \bar{\Omega}^{1/2}} \left( dI_n(t) - G(t) \frac{\bar{\tau}_n^{1/2}}{\sigma_G \bar{\Omega}^{1/2}} dt \right). \quad (\text{A11})$$

Rearranging terms yields equation (A10). The mean-square filtering error of the estimate  $G(t)$ , denoted  $\sigma_G^2 \bar{\Omega}(t)$ , is defined by the Riccati differential equation

$$\sigma_G^2 \frac{d\bar{\Omega}(t)}{dt} = -2\alpha_G \sigma_G^2 \bar{\Omega}(t) + \sigma_G^2 - \sigma_G^4 \bar{\Omega}(t)^2 \sum_{n=0}^N \left( \frac{\bar{\tau}_n^{1/2}}{\sigma_G \bar{\Omega}(t)^{1/2}} \right)^2. \quad (\text{A12})$$

Let  $\bar{\Omega}$  denote the steady state of the function of time  $\bar{\Omega}(t)$ . Using the steady-state assumption  $d\bar{\Omega}(t)/dt = 0$ , solve the second equation for the steady state value  $\bar{\Omega} = \bar{\Omega}(t)$  to obtain equation (A9).  $\parallel$

14. When there is not enough disagreement to sustain an equilibrium with trade using pure strategies, one might imagine that it is possible to have an equilibrium with trade using mixed strategies. For mixed strategies to generate an equilibrium with trade, the trader must be indifferent across the various randomized choices of quantities he trades. For example, if we add normally distributed noise to quantities traded, symmetrically across all traders, a mixed strategy equilibrium requires the second-order condition to be exactly zero. This means that the quadratic objective function reduces to a linear function, so that the denominator of equation (A1) is zero. Since the trader has to be indifferent across various randomizations, this further implies that the linear function must be a constant, distributed independently from the quantity traded. This assumption cannot hold, because a trader with a positive value of  $i_n$  would always want to buy unlimited quantities, and a trader with a negative  $i_n$  would always want to sell unlimited quantities. Thus, an equilibrium with symmetric normally distributed noise cannot generate trade. When noise is not normally distributed or the equilibrium is not symmetric, the objective function is not quadratic any more, but it will still be difficult to find a mixed strategy equilibrium with trade, given that the sensitivity of utility to a trader’s own private information must be well-defined.

The error variance  $\bar{\Omega}$  corresponds to a steady state that balances an increase in error variance due to innovations  $dB_G(t)$  in the true growth rate with a reduction in error variance due to (1) mean reversion of the true growth rate at rate  $\alpha_G$  and (2) arrival of new information with total precision  $\bar{\tau}$ .

Note that  $\bar{\Omega}$  is not a free parameter but is instead determined as an endogenous function of the other parameters. Equation (A9) implies that  $\bar{\Omega}$  turns out to be the solution to the quadratic equation  $\bar{\Omega}^{-1} = 2\alpha_G + \bar{\Omega}\sigma_G^2/\sigma_D^2 + \sum_{n=1}^N \bar{\tau}_n$ . In equations (21) and (24), we scaled the units with which precision is measured by the endogenous parameter  $\bar{\Omega}$  because this leads to simpler filtering expressions which more clearly bring out intuition about signal processing.

From equation (A10), the estimate  $G(t)$  can be conveniently written as the weighted sum of  $N+1$  sufficient statistics  $H_n(t)$  corresponding to  $N+1$  information flows  $dI_n$ . Define the sufficient statistics or “signals”  $H_n(t)$  by

$$H_n(t) := \int_{u=-\infty}^t e^{-(\alpha_G + \bar{\tau})(t-u)} dI_n(u), \quad n=0, 1, \dots, N, \quad (\text{A13})$$

which implies

$$dH_n(t) = -(\alpha_G + \bar{\tau})H_n(t)dt + dI_n(t), \quad n=0, 1, \dots, N. \quad (\text{A14})$$

Then  $G(t)$  becomes a linear combination of sufficient statistics  $H_n(t)$  with weights proportional to the square roots of the precisions  $\bar{\tau}_n^{1/2}$ :

$$G(t) = \sigma_G \bar{\Omega}^{1/2} \sum_{n=0}^N \bar{\tau}_n^{1/2} H_n(t). \quad (\text{A15})$$

The importance of each bit of information  $dI_n$  about the growth rate  $G(t)$  decays exponentially at a rate  $\alpha_G + \bar{\tau}$ , which is the same for all of the signals. The half-life of a signal  $\ln 2/(\alpha_G + \bar{\tau})$  decreases as “aggregate precision”  $\bar{\tau}$  increases. Even though the true unobserved growth rate may have a long half-life (small  $\alpha_G$ ), signals predicting this growth rate decay rapidly if aggregate precision  $\bar{\tau}$  is large.

Equations (21), (24), and (A10) imply that the estimate  $G(t)$  mean-reverts to zero at a rate  $\alpha_G$  while moving towards the true value  $G^*(t)$  at rate  $\bar{\tau}$ :

$$dG(t) = -\alpha_G G(t)dt + \bar{\tau} (G^*(t) - G(t))dt + \sigma_G \bar{\Omega}^{1/2} \sum_{n=0}^N \bar{\tau}_n^{1/2} dB_n(t). \quad (\text{A16})$$

### A.3. Proof of Theorem 3

Let  $\mathbb{E}_t^n[\dots]$  denote the conditional expectations operator  $\mathbb{E}[\dots|\mathcal{F}_n(t)]$  based on trader  $n$ 's beliefs. Let  $J^n(\mathcal{F}_n(t); X_n, C_n; X_m, m \neq n)$  denote the expected utility trader  $n$  receives as a function of his own consumption and trading strategies  $(C_n, X_n)$  and the  $N-1$  other traders' trading strategies  $(X_m)$ , conditional on his information set  $\mathcal{F}_n(t)$ . In this particular model, exponential utility functions with fixed interest rates make it unnecessary for  $J^n(\dots)$  to depend on other traders' consumption strategies. Define an *equilibrium* as a set of trading strategies  $X_1^*, \dots, X_N^*$  and consumption strategies  $C_1^*, \dots, C_N^*$  such that, for  $n=1, \dots, N$ , trader  $n$ 's optimal consumption and trading strategies  $X_n = X_n^*$  and  $C_n = C_n^*$  solve the maximization problem

$$J^n(\mathcal{F}_n(t); X_n^*, C_n^*; X_m^*, m \neq n) = \max_{[C_n, X_n]} \mathbb{E}_t^n \left[ \int_{s=t}^{\infty} e^{-\rho(s-t)} U(C_n(s)) ds \right], \quad (\text{A17})$$

subject to inventories following

$$dS_n(t) = x_n(t)dt \quad (\text{A18})$$

and money holdings following

$$dM_n(t) = (rM_n(t) + S_n(t)D(t) - c_n(t) - P(t)x_n(t))dt. \quad (\text{A19})$$

Conditional on any information set  $\mathcal{F}_n(t)$ , trader  $n$ 's money and asset holdings must, with probability one, satisfy the budget constraint

$$\liminf_{T \rightarrow \infty} \mathbb{E}_t^n \left[ e^{-r(T-t)} M_n(T) + \int_{u=T}^{\infty} e^{-r(u-t)} D(u) S_n(T) du \right] \geq 0. \quad (\text{A20})$$

In equations (A18) and (A19), the price  $P(t)$ , quantity  $x_n(t)$ , and consumption  $c_n(t)$  are the abbreviations

$$P(t) := P[X_1, \dots, X_N](t), \quad x_n(t) := \frac{dS_n(t)}{dt} = X_n(t, P(t); \mathcal{F}_n(t)), \quad c_n(t) := C_n(t; \mathcal{F}_n(t)). \quad (\text{A21})$$

This implies that trader  $n$  takes as given the strategies of the other traders when he chooses his optimal strategy. It also implies that when he chooses his optimal strategy, he takes into account how his strategy choice affects the price at which he trades and his trading opportunities in the future.

The budget constraint (A20) says that if the trader calculates the fundamental value of his wealth using information at time  $t$ , then he does not engage in Ponzi finance. Also note that the optimal strategy will satisfy the transversality condition  $\mathbb{E}_t^n[e^{-\rho(T-t)} J^n(\mathcal{F}_n(T), X_n^*, C_n^*, \dots)] \rightarrow 0$  as  $T \rightarrow \infty$ .

Instead of calculating the solution  $J^n(\dots)$  directly, we use the no regret approach, which assumes that trader  $n$  observes his residual supply schedule at each point in time, then picks an optimal point on the residual supply schedule. We then show that the solution to this less constrained problem also implements the optimal solution to the more constrained problem which defines  $J^n(\dots)$ .

For the less constrained problem, we conjecture a steady-state value function of the form  $V(M_n, S_n, D, H_0, H_n, H_{-n})$ , where  $M_n$  denotes trader  $n$ 's cash holdings (measured in dollars) and  $S_n$  denotes trader  $n$ 's holdings of the traded asset (measured in shares).

In a competitive model, a trader's value function depends on his wealth but does not depend on the decomposition of his wealth into his various security holdings. With imperfect competition, the decomposition of a trader's wealth into various security holdings does affect his value function because the trader cannot costlessly convert one security holding into cash or another security holding by trading at market prices. Wealth does not appear in the value function because wealth is not well-defined. Trader  $n$  is always influencing the mark-to-market value of his risky inventory by choosing his rate of trading. It is therefore necessary to keep track of the two components of wealth—cash  $M_n$  and inventories  $S_n$ —separately.

Also, we expect the asset price to be a linear combination of two components: (1) a dividend level component linear in dividend flow  $D(t)$  and (2) a dividend-growth component linear in the variables  $H_0(t)$ ,  $H_n(t)$ , and  $H_{-n}(t)$ . Given the symmetric linear conjectured form of the residual supply function, observing the average of other traders' signals  $H_{-n}(t)$  is informationally equivalent to observing the intercept of the residual supply schedule (when  $S_n(t) = S'_n(t) = 0$ ). Therefore we include  $H_{-n}(t)$  as a state variable in the value function and omit the price  $P(t)$ .

The values of all ten exogenous parameters  $\alpha_D$ ,  $\sigma_D$ ,  $\alpha_G$ ,  $\sigma_G$ ,  $\tau_H$ ,  $\tau_L$ ,  $N$ ,  $r$ ,  $A$ , and  $\rho$  are common knowledge. It is also common knowledge that each trader believes that  $dB_D(t)$ ,  $dB_G(t)$ ,  $dB_1(t), \dots, dB_N(t)$  are independently distributed Brownian motions, given traders' beliefs. Note that since traders disagree about whether a signal has precision  $\tau_H$  or  $\tau_L$ , they also disagree about how to construct the Brownian motions  $dB_n(t)$  from the information  $dI_n(t)$ . Symmetry of parameter values prevents the number of state variables from exploding, avoiding the forecasting-the-forecasts-of-others problem described by Townsend (1983).

In the derivations below, mathematical notation is simplified if the three state variables  $H_0(t)$ ,  $H_n(t)$ , and  $H_{-n}(t)$  are replaced with two composite signals, denoted  $\hat{H}_n(t)$  and  $\hat{H}_{-n}(t)$ . Repeating equations (29), define the weighting constant  $\hat{a}$  by

$$\hat{a} := \frac{\tau_0^{1/2}}{\tau_H^{1/2} + (N-1)\tau_L^{1/2}}, \quad (\text{A22})$$

and define the two composite signals  $\hat{H}_n(t)$  and  $\hat{H}_{-n}(t)$  by

$$\hat{H}_n(t) := H_n(t) + \hat{a}H_0(t), \quad (\text{A23})$$

$$\hat{H}_{-n}(t) := H_{-n}(t) + \hat{a}H_0(t). \quad (\text{A24})$$

Trader  $n$ 's estimate of the dividend growth rate can now be expressed as a function of the two composite signals  $\hat{H}_n(t)$  and  $\hat{H}_{-n}(t)$  as

$$G_n(t) = \sigma_G \Omega^{1/2} \left( \tau_H^{1/2} \hat{H}_n(t) + (N-1)\tau_L^{1/2} \hat{H}_{-n}(t) \right). \quad (\text{A25})$$

Define the  $N+1$  processes  $dB_0^n$ ,  $dB_m^n$ , and  $dB_n^n$ ,  $m = 1, \dots, N$ ,  $m \neq n$ , by

$$dB_0^n(t) = \tau_0^{1/2} \frac{G^*(t) - G_n(t)}{\sigma_G \Omega^{1/2}} dt + dB_D(t), \quad (\text{A26})$$

$$dB_n^n(t) = \tau_H^{1/2} \frac{G^*(t) - G_n(t)}{\sigma_G \Omega^{1/2}} dt + dB_n(t), \quad (\text{A27})$$

and

$$dB_m^n(t) = \tau_L^{1/2} \frac{G^*(t) - G_n(t)}{\sigma_G \Omega^{1/2}} dt + dB_m(t). \quad (\text{A28})$$

The superscript  $n$  indicates conditioning on beliefs of trader  $n$ . Since trader  $n$ 's forecast of the error  $G^*(t) - G_n(t)$  is zero given his information set, these  $N+1$  processes are independently distributed Brownian motions from the perspective of trader  $n$ . In terms of these Brownian motions, trader  $n$  believes that signals change as follows:

$$dH_0(t) = -(\alpha_G + \tau)H_0(t)dt + \tau_0^{1/2} \frac{G_n(t)}{\sigma_G \Omega^{1/2}} dt + dB_0^n(t), \quad (\text{A29})$$

$$dH_n(t) = -(\alpha_G + \tau)H_n(t)dt + \tau_H^{1/2} \frac{G_n(t)}{\sigma_G \Omega^{1/2}} dt + dB_n^n(t), \quad (\text{A30})$$

$$dH_{-n}(t) = -(\alpha_G + \tau)H_{-n}(t)dt + \tau_L^{1/2} \frac{G_n(t)}{\sigma_G \Omega^{1/2}} dt + \frac{1}{N-1} \sum_{\substack{m=1 \\ m \neq n}}^N dB_m^n(t). \quad (\text{A31})$$

Note that each signal drifts towards zero at rate  $\alpha_G + \tau$  and drifts towards the optimal forecast  $G_n(t)$  at a rate proportional to the square root of the signal's precision  $\tau_0^{1/2}$ ,  $\tau_H^{1/2}$ , or  $\tau_L^{1/2}$ , respectively.

In terms of the composite variables  $\hat{H}_n$  and  $\hat{H}_{-n}$ , we conjecture (and verify below) a steady-state value function of the form  $V(M_n, S_n, D, \hat{H}_n, \hat{H}_{-n})$ . Letting  $(c_n(t), x_n(t))$  denote consumption and investment choices, write

$$V(M_n, S_n, D, \hat{H}_n, \hat{H}_{-n}) := \max_{[c_n(t), x_n(t)]} \mathbb{E}_t^n \left[ \int_{s=t}^{\infty} -e^{-\rho(s-t) - A c_n(s)} ds \right], \quad (\text{A32})$$

where  $P(x_n(t))$  is a linear function of  $x_n(t)$  given by equation (32), dividends follow equation (19), inventories follow  $dS_n(t) = x_n(t)dt$ , the change in cash holdings  $dM_n(t)$  is a quadratic function of  $x_n(t)$  following

$$dM_n(t) = (rM_n(t) + S_n(t)D(t) - c_n(t) - P(x_n(t))x_n(t))dt, \quad (\text{A33})$$

and signals  $\hat{H}_n$  and  $\hat{H}_{-n}$  follow a bivariate vector auto-regression given by

$$\begin{aligned} d\hat{H}_n(t) = & -(\alpha_G + \tau)\hat{H}_n(t)dt \\ & + (\tau_H^{1/2} + \hat{a}\tau_0^{1/2}) \left( \tau_H^{1/2} \hat{H}_n(t) + (N-1)\tau_L^{1/2} \hat{H}_{-n}(t) \right) dt \\ & + \hat{a}dB_0^n(t) + dB_n^n(t), \end{aligned} \quad (\text{A34})$$

$$\begin{aligned} d\hat{H}_{-n}(t) = & -(\alpha_G + \tau)\hat{H}_{-n}(t)dt \\ & + (\tau_L^{1/2} + \hat{a}\tau_0^{1/2}) \left( \tau_H^{1/2} \hat{H}_n(t) + (N-1)\tau_L^{1/2} \hat{H}_{-n}(t) \right) dt \\ & + \hat{a}dB_0^n(t) + \frac{1}{N-1} \sum_{\substack{m=1 \\ m \neq n}}^N dB_m^n(t). \end{aligned} \quad (\text{A35})$$

The dynamics of  $\hat{H}_n$  and  $\hat{H}_{-n}$  in equations (A34) and (A35) can be derived from equations (A29), (A30), and (A31).

Note that the coefficient  $\tau_H^{1/2} + \hat{a}\tau_0^{1/2}$  in the second line of equation (A34) is different from the coefficient  $\tau_L^{1/2} + \hat{a}\tau_0^{1/2}$  in the second line of equation (A35). This difference is the key driving force behind the price-dampening effect resulting from the Keynesian beauty contest. It captures the fact that—in addition to disagreeing about the value of the asset in the present—traders also disagree about the dynamics of their future valuations. As shown in equations (C33) and (C34) in Supplementary Appendix C.3, these two different coefficients are the same in an otherwise similar private-values model. As a result, prices are not dampened in the private-values model.

Using the definition of  $G_n(t)$  in equation (28) and the definition of  $\hat{a}$  in equation (29), it can be shown that trader  $n$  believes the stochastic process  $\hat{H}_n - \hat{H}_{-n}$  satisfies

$$\begin{aligned} d(\hat{H}_n - \hat{H}_{-n}) = & -(\alpha_G + \tau)(\hat{H}_n - \hat{H}_{-n})dt + \frac{\tau_H^{1/2} - \tau_L^{1/2}}{\sigma_G \Omega^{1/2}} G_n(t)dt \\ & + dB_n^n(t) - \frac{1}{N-1} \sum_{\substack{m=1 \\ m \neq n}}^N dB_m^n(t). \end{aligned} \quad (\text{A36})$$

In equation (A36), the term with  $G_n(t)dt$  implies that each trader believes that  $\hat{H}_n - \hat{H}_{-n}$  does not follow an AR-1 process. Because traders have different expectations  $G_n(t)$ , they agree in the present about how they will disagree in the future. If traders only disagreed about the value of  $G_n(t)$  in the present but agreed about the evolution of  $\hat{H}_n - \hat{H}_{-n}$  in the future, then the coefficient of the  $G_n(t)dt$  term would be zero,  $\hat{H}_n - \hat{H}_{-n}$  would follow an AR-1 (Ornstein–Uhlenbeck) process, and traders would not disagree about the dynamics of process  $\hat{H}_n - \hat{H}_{-n}$ . In the otherwise similar model with private values, the term involving  $G_n(t)dt$  becomes zero.

We conjecture and verify that the value function  $V(M_n, S_n, D, \hat{H}_n, \hat{H}_{-n})$  has the specific quadratic exponential form

$$V(M_n, S_n, D, \hat{H}_n, \hat{H}_{-n}) = -\exp\left(\psi_0 + \psi_M M_n + \frac{1}{2} \psi_{SS} S_n^2 + \psi_{SD} S_n D + \psi_{Sn} S_n \hat{H}_n + \psi_{Sx} S_n \hat{H}_{-n} + \frac{1}{2} \psi_{nn} \hat{H}_n^2 + \frac{1}{2} \psi_{xx} \hat{H}_{-n}^2 + \psi_{nx} \hat{H}_n \hat{H}_{-n}\right). \quad (\text{A37})$$

The nine constants  $\psi_0$ ,  $\psi_M$ ,  $\psi_{SS}$ ,  $\psi_{SD}$ ,  $\psi_{Sn}$ ,  $\psi_{Sx}$ ,  $\psi_{nn}$ ,  $\psi_{xx}$ , and  $\psi_{nx}$  have values consistent with a steady-state equilibrium. The term  $\psi_M$  measures the utility value of cash. The terms  $\psi_{SS}$ ,  $\psi_{SD}$ ,  $\psi_{Sn}$ , and  $\psi_{Sx}$  measure the utility value of risky asset holdings. The terms  $\psi_{nn}$ ,  $\psi_{xx}$ , and  $\psi_{nx}$  capture the value of future trading opportunities based on current public and private information. The value of trading on innovations to future information is built into the constant term  $\psi_0$ .

The Hamilton–Jacobi–Bellman (HJB) equation corresponding to the conjectured value function  $V(M_n, S_n, D, \hat{H}_n, \hat{H}_{-n})$  in equation (A32) is

$$\begin{aligned} 0 = \max_{c_n, x_n} & \left[ U(c_n) - \rho V + \frac{\partial V}{\partial M_n} (rM_n + S_n D - c_n - P(x_n)x_n) + \frac{\partial V}{\partial S_n} x_n \right] \\ & + \frac{\partial V}{\partial D} \left( -\alpha_D D + \sigma_G \Omega^{1/2} \tau_H^{1/2} \hat{H}_n + \sigma_G \Omega^{1/2} (N-1) \tau_L^{1/2} \hat{H}_{-n} \right) \\ & + \frac{\partial V}{\partial \hat{H}_n} \left( -(\alpha_G + \tau) \hat{H}_n(t) + (\tau_H^{1/2} + \hat{a} \tau_0^{1/2}) (\tau_H^{1/2} \hat{H}_n + (N-1) \tau_L^{1/2} \hat{H}_{-n}) \right) \\ & + \frac{\partial V}{\partial \hat{H}_{-n}} \left( -(\alpha_G + \tau) \hat{H}_{-n}(t) + (\tau_L^{1/2} + \hat{a} \tau_0^{1/2}) (\tau_H^{1/2} \hat{H}_n + (N-1) \tau_L^{1/2} \hat{H}_{-n}) \right) + \frac{1}{2} \frac{\partial^2 V}{\partial D^2} \sigma_D^2 \\ & + \frac{1}{2} \frac{\partial^2 V}{\partial \hat{H}_n^2} (1 + \hat{a}^2) + \frac{1}{2} \frac{\partial^2 V}{\partial \hat{H}_{-n}^2} \left( \frac{1}{N-1} + \hat{a}^2 \right) + \left( \frac{\partial^2 V}{\partial D \partial \hat{H}_n} + \frac{\partial^2 V}{\partial D \partial \hat{H}_{-n}} \right) \hat{a} \sigma_D + \frac{\partial^2 V}{\partial \hat{H}_n \partial \hat{H}_{-n}} \hat{a}^2. \end{aligned} \quad (\text{A38})$$

For the specific quadratic specification of the value function in equation (A37), the HJB equation becomes

$$\begin{aligned} 0 = \min_{c_n, x_n} & \left[ -\frac{e^{-Ac_n}}{V} - \rho + \psi_M (rM_n + S_n D - c_n - P(x_n)x_n) \right. \\ & \left. + (\psi_{SS} S_n + \psi_{SD} D + \psi_{Sn} \hat{H}_n + \psi_{Sx} \hat{H}_{-n}) x_n \right] \\ & + \psi_{SD} S_n (-\alpha_D D + \sigma_G \Omega^{1/2} \tau_H^{1/2} \hat{H}_n + \sigma_G \Omega^{1/2} (N-1) \tau_L^{1/2} \hat{H}_{-n}) \\ & + \left( \psi_{Sn} S_n + \psi_{nn} \hat{H}_n + \psi_{nx} \hat{H}_{-n} \right) \left( -(\alpha_G + \tau) \hat{H}_n(t) + (\tau_H^{1/2} + \hat{a} \tau_0^{1/2}) (\tau_H^{1/2} \hat{H}_n + (N-1) \tau_L^{1/2} \hat{H}_{-n}) \right) \\ & + \left( \psi_{Sx} S_n + \psi_{xx} \hat{H}_{-n} + \psi_{nx} \hat{H}_n \right) \left( -(\alpha_G + \tau) \hat{H}_{-n}(t) + (\tau_L^{1/2} + \hat{a} \tau_0^{1/2}) (\tau_H^{1/2} \hat{H}_n + (N-1) \tau_L^{1/2} \hat{H}_{-n}) \right) \\ & + \frac{1}{2} \psi_{SD}^2 S_n^2 \sigma_D^2 + \frac{1}{2} \left( (\psi_{Sn} S_n + \psi_{nn} \hat{H}_n + \psi_{nx} \hat{H}_{-n})^2 + \psi_{nn} \right) (1 + \hat{a}^2) \\ & + \frac{1}{2} \left( (\psi_{Sx} S_n + \psi_{xx} \hat{H}_{-n} + \psi_{nx} \hat{H}_n)^2 + \psi_{xx} \right) \left( \frac{1}{N-1} + \hat{a}^2 \right) \\ & + \left( (\psi_{Sn} + \psi_{Sx}) S_n + (\psi_{nn} + \psi_{nx}) \hat{H}_n + (\psi_{xx} + \psi_{nx}) \hat{H}_{-n} \right) \psi_{SD} S_n \hat{a} \sigma_D \\ & + \left( (\psi_{Sn} S_n + \psi_{nn} \hat{H}_n + \psi_{nx} \hat{H}_{-n}) (\psi_{Sx} S_n + \psi_{xx} \hat{H}_{-n} + \psi_{nx} \hat{H}_n) + \psi_{nx} \right) \hat{a}^2. \end{aligned} \quad (\text{A39})$$

The solution for optimal consumption is

$$c_n^*(t) = -\frac{1}{A} \log \left( \frac{\psi_M V(t)}{A} \right). \quad (\text{A40})$$

In the HJB equation (A39), the price  $P(x_n)$  is linear in  $x_n$  based on equation (32). Plugging  $P(x_n)$  from equation (32) into the HJB equation (A39) yields a quadratic function of  $x_n$  which captures the effect of trader  $n$ 's trading rate  $x_n$  on



prices. The optimal trading strategy is a linear function of the state variables given by

$$x_n^*(t) = \frac{(N-1)\gamma_P}{2\psi_M} \left( \left( \psi_{SD} - \frac{\psi_M \gamma_D}{\gamma_P} \right) D(t) + \left( \psi_{SS} - \frac{\psi_M \gamma_S}{(N-1)\gamma_P} \right) S_n(t) \right. \\ \left. + \psi_{Sn} \hat{H}_n(t) + \left( \psi_{Sx} - \frac{\psi_M \gamma_H}{\gamma_P} \right) \hat{H}_{-n}(t) \right). \quad (\text{A41})$$

Because the exponent of the conjectured value function is a quadratic function of the state variables, the best linear strategy will dominate any non-linear strategy or a mixed strategy.

The derivation of this optimal trading strategy assumes that trader  $n$  observes the values of  $D(t)$ ,  $S_n(t)$ ,  $\hat{H}_n(t)$ , and  $\hat{H}_{-n}(t)$ . Although trader  $n$  does not actually observe  $\hat{H}_{-n}(t)$ , he can implement the optimal quantity  $x_n^*$  by submitting an appropriate linear demand schedule. We can think of this demand schedule as a linear function of  $P(t)$  whose intercept is a linear function of  $D(t)$ ,  $S_n(t)$ , and  $\hat{H}_n(t)$ . Trader  $n$  can infer from the market-clearing condition (31) that  $\hat{H}_{-n}$  is given by

$$\hat{H}_{-n}(t) = \frac{\gamma_P}{\gamma_H} \left( P(t) - D(t) \frac{\gamma_D}{\gamma_P} \right) - \frac{1}{(N-1)\gamma_H} x_n^*(t) - \frac{\gamma_S}{(N-1)\gamma_H} S_n(t). \quad (\text{A42})$$

Plugging equation (A42) into equation (A41) and solving for  $x_n^*(t)$  implements the optimal trading strategy  $x_n^*(t)$  as a linear demand schedule which depends on the price  $P(t)$  and state variables  $\hat{H}_n$ ,  $S_n(t)$ , and  $D(t)$ , which the trader directly observes. This schedule is given by

$$x_n^*(t) = \frac{(N-1)\gamma_P}{\psi_M} \left( 1 + \frac{\psi_{Sx} \gamma_P}{\psi_M \gamma_H} \right)^{-1} \\ \cdot \left( \left( \psi_{SD} - \psi_{Sx} \frac{\gamma_D}{\gamma_H} \right) D(t) + \left( \psi_{SS} - \psi_{Sx} \frac{\gamma_S}{(N-1)\gamma_H} \right) S_n(t) \right. \\ \left. + \psi_{Sn} \hat{H}_n(t) + \left( \psi_{Sx} \frac{\gamma_P}{\gamma_H} - \psi_M \right) P(t) \right). \quad (\text{A43})$$

Symmetry requires that this demand schedule be the same as the demand schedule conjectured for the  $N-1$  other traders. Equating the coefficients of  $D(t)$ ,  $\hat{H}_n(t)$ ,  $S_n(t)$ , and  $P(t)$  in equation (A43) to the conjectured coefficients  $\gamma_D$ ,  $\gamma_H$ ,  $-\gamma_S$ , and  $-\gamma_P$  results in the following four restrictions that the values of the  $\gamma$ -parameters and  $\psi$ -parameters must satisfy in a symmetric equilibrium with linear trading strategies:

$$\frac{(N-1)\gamma_P}{\psi_M} \left( 1 + \frac{\psi_{Sx} \gamma_P}{\psi_M \gamma_H} \right)^{-1} \left( \psi_{SD} - \psi_{Sx} \frac{\gamma_D}{\gamma_H} \right) = \gamma_D, \quad (\text{A44})$$

$$\frac{(N-1)\gamma_P}{\psi_M} \left( 1 + \frac{\psi_{Sx} \gamma_P}{\psi_M \gamma_H} \right)^{-1} \psi_{Sn} = \gamma_H, \quad (\text{A45})$$

$$\frac{(N-1)\gamma_P}{\psi_M} \left( 1 + \frac{\psi_{Sx} \gamma_P}{\psi_M \gamma_H} \right)^{-1} \left( \psi_{SS} - \psi_{Sx} \frac{\gamma_S}{(N-1)\gamma_H} \right) = -\gamma_S, \quad (\text{A46})$$

$$\frac{(N-1)\gamma_P}{\psi_M} \left( 1 + \frac{\psi_{Sx} \gamma_P}{\psi_M \gamma_H} \right)^{-1} \left( \psi_{Sx} \frac{\gamma_P}{\gamma_H} - \psi_M \right) = -\gamma_P. \quad (\text{A47})$$

Note that it is not possible to solve this system for the four  $\gamma$ -parameters  $\gamma_H$ ,  $\gamma_S$ ,  $\gamma_D$ , and  $\gamma_P$  because this system of four equations can be written so that the four  $\gamma$ -parameters enter only as the three ratios  $\gamma_H/\gamma_P$ ,  $\gamma_S/\gamma_P$ , and  $\gamma_D/\gamma_P$ . Therefore, we solve the system instead for the four unknowns  $\psi_{Sx}$ ,  $\gamma_H$ ,  $\gamma_S$ , and  $\gamma_D$ . The solution is

$$\psi_{Sx} = \frac{N-2}{2} \psi_{Sn}, \quad \gamma_H = \frac{N\gamma_P}{2\psi_M} \psi_{Sn}, \quad \gamma_S = -\frac{(N-1)\gamma_P}{\psi_M} \psi_{SS}, \quad \gamma_D = \frac{\gamma_P}{\psi_M} \psi_{SD}. \quad (\text{A48})$$

Define the constants  $C_L$  and  $C_G$  by

$$C_L := -\frac{\psi_{Sn}}{2\psi_{SS}}, \quad C_G := \frac{\psi_{Sn}}{2\psi_M} \frac{N(r+\alpha_D)(r+\alpha_G)}{\sigma_G \Omega^{1/2} (\tau_H^{1/2} + (N-1)\tau_L^{1/2})}. \quad (\text{A49})$$

When  $\gamma_D$  in equation (A48) is plugged into equation (A41), the coefficient on  $D(t)$  zeros out; this implies that traders will not trade on public information. It is intuitively obvious that traders cannot trade on the basis of the public information

$D(t)$  because all traders would want to trade in the same direction and this would be inconsistent with market clearing. Substituting equation (A48) into equation (A41) yields the solution for optimal strategy.

$$x_n^*(t) = \gamma_S \left( C_L (H_n(t) - H_{-n}(t)) - S_n(t) \right). \quad (\text{A50})$$

Define the average of traders' expected growth rates  $\bar{G}(t)$  by

$$\bar{G}(t) := \frac{1}{N} \sum_{n=1}^N G_n(t), \quad (\text{A51})$$

Then, the equilibrium price is

$$P^*(t) = \frac{D(t)}{r + \alpha_D} + \frac{C_G \bar{G}(t)}{(r + \alpha_D)(r + \alpha_G)}. \quad (\text{A52})$$

One might expect that the solution of the maximization problem would yield solutions for the nine  $\psi$ -parameters as functions of the four  $\gamma$ -parameters. One might also expect that imposing symmetry by equating the four optimal  $\gamma$ -parameters (implied by trader  $n$ 's optimal trading strategy) to the four conjectured  $\gamma$ -parameters would yield solutions for the four  $\gamma$ -parameters as functions of the nine  $\psi$ -parameters. In principle, one could then expect a solution to the thirteen equations in thirteen unknowns to describe a steady-state equilibrium, if one exists.

Although this is the intuition for the solution methodology, the solution does not work in this straightforward manner. As mentioned above, the four equations for the  $\gamma$ -parameters do not determine  $\gamma_P$  as a function of the nine  $\psi$ -parameters. Instead, the solution to the four  $\gamma$ -equations (A48) implies a restriction on the  $\psi$ -parameters (the first of equations (A48)), which must hold in a steady-state equilibrium. This restriction insures that the incentives to demand and supply liquidity are balanced, but it does not define a level of liquidity  $\gamma_P$ .

Plugging (A40) and (A41) back into the Bellman equation and setting the constant term and the coefficients of  $M_n$ ,  $S_n D$ ,  $S_n^2$ ,  $S_n \hat{H}_n$ ,  $S_n \hat{H}_{-n}$ ,  $\hat{H}_n^2$ ,  $\hat{H}_{-n}^2$ , and  $\hat{H}_n \hat{H}_{-n}$  to be zero, we obtain nine equations. Using the first equation (A48) to substitute  $\psi_{S_n}$  for  $\psi_{S_x}$ , there are in total nine equations in nine unknowns  $\gamma_P$ ,  $\psi_0$ ,  $\psi_M$ ,  $\psi_{SD}$ ,  $\psi_{SS}$ ,  $\psi_{S_n}$ ,  $\psi_{nn}$ ,  $\psi_{xx}$ , and  $\psi_{nx}$ .

By setting the constant term, coefficient of  $M$ , and coefficient of  $SD$  to be zero, we obtain

$$\psi_M = -rA, \quad (\text{A53})$$

$$\psi_{SD} = -\frac{rA}{r + \alpha_D}, \quad (\text{A54})$$

$$\psi_0 = 1 - \log(r) + \frac{1}{r} \left( -\rho + \frac{1}{2}(1 + \hat{a}^2)\psi_{nn} + \frac{1}{2} \left( \frac{1}{N-1} + \hat{a}^2 \right) \psi_{xx} + \hat{a}^2 \psi_{nx} \right). \quad (\text{A55})$$

In addition, by setting the coefficients of  $S_n^2$ ,  $S_n \hat{H}_n$ ,  $S_n \hat{H}_{-n}$ ,  $\hat{H}_n^2$ ,  $\hat{H}_{-n}^2$ , and  $\hat{H}_n \hat{H}_{-n}$  to be zero, we obtain six polynomial equations in the six unknowns  $\gamma_P$ ,  $\psi_{SS}$ ,  $\psi_{S_n}$ ,  $\psi_{nn}$ ,  $\psi_{xx}$ , and  $\psi_{nx}$ . Defining the constants  $a_1$ ,  $a_2$ ,  $a_3$ , and  $a_4$  by

$$\begin{aligned} a_1 &:= -\alpha_G - \tau + \tau_H^{1/2}(\tau_H^{1/2} + \hat{a}\tau_0^{1/2}), \\ a_2 &:= -\alpha_G - \tau + (N-1)\tau_L^{1/2}(\tau_L^{1/2} + \hat{a}\tau_0^{1/2}), \\ a_3 &:= (\tau_H^{1/2} + \hat{a}\tau_0^{1/2})(N-1)\tau_L^{1/2}, \\ a_4 &:= (\tau_L^{1/2} + \hat{a}\tau_0^{1/2})\tau_H^{1/2}, \end{aligned} \quad (\text{A56})$$

these six equations in six unknowns can be written

$$\begin{aligned} 0 &= -\frac{1}{2}r\psi_{SS} - \frac{\gamma_P(N-1)}{rA}\psi_{SS}^2 + \frac{r^2A^2\sigma_D^2}{2(r+\alpha_D)^2} + \frac{1}{2}(1+\hat{a}^2)\psi_{S_n}^2 \\ &\quad + \frac{1}{2} \left( \frac{1}{N-1} + \hat{a}^2 \right) \frac{(N-2)^2}{4}\psi_{S_n}^2 - \frac{rA}{r+\alpha_D}\hat{a}\sigma_D\frac{N}{2}\psi_{S_n} + \hat{a}^2\frac{N-2}{2}\psi_{S_n}^2, \end{aligned} \quad (\text{A57})$$

$$\begin{aligned} 0 &= -r\psi_{S_n} - \frac{\gamma_P(N-1)}{rA}\psi_{SS}\psi_{S_n} - \frac{rA}{r+\alpha_D}\sigma_G\Omega^{1/2}\tau_H^{1/2} + a_1\psi_{S_n} \\ &\quad + \frac{N-2}{2}a_4\psi_{S_n} + (1+\hat{a}^2)\psi_{nn}\psi_{S_n} + \frac{N-2}{2} \left( \frac{1}{N-1} + \hat{a}^2 \right) \psi_{nx}\psi_{S_n} \\ &\quad - \frac{rA}{r+\alpha_D}\hat{a}\sigma_D(\psi_{nn} + \psi_{nx}) + \hat{a}^2\psi_{nx}\psi_{S_n} + \frac{N-2}{2}\hat{a}^2\psi_{nn}\psi_{S_n}, \end{aligned} \quad (\text{A58})$$

$$\begin{aligned}
0 = & -r \frac{N-2}{2} \psi_{Sn} + \frac{\gamma_P(N-1)}{rA} \psi_{SS} \psi_{Sn} - \frac{rA}{r+\alpha_D} \sigma_G \Omega^{1/2} (N-1) \tau_L^{1/2} \\
& + \left( a_3 + \frac{N-2}{2} a_2 \right) \psi_{Sn} + (1 + \hat{a}^2) \psi_{Sn} \psi_{nx} + \frac{N-2}{2} \left( \frac{1}{N-1} + \hat{a}^2 \right) \psi_{xx} \psi_{Sn} \\
& - \frac{rA}{r+\alpha_D} \hat{a} \sigma_D (\psi_{xx} + \psi_{nx}) + \hat{a}^2 \psi_{xx} \psi_{Sn} + \frac{N-2}{2} \hat{a}^2 \psi_{nx} \psi_{Sn},
\end{aligned} \tag{A59}$$

$$\begin{aligned}
0 = & -\frac{r}{2} \psi_{nn} - \frac{\gamma_P(N-1)}{4rA} \psi_{Sn}^2 + a_1 \psi_{nn} + a_4 \psi_{nx} + \frac{1}{2} (1 + \hat{a}^2) \psi_{nn}^2 \\
& + \frac{1}{2} \left( \frac{1}{N-1} + \hat{a}^2 \right) \psi_{nx}^2 + \hat{a}^2 \psi_{nn} \psi_{nx},
\end{aligned} \tag{A60}$$

$$\begin{aligned}
0 = & -\frac{r}{2} \psi_{xx} - \frac{\gamma_P(N-1)}{4rA} \psi_{Sn}^2 + a_2 \psi_{xx} + a_3 \psi_{nx} + \frac{1 + \hat{a}^2}{2} \psi_{nx}^2 \\
& + \frac{1}{2} \left( \frac{1}{N-1} + \hat{a}^2 \right) \psi_{xx}^2 + \hat{a}^2 \psi_{xx} \psi_{nx},
\end{aligned} \tag{A61}$$

$$\begin{aligned}
0 = & -r \psi_{nx} + \frac{\gamma_P(N-1)}{2rA} \psi_{Sn}^2 + a_3 \psi_{nn} + a_4 \psi_{xx} + (a_1 + a_2) \psi_{nx} \\
& + (1 + \hat{a}^2) \psi_{nn} \psi_{nx} + \left( \frac{1}{N-1} + \hat{a}^2 \right) \psi_{xx} \psi_{nx} + \hat{a}^2 (\psi_{nn} \psi_{xx} + \psi_{nx}^2).
\end{aligned} \tag{A62}$$

We have not discovered a simple closed-form solution for equations (A57)–(A62); instead, we attempt to solve these equations numerically.

Equations (A57)–(A62) are necessary but not sufficient conditions for steady-state equilibrium with symmetric, linear flow-strategies. For a solution to the six polynomial equations to define a stationary equilibrium, it is sufficient for the solution to satisfy (1) a second-order condition implying  $\gamma_P > 0$ , (2) a stationarity condition implying  $\gamma_S > 0$ , (3) a transversality condition requiring  $r > 0$ , and (4) a budget constraint ruling out Ponzi schemes (implied by  $r > 0$  and stationarity of inventories).

(1) The second-order condition requires  $\gamma_P > 0$ . For the minimum in the optimization problem (A39) to exist, the second-order condition requires the  $2 \times 2$  matrix

$$\begin{pmatrix} -\frac{A^2}{V} & 0 \\ 0 & \frac{2rA}{(N-1)\gamma_P} \end{pmatrix} \tag{A63}$$

to be positive definite. Since the value function  $V$  is negative, this condition holds if and only if  $\gamma_P > 0$ . This is equivalent to requiring downward-sloping flow-demand schedules; it is also equivalent to requiring temporary price impact to be positive.

(2) If  $\gamma_P > 0$  but  $\gamma_S < 0$ , then permanent price impact slopes the wrong way. Each trader's inventories grow exponentially over time, violating the requirement that inventories have a stationary distribution.

(3) The transversality condition for the value function  $V(\dots)$  is

$$\lim_{T \rightarrow +\infty} \mathbb{E}_t^n \left[ e^{-\rho(T-t)} V \left( M_n(T), S_n(T), D(T), \hat{H}_n(T), \hat{H}_{-n}(T) \right) \right] = 0. \tag{A64}$$

From the HJB equation and equations (A57)–(A62), we have

$$\begin{aligned}
\mathbb{E}_t^n \left[ dV \left( M_n(t), S_n(t), D(t), \hat{H}_n(t), \hat{H}_{-n}(t) \right) \right] = \\
-(r - \rho) V \left( M_n(t), S_n(t), D(t), \hat{H}_n(t), \hat{H}_{-n}(t) \right) dt.
\end{aligned} \tag{A65}$$

This yields

$$\begin{aligned}
\mathbb{E}_t^n \left[ e^{-\rho(T-t)} V \left( M_n(T), S_n(T), D(T), \hat{H}_n(T), \hat{H}_{-n}(T) \right) \right] = \\
e^{-r(T-t)} V \left( M_n(t), S_n(t), D(t), \hat{H}_n(t), \hat{H}_{-n}(t) \right),
\end{aligned} \tag{A66}$$

which implies that the transversality condition (A64) is satisfied if  $r > 0$ .

(4) The budget constraint ruling out Ponzi schemes (A20) is automatically satisfied if  $r > 0$  and the state variables are stationary.

Since the model assumes  $r > 0$ , the inequalities  $\gamma_P > 0$  and  $\gamma_S > 0$  are necessary and sufficient conditions for a solution to the six equations to characterize the desired equilibrium.

Under the assumptions  $\gamma_P > 0$  and  $\gamma_S > 0$ , analytical results imply  $\gamma_D > 0$ ,  $\psi_M < 0$ , and  $\psi_{SD} < 0$ , consistent with the intuition that traders prefer more to less; we also obtain  $\psi_{SS} > 0$ , consistent with the intuition that traders are averse to inventory risk. Our numerical results indicate that all endogenous parameters have the intuitively correct signs. For example, numerical results indicate that  $\gamma_H > 0$ ,  $\psi_{Sn} < 0$ ,  $\psi_{Sx} < 0$ ,  $\psi_{nm} < 0$ , and  $\psi_{xx} < 0$ , consistent with the intuition that traders buy when they have bullish information, value greater expected dividends, and make greater profits (whether long or short) from more extreme signals. The sign of  $\psi_{nx}$  is intuitively and numerically ambiguous.

#### A.4. Proof of Theorem 4

Let a vector  $(\gamma_P^*, \psi_{SS}^*, \psi_{Sn}^*, \psi_{nm}^*, \psi_{nx}^*, \psi_{xx}^*)$  be a solution to the system (A57)–(A62) for exogenous parameters  $A$ ,  $\sigma_D$ ,  $\sigma_G$ ,  $r$ ,  $\alpha_G$ ,  $\alpha_D$ ,  $\tau_0$ ,  $\tau_L$ , and  $\tau_H$ . If risk aversion is rescaled by factor  $F$  from  $A$  to  $A/F$  and other exogenous parameters are kept unchanged, then it is straightforward to show that a vector  $(\gamma_P^*F, \psi_{SS}^*/F^2, \psi_{Sn}^*/F, \psi_{nm}^*, \psi_{nx}^*, \psi_{xx}^*)$  is the solution to the system (A57)–(A62). From equations (A48), (A49), and (33), it then follows that  $C_L$  changes to  $C_L F$ ,  $\lambda$  changes to  $\lambda/F$ ,  $\kappa$  changes to  $\kappa/F$ , but  $\gamma_S$  and  $C_G$  remain the same.

#### A.5. Proof of Corollary 1

With  $\gamma_P = 0$ , it is clear that  $\psi_{nm} = \psi_{nx} = \psi_{xx} = 0$  solves the last three equations (A60)–(A62) of the six equations (A57)–(A62), consistent with the intuition that information has no value if there is no market liquidity. With  $\gamma_P = \psi_{nm} = \psi_{xx} = \psi_{nx} = 0$ , then the first three equations (A57)–(A59) become

$$0 = -\frac{1}{2}r\psi_{SS} + \frac{r^2A^2\sigma_D^2}{2(r+\alpha_D)^2} + \frac{1}{2}\left(\frac{1}{N-1} + \hat{a}^2\right)\frac{(N-2)^2}{4}\psi_{Sn}^2 + \frac{1}{2}(1+\hat{a}^2)\psi_{Sn}^2 - \frac{rA}{r+\alpha_D}\hat{a}\sigma_D\frac{N}{2}\psi_{Sn} + \hat{a}^2\frac{N-2}{2}\psi_{Sn}^2, \quad (\text{A67})$$

$$0 = -r\psi_{Sn} - \frac{rA}{r+\alpha_D}\sigma_G\Omega^{1/2}\tau_H^{1/2} + a_1\psi_{Sn} + \frac{N-2}{2}a_4\psi_{Sn}, \quad (\text{A68})$$

$$0 = -r\frac{N-2}{2}\psi_{Sn} - \frac{rA}{r+\alpha_D}\sigma_G\Omega^{1/2}(N-1)\tau_L^{1/2} + \left(a_3 + \frac{N-2}{2}a_2\right)\psi_{Sn}. \quad (\text{A69})$$

Equations (A68) and (A69) are both linear equations in  $\psi_{Sn}$ . They have the same solution if and only if the existence condition is satisfied as an equality,  $\Delta_H = 0$ , in which case the unique solution for  $\psi_{Sn}$  is

$$\psi_{Sn} = -\frac{rA\sigma_G\Omega^{1/2}\tau_H^{1/2}}{(r+\alpha_D)(r+\alpha_G)}. \quad (\text{A70})$$

Substituting (A70) into (A67) yields

$$\psi_{SS} = \frac{rA^2}{(r+\alpha_D)^2} \left( \left( \sigma_D + \frac{\sigma_G\Omega^{1/2}\tau_0^{1/2}}{r+\alpha_G} \right)^2 + \frac{(\tau-\tau_0)\sigma_G^2\Omega}{(r+\alpha_G)^2} \right). \quad (\text{A71})$$

This implies  $C_G = 1$ :

$$C_G = -\frac{\psi_{Sn}}{2rA} \frac{N(r+\alpha_D)(r+\alpha_G)}{\sigma_G\Omega^{1/2}(\tau_H^{1/2} + (N-1)\tau_L^{1/2})} = \frac{N\tau_H^{1/2}}{2(\tau_H^{1/2} + (N-1)\tau_L^{1/2})} = 1. \quad (\text{A72})$$

#### A.6. Limiting case with $N \rightarrow \infty$ , $\tau_L = 0$ , and $\hat{a} \rightarrow 0$

Set  $\tau_L = 0$ , and then evaluate the solution in the limit as  $N \rightarrow \infty$  and  $\hat{a} \rightarrow 0$ . We conjecture and verify that  $\gamma_P = N\bar{\gamma}_P$ ,  $\psi_{Sn} = N^{-1}\bar{\psi}_{Sn}$ ,  $\psi_{SS} = N^{-1}\bar{\psi}_{SS}$ ,  $\psi_{nm} = \bar{\psi}_{nm}$ ,  $\psi_{nx} = \bar{\psi}_{nx}$ , and  $\psi_{xx} = \bar{\psi}_{xx}$ , where  $\bar{\gamma}_P$ ,  $\bar{\psi}_{Sn}$ ,  $\bar{\psi}_{SS}$ ,  $\bar{\psi}_{nm}$ ,  $\bar{\psi}_{nx}$ , and  $\bar{\psi}_{xx}$  are constants that do not depend on  $N$ .

Solving the system of equations (A57)–(A62) yields

$$\bar{\psi}_{Sn} = -\frac{2Ar\Omega^{1/2}\sigma_G\tau_H^{1/2}}{(r+\alpha_D)(r+\alpha_G+\tau)}, \quad \bar{\psi}_{SS} = \frac{A^2r^2\sigma_D^2}{(r+\alpha_D)^2(r+\alpha_G+\tau)}, \quad (\text{A73})$$

$$\bar{\gamma}_P = \frac{(r+\alpha_D)^2(r+\alpha_G+\tau)^2}{2Ar\sigma_D^2}, \quad (\text{A74})$$

$$\bar{\psi}_{nn} = \frac{1}{2} \left( r + 2(\alpha_G + \tau - \tau_H) - \left( (r + 2(\alpha_G + \tau - \tau_H))^2 + \frac{4\Omega\sigma_G^2\tau_H}{\sigma_D^2} \right)^{1/2} \right), \quad (\text{A75})$$

$$\bar{\psi}_{nx} = \frac{\Omega\sigma_G^2\tau_H/\sigma_D^2}{r + 2(\alpha_G + \tau) - \tau_H - \bar{\psi}_{nn}}, \quad (\text{A76})$$

$$\bar{\psi}_{xx} = \frac{1}{r + 2\alpha_G + 2\tau} \left( \bar{\psi}_{nx}^2 - \frac{\Omega\sigma_G^2\tau_H}{\sigma_D^2} \right). \quad (\text{A77})$$

This implies

$$C_G \rightarrow \frac{r+\alpha_G}{r+\alpha_G+\tau} < 1, \quad \lambda \rightarrow 0, \quad \kappa \rightarrow 0, \quad (\text{A78})$$

$$C_L = \frac{\Omega^{1/2}\sigma_G\tau_H^{1/2}(r+\alpha_D)}{Ar\sigma_D^2}, \quad (\text{A79})$$

$$\gamma_S = \frac{(N-1)\bar{\gamma}_P}{rA} \bar{\psi}_{SS} = \frac{(N-1)(r+\alpha_G+\tau)}{2} \rightarrow \infty. \quad (\text{A80})$$

#### A.7. Dampening effect: the present value of expected cumulative dividends and cash flow

From (A25), (A29), (A30), and (A31), we can derive the stochastic process for  $G_n(t)$  and  $G_{-n}(t) := \frac{1}{N-1} \sum_{m=1, m \neq n}^N G_m(t)$  as follows:

$$dG_n(t) = -\alpha_G G_n(t)dt + \sigma_G \Omega^{1/2} \left( \tau_0^{1/2} dB_0^n(t) + \tau_H^{1/2} dB_n^n(t) + \tau_L^{1/2} \sum_{m=1, m \neq n}^N dB_m^n(t) \right), \quad (\text{A81})$$

$$dG_{-n}(t) = -(\alpha_G + \tau)G_{-n}(t)dt + \left( \tau_0 + \tau_L^{1/2} \left( 2\tau_H^{1/2} + (N-2)\tau_L^{1/2} \right) \right) G_n(t)dt \\ + \sigma_G \Omega^{1/2} \left( \tau_0^{1/2} dB_0^n(t) + \tau_L^{1/2} dB_n^n(t) + \frac{\tau_H^{1/2} + (N-2)\tau_L^{1/2}}{N-1} \sum_{m=1, m \neq n}^N dB_m^n(t) \right). \quad (\text{A82})$$

From (A82), when  $G_m(t) = G_n(t)$ , trader  $n$  believes that other traders' estimates of expected growth rates  $G_m(t)$  will mean-revert to zero at a rate  $\alpha_G + (\tau_H^{1/2} - \tau_L^{1/2})^2 > \alpha_G$ . From (A81), trader  $n$  believes that his own estimate of expected growth rate  $G_n(t)$  will mean-revert to zero at a rate  $\alpha_G$ .

From (A81), (A82), and (19), the expected dynamics of  $G_n(t)$ ,  $G_{-n}(t)$ , and  $D(t)$  are given by

$$\mathbb{E}_0^n[G_n(t)] = e^{-\alpha_G t} G_n(0), \quad (\text{A83})$$

$$\mathbb{E}_0^n[G_{-n}(t)] = \frac{\tau_0 + \tau_L^{1/2} \left( 2\tau_H^{1/2} + (N-2)\tau_L^{1/2} \right)}{\tau} \left( e^{-\alpha_G t} - e^{-(\alpha_G + \tau)t} \right) G_n(0) + e^{-(\alpha_G + \tau)t} G_{-n}(0), \quad (\text{A84})$$

$$\mathbb{E}_0^n[D(t)] = \frac{1}{\alpha_D - \alpha_G} \left( e^{-\alpha_G t} - e^{-\alpha_D t} \right) G_n(0) + e^{-\alpha_D t} D(0). \quad (\text{A85})$$

The present value of expected cumulative dividends and cash flow from liquidating one share of the stock at date  $t$  using trader  $n$ 's valuation is

$$PV_n(0, t) := \mathbb{E}_0^n \left[ \int_0^t e^{-ru} D(u) du + e^{-rt} \left( \frac{D(t)}{r + \alpha_D} + \frac{G_n(t)}{(r + \alpha_D)(r + \alpha_G)} \right) \right]. \quad (\text{A86})$$

Substituting (A83) and (A85) into (A86), it can be shown that (A86) is equal to

$$F_n(0) = \frac{D(0)}{r + \alpha_D} + \frac{G_n(0)}{(r + \alpha_D)(r + \alpha_G)}. \quad (\text{A87})$$

The present value of expected cumulative dividends and cash flow from liquidating one share of the stock at date  $t$  using others' valuations  $\sum_{m=1, m \neq n}^N F_m(t)/(N-1)$  is

$$PV_{-n}(0, t) := \mathbb{E}_0^n \left[ \int_0^t e^{-ru} D(u) du + e^{-rt} \left( \frac{D(t)}{r + \alpha_D} + \frac{G_{-n}(t)}{(r + \alpha_D)(r + \alpha_G)} \right) \right]. \quad (\text{A88})$$

Assuming  $G_m(0) = G_n(0) = \bar{G}(0)$  and substituting (A83)–(A85) into (A88), it can be shown that equation (A88) is equal to

$$PV_{-n}(0, t) = F_n(0) + \frac{(\tau_H^{1/2} - \tau_L^{1/2})^2}{\tau(r + \alpha_G)(r + \alpha_D)} \left( e^{-(r + \alpha_G + \tau)t} - e^{-(r + \alpha_G)t} \right) G_n(0). \quad (\text{A89})$$

Similarly, the present value of expected cumulative dividends and cash flow from liquidating one share of the stock at date  $t$  at the equilibrium price  $P(t)$  is

$$PV_p(0, t) := \mathbb{E}_0^n \left[ \int_0^t e^{-ru} D(u) du + e^{-rt} \left( \frac{D(t)}{r + \alpha_D} + \frac{C_G \bar{G}(t)}{(r + \alpha_D)(r + \alpha_G)} \right) \right]. \quad (\text{A90})$$

Substituting (A83)–(A85) into (A90), it can be shown that (A90) is equivalent to

$$\begin{aligned} PV_p(0, t) = F_n(0) + & \frac{C_G \left( N - (\tau_H^{1/2} - \tau_L^{1/2})^2 \tau^{-1} (N-1) \right) - N}{N(r + \alpha_G)(r + \alpha_D)} e^{-(r + \alpha_G)t} G_n(0) \\ & + \frac{C_G (\tau_H^{1/2} - \tau_L^{1/2})^2 \tau^{-1} (N-1)}{N(r + \alpha_G)(r + \alpha_D)} e^{-(r + \alpha_G + \tau)t} G_n(0). \end{aligned} \quad (\text{A91})$$

From (A91), it follows that

$$\begin{aligned} \frac{dPV_p(0, t)}{dt} = & \frac{G_n(0) e^{-(r + \alpha_G)t}}{N(r + \alpha_G)(r + \alpha_D)} \left( \left( N - C_G \left( N - (\tau_H^{1/2} - \tau_L^{1/2})^2 \tau^{-1} (N-1) \right) \right) (r + \alpha_G) \right. \\ & \left. - C_G (\tau_H^{1/2} - \tau_L^{1/2})^2 \tau^{-1} (N-1) (r + \alpha_G + \tau) e^{-\tau t} \right). \end{aligned} \quad (\text{A92})$$

Clearly, (A92) implies  $\frac{dPV_p(0, t)}{dt} \rightarrow 0$  when  $t \rightarrow \infty$ . Define

$$\hat{t} := -\frac{1}{\tau} \ln \left( \left( 1 + \frac{(1 - C_G)N\tau}{C_G(\tau_H^{1/2} - \tau_L^{1/2})^2(N-1)} \right) \frac{r + \alpha_G}{r + \alpha_G + \tau} \right). \quad (\text{A93})$$

Equation (A92) implies  $\frac{dPV_p(0, t)}{dt} > 0$  if and only if  $t > \hat{t}$ . It can be shown that  $\hat{t} > 0$  if and only if  $C_G > \hat{C}_G := \left( 1 + (1 - 1/N)(\tau_H^{1/2} - \tau_L^{1/2})^2(r + \alpha_G)^{-1} \right)^{-1}$ . This yields the following results:

- If  $C_G \leq \hat{C}_G$ , then  $\frac{dPV_p(0, t)}{dt} > 0$  for all  $t > 0$ .
- If  $C_G > \hat{C}_G$ , then  $\frac{dPV_p(0, t)}{dt} = 0$  for  $t = \hat{t}$ ,  $\frac{dPV_p(0, t)}{dt} > 0$  for  $t > \hat{t}$ , and  $\frac{dPV_p(0, t)}{dt} < 0$  for  $t < \hat{t}$ .

From (A88), it follows that

$$\frac{dPV_{-n}(0, t)}{dt} = \frac{(\tau_H^{1/2} - \tau_L^{1/2})^2 G_n(0) e^{-(r + \alpha_G)t}}{\tau(r + \alpha_G)(r + \alpha_D)} \left( (r + \alpha_G) - (r + \alpha_G + \tau) e^{-\tau t} \right). \quad (\text{A94})$$

(A94) implies that  $\frac{dPV_{-n}(0, t)}{dt} < 0$  if and only if  $t < \frac{1}{\tau} \ln \left( 1 + \frac{\tau}{r + \alpha_G} \right)$ .

#### A.8. The dynamics of $\hat{H}_n(t)$ and $\hat{H}_{-n}(t)$ under an alternative information structure

The dampening effect is related to an important conceptual point about how to model information. If we modeled the private information (21) and public information (24) as

$$dI_n(t) := \frac{G^*(t)}{\sigma_G \Omega^{1/2}} dt + \tau_n^{-1/2} dB_n(t), \quad n = 0, 1, \dots, N, \quad (\text{A95})$$

then our model would not generate a dampening effect. The dynamics of  $\hat{H}_n(t)$  and  $\hat{H}_{-n}(t)$  in equations (A34) and (A35) become

$$\begin{aligned} d\hat{H}_n(t) = & -(\alpha_G + \tau) \hat{H}_n(t) dt + (1 + \hat{a}) \left( \tau_H \hat{H}_n(t) + (N-1) \tau_L \hat{H}_{-n}(t) \right) dt \\ & + \hat{a} \tau_0^{-1/2} dB_0^n(t) + \tau_H^{-1/2} dB_n^n(t), \end{aligned} \quad (\text{A96})$$



$$d\hat{H}_{-n}(t) = -(\alpha_G + \tau)\hat{H}_{-n}(t)dt + (1 + \hat{a})\left(\tau_H\hat{H}_n(t) + (N-1)\tau_L\hat{H}_{-n}(t)\right)dt \\ + \hat{a}\tau_0^{-1/2}dB_0^n(t) + \frac{\tau_L}{N-1}\sum_{\substack{m=1 \\ m \neq n}}^N dB_m^n(t), \quad (\text{A97})$$

where  $\hat{a} := \tau_0/(\tau_H + (N-1)\tau_L)$ . The dynamics of  $\hat{H}_n$  and  $\hat{H}_{-n}$  then imply that traders only disagree about the value of the asset in the present but not about the dynamics of their future valuations. We have solved the equilibrium under this alternative setup. Most results remain qualitatively the same, except that there is no price dampening ( $C_G = 1$ ).

This alternative setup is not consistent with minimal rationality. In a continuous-time model, a trader can infer the diffusion variance with high accuracy by observing the information process over short periods of time. Therefore, it does not make economic sense for one trader to assume that another trader observes the diffusion variance of his signal incorrectly.

### A.9. Proof of Theorem 6

If the empirically correct precision of all traders is the same—such as  $\tau_H$  or  $\tau_L$  or  $(\tau_H + (N-1)\tau_L)/N$ —it can be shown that

$$d\hat{H}_n(t) - d\hat{H}_{-n}(t) = -(\alpha_G + \tau)(\hat{H}_n(t) - \hat{H}_{-n}(t))dt + dB_n(t) - \frac{1}{N-1}\sum_{\substack{m=1 \\ m \neq n}}^N dB_m(t). \quad (\text{A98})$$

This follows directly from the definitions of  $I_n(t)$ ,  $\hat{H}_n(t)$ , and  $\hat{H}_{-n}(t)$ . In the dynamics of  $d\hat{H}_n(t) = -(\alpha_G + \tau)\hat{H}_n(t)dt + dI_n(t)$ , we plug in the empirically correct beliefs about the dynamics of the information  $dI_n(t)$ , where the precision is the same for all information processes.

Equations (36) and (37) imply the target inventories and actual inventories follow the bivariate vector autoregression

$$\begin{pmatrix} dS_n^{TI}(t) \\ dS_n(t) \end{pmatrix} = \begin{pmatrix} -(\alpha_G + \tau) & 0 \\ \gamma_S & -\gamma_S \end{pmatrix} \begin{pmatrix} S_n^{TI}(t) \\ S_n(t) \end{pmatrix} dt + \begin{pmatrix} C_L \\ 0 \end{pmatrix} \left( dB_n(t) - \frac{1}{N-1}\sum_{\substack{m=1 \\ m \neq n}}^N dB_m(t) \right). \quad (\text{A99})$$

Simple calculations yield

$$S_n^{TI}(t) = C_L \int_{-\infty}^t e^{-(\alpha_G + \tau)(t-s)} \left( dB_n(s) - \frac{1}{N-1}\sum_{\substack{m=1 \\ m \neq n}}^N dB_m(s) \right), \quad (\text{A100})$$

$$S_n(t) = C_L \gamma_S \int_{-\infty}^t \frac{e^{-(\alpha_G + \tau)(t-s)} - e^{-\gamma_S(t-s)}}{\gamma_S - \alpha_G - \tau} \left( dB_n(s) - \frac{1}{N-1}\sum_{\substack{m=1 \\ m \neq n}}^N dB_m(s) \right). \quad (\text{A101})$$

From (A100) and (A101), simple calculations yield

$$\text{Corr}[S_n(t), S_n(t + \Delta t)] = \frac{(\alpha_G + \tau)e^{-\gamma_S \Delta t} - \gamma_S e^{-(\alpha_G + \tau)\Delta t}}{\alpha_G + \tau - \gamma_S}, \quad (\text{A102})$$

$$\text{Corr}[S_n(t), S_n^{TI}(t)] = \left( \frac{\gamma_S}{\gamma_S + \alpha_G + \tau} \right)^{1/2}. \quad (\text{A103})$$

### A.10. Auto-covariance of the Holding-Period return

Consider the case when traders are correct on average, so that the econometrician assigns precisions  $\check{\tau}_I = (\tau_H + (N-1)\tau_L)/N$  to each signal. Then the econometrician agrees with the traders about the total precision of the signals ( $\check{\tau} = \check{\tau}_0 + N\check{\tau}_I = \tau$ ). The error variance  $\check{\Omega}$  satisfies  $\check{\Omega}^{-1} = 2\alpha_G + \check{\tau}$  with  $\check{\tau}_0 = \check{\Omega}\sigma_G^2/\sigma_D^2$ . The signal  $\check{H}_n(t)$  for  $n = 0, 1, \dots, N$  is obtained by replacing  $\tau_0$ ,  $\tau$ ,  $\Omega$ , and  $\tau_n$  with  $\check{\tau}_0$ ,  $\check{\tau}$ ,  $\check{\Omega}$ , and  $\check{\tau}_I$  in (21), (24), and (26). Since  $\check{\tau}_0 = \tau_0$  and  $\check{\Omega} = \Omega$  hold in this case, the traders' statistics and econometrician's statistics coincide, yielding  $\check{H}_n(t) = H_n(t)$ .

The information flow and the econometrician's estimate of the growth rate  $\check{G}(t)$  are given as

$$dI_n(t) := \check{\tau}_I^{1/2} \frac{G^*(t)}{\sigma_G \check{\Omega}^{1/2}} dt + d\check{B}_n(t), \quad n = 1, \dots, N, \quad (\text{A104})$$

$$d\check{B}_n(t) = dB_n(t) + \left( \frac{\tau_n^{1/2}}{\sigma_G \Omega^{1/2}} - \frac{\check{\tau}_I^{1/2}}{\sigma_G \Omega^{1/2}} \right) G^*(t) dt, \quad (\text{A105})$$

$$\check{G}(t) := \mathbb{E}[G^*(t)] = \sigma_G \Omega^{1/2} \left( \tau_0^{1/2} H_0(t) + \sum_{n=1}^N \check{\tau}_I^{1/2} H_n(t) \right). \quad (\text{A106})$$

Using equations (35) and (28), we can write an equation for  $dP(t)$ , plug in  $dH_n(t)$  using equation (A13), and plug in the econometrician's beliefs about the dynamics of  $dH_n(t)$  from equation (A104) and the econometrician's estimate  $\check{G}(t)$  from equation (A106). The instantaneous return can be written in terms of  $H_0(t)$  and  $H_n(t)$  as

$$dP(t) + D(t)dt = rP(t)dt + c_0 H_0(t)dt + c_H \sum_{n=1}^N H_n(t)dt + d\check{B}_r(t). \quad (\text{A107})$$

The coefficients of  $H_0(t)$  and  $\sum_{n=1}^N H_n(t)$  are given as

$$c_0 = \frac{\sigma_G \Omega^{1/2}}{r + \alpha_D} \left( (1 - C_G) \tau_0^{1/2} - \frac{C_G N \check{\tau}_I^{1/2} \tau_0^{1/2}}{r + \alpha_G} (\check{\tau}_I^{1/2} - \tau_I^{1/2}) \right), \quad (\text{A108})$$

$$c_H = \frac{\sigma_G \Omega^{1/2}}{r + \alpha_D} \left( (1 - C_G) \tau_I^{1/2} + \frac{r + \alpha_G + C_G \tau_0}{r + \alpha_G} (\check{\tau}_I^{1/2} - \tau_I^{1/2}) \right), \quad (\text{A109})$$

where  $\tau_I^{1/2} := \frac{1}{N} (\tau_H^{1/2} + (N-1)\tau_L^{1/2})$ . It can be shown that the coefficient on  $\sum_{n=1}^N H_n(t)$  in this expression is always positive. Indeed, its first term with  $1 - C_G > 0$  results from the price-dampening effect of the Keynesian beauty contest, and its second term with  $\check{\tau}_I^{1/2} - \tau_I^{1/2} > 0$  results from Jensen's inequality. Thus, there will be momentum in price dynamics even when traders and the economist agree on the total precision of the information flow.

The uncertainty term  $d\check{B}_r(t)$  in equation (A107) is defined as

$$d\check{B}_r(t) := \frac{\sigma_G C_G \Omega^{1/2}}{(r + \alpha_D)(r + \alpha_G)} \left( \tau_0^{1/2} dB_0^*(t) + \tau_I^{1/2} N d\bar{B}^*(t) \right) + \frac{\sigma_D}{(r + \alpha_D)} dB_0^*(t). \quad (\text{A110})$$

The processes  $d\bar{B}^*(t)$  and  $dB_0^*(t)$ , defined as

$$d\bar{B}^*(t) := \check{\tau}_I^{1/2} (\sigma_G \Omega^{1/2})^{-1} (G^*(t) - \check{G}(t))dt + \frac{1}{N} \sum_{n=1}^N d\check{B}_n(t), \quad (\text{A111})$$

$$dB_0^*(t) := \tau_0^{1/2} (\sigma_G \Omega^{1/2})^{-1} (G^*(t) - \check{G}(t))dt + dB_0(t), \quad (\text{A112})$$

are Brownian motions under the empirically correct beliefs. Note that the variance of  $dB_0^*(t)$  is equal to one, but the variance of  $d\bar{B}^*(t)$  is equal to  $1/N$  per unit of time.

Let  $R(t, t + \Delta t)$  denote the cumulative undiscounted holding-period mark-to-market cash flow per share on a fully levered investment in the risky asset from time  $t$  to time  $t + \Delta t$ :

$$R(t, t + \Delta t) = \int_{u=t}^{t+\Delta t} (dP(u) + D(u)du - rP(u)du). \quad (\text{A113})$$

We compute the holding-period return  $R(t, t + \Delta t)$  as

$$R(t, t + \Delta t) = \beta_0(\Delta t) H_0(t) + \beta_H(\Delta t) \sum_{n=1}^N H_n(t) + \bar{B}(t, t + \Delta t), \quad (\text{A114})$$

where the coefficient of  $H_0(t)$  and the coefficient of  $\sum_{n=1}^N H_n(t)$  are given as

$$\begin{aligned} \beta_0(\Delta t) = & \frac{\sigma_G \Omega^{1/2} \tau_0^{1/2}}{r + \alpha_D} C_G \left( \frac{\tau_0 + N \check{\tau}_I^{1/2} \tau_I^{1/2}}{\tau \alpha_G} e^{-\alpha_G \Delta t} + \left( \frac{1}{\tau} + \frac{1}{r + \alpha_G} \right) N \check{\tau}_I^{1/2} (\check{\tau}_I^{1/2} - \tau_I^{1/2}) \frac{e^{-(\alpha_G + \tau) \Delta t}}{\alpha_G + \tau} \right) \\ & + \frac{\sigma_G \Omega^{1/2} \tau_0^{1/2} (1 - C_G) \alpha_G (r + \alpha_G + \tau) + r \left( \tau - C_G (\tau_0 + N \check{\tau}_I^{1/2} \tau_I^{1/2}) \right)}{(r + \alpha_D) (r + \alpha_G) (\alpha_G + \tau) \alpha_G}, \end{aligned} \quad (\text{A115})$$

$$\begin{aligned} \beta_H(\Delta t) = & \frac{\sigma_G \Omega^{1/2}}{r + \alpha_D} (1 - C_G) \tau_I^{1/2} \frac{1 - e^{-\alpha_G \Delta t}}{\alpha_G} + \frac{\sigma_G \Omega^{1/2}}{(r + \alpha_D) \alpha_G (r + \alpha_G)} (\check{\tau}_I^{1/2} - \tau_I^{1/2}) \\ & \cdot \left( (1 - e^{-\alpha_G \Delta t}) \left( r + \alpha_G - \frac{C_G r \tau_0}{\alpha_G + \tau} \right) + \frac{C_G (\alpha_G + r + \tau) \alpha_G \tau_0}{(\alpha_G + \tau) \tau} (e^{-\alpha_G \Delta t} - e^{-(\alpha_G + \tau) \Delta t}) \right), \end{aligned} \quad (\text{A116})$$

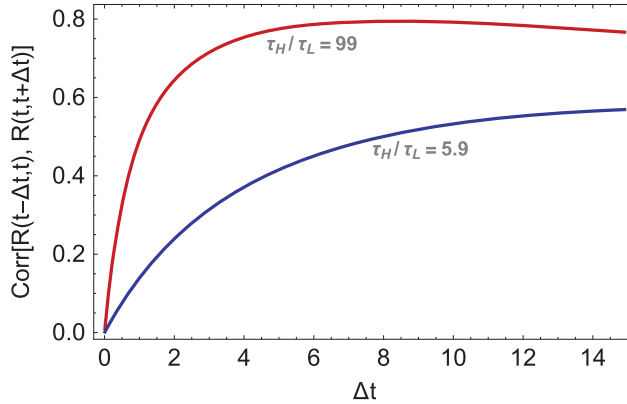


FIGURE 9

Auto-correlation of holding-period return.

$$\bar{B}(t, t + \Delta t) := \int_{s=t}^{t+\Delta t} \int_{u=s}^{t+\Delta t} [c_0, c_H] e^{K(u-s)} du dZ(s) + \int_{s=t}^{t+\Delta t} d\bar{B}_r(s), \quad (\text{A117})$$

where  $dZ(s) = [dB_0^*(s), Nd\bar{B}^*(s)]'$  and

$$K = \begin{pmatrix} -\alpha_G - N\check{\tau}_I & \tau_0^{1/2} \check{\tau}_I^{1/2} \\ N\tau_0^{1/2} \check{\tau}_I^{1/2} & -\alpha_G - \tau_0 \end{pmatrix}. \quad (\text{A118})$$

It can be shown that

$$\begin{aligned} \text{Cov}[R(t - \Delta t, t), R(t, t + \Delta t)] &= \mathbb{E}[R(t - \Delta t, t)R(t, t + \Delta t)] \\ &= \frac{\sigma_G^2 \Omega}{(r + \alpha_D)^2} \frac{(1 - e^{-\alpha_G \Delta t})^2}{\alpha_G^2} \frac{\tau - C_G(\tau_0 + N\check{\tau}_I^{1/2} \tau_I^{1/2})}{\tau} \left( 1 + \frac{\tau}{2\alpha_G} + \frac{C_G(\alpha_G - r)(\tau_0 + N\check{\tau}_I^{1/2} \tau_I^{1/2})}{2\alpha_G(r + \alpha_G)} \right) \\ &\quad - \frac{\sigma_G^2 \Omega N(\check{\tau}_I^{1/2} - \tau_I^{1/2}) C_G}{(r + \alpha_D)^2} \frac{(1 - e^{-(\alpha_G + \tau)\Delta t})^2}{(\alpha_G + \tau)^2} \frac{r + \alpha_G + \tau}{(r + \alpha_G)\tau} \left( \frac{C_G \tau_0(\check{\tau}_I^{1/2} - \tau_I^{1/2})}{2(r + \alpha_G)} \frac{\alpha_G + \tau - r}{\alpha_G + \tau} + \check{\tau}_I^{1/2} \right). \end{aligned} \quad (\text{A119})$$

The dampening effect ( $C_G < 1$ ) leads to momentum (positive autocorrelation) in returns. For the limiting case studied in section 4.4 with  $\tau_L = 0$ ,  $\tau_0 \rightarrow 0$ , and  $N \rightarrow \infty$ , the dampening effect is substantial; it can be shown that

$$\text{Cov}[R(t - \Delta t, t), R(t, t + \Delta t)] = \frac{(1 - e^{-\alpha_G \Delta t})^2}{\alpha_G^2} \left( 1 + \frac{\tau}{2\alpha_G} \right) - \frac{(1 - e^{-(\alpha_G + \tau)\Delta t})^2}{(\alpha_G + \tau)^2} > 0. \quad (\text{A120})$$

For general cases, from (A119), the auto-covariance of holding-period return tends to be positive when  $\tau$  is large relative to  $\tau_0$  and  $\alpha_G$ . Our extensive numerical analysis of equation (A119) shows that the auto-covariance is positive for a large range of parameter values. In addition, as illustrated in Figure 9, the auto-correlation of holding-period return tends to be larger with more disagreement. This implies that momentum is more pronounced in more liquid markets.<sup>15</sup>

## B. A COMPETITIVE MODEL OF TRADING

In this section, we consider a model that is similar to the smooth trading model with the only difference that traders are perfectly competitive. The competitive equilibrium is different from the equilibrium with imperfect competition. Traders adjust their inventories immediately; they do not smooth out their trading over time. The one-period model is discussed next, followed by the continuous-time model of perfect competition.

15. Numerical calculations in Figure 9 are based on the exogenous parameter values  $\alpha_D = 0.1$ ,  $\alpha_G = 0.02$ ,  $\sigma_D = 0.5$ ,  $\sigma_G = 0.1$ , and  $N = 100$ , with  $\tau = 8.9$  and  $\tau_0 = \Omega \sigma_G^2 / \sigma_D^2 = 0.0045$  in both cases;  $\tau_H = 4.46$  and  $\tau_L = 0.045$  for the case with higher disagreement level;  $\tau_H = 0.5$  and  $\tau_L = 0.085$  for the case with lower disagreement level.

### B.1. One-period model

The setting is almost identical to the setting of our model of imperfect competition. For clarity, we repeat analogous assumptions here.

A risky asset with random liquidation value  $v \sim N(0, 1/\tau_v)$  is traded for a safe numeraire asset. Each of  $N$  traders  $n = 1, \dots, N$  is endowed with  $S_n$  shares of a zero-net-supply risky asset, implying  $\sum_{n=1}^N S_n = 0$ . Traders observe signals about the normalized liquidation value  $\tau_v^{1/2} v$ . All traders observe a public signal  $i_0 := \tau_0^{1/2} (\tau_v^{1/2} v) + e_0$  with  $e_0 \sim N(0, 1)$ . Each trader  $n$  observes a private signal  $i_n := \tau_n^{1/2} (\tau_v^{1/2} v) + e_n$  with  $e_n \sim N(0, 1)$ . The asset payoff  $v$ , the public signal error  $e_0$ , and  $N$  private signal errors  $e_1, \dots, e_N$  are independently distributed.

Traders agree about the precision of the public signal  $\tau_0$  and agree to disagree about the precisions of private signals  $\tau_n$ . Each trader is “relatively overconfident”, believing his own signal has a high precision  $\tau_n = \tau_H$  and other traders’ signals have low precision  $\tau_m = \tau_L$  for  $m \neq n$ , with  $\tau_H > \tau_L \geq 0$ .

Each trader submits a demand schedule  $X_n(p) := X_n(i_0, i_n, S_n, p)$  to a single-price auction. An auctioneer calculates the market-clearing price  $p := p[X_1, \dots, X_N]$ .

Trader  $n$ ’s terminal wealth is

$$W_n := v(S_n + X_n(p)) - pX_n(p). \quad (\text{B1})$$

The difference from equation (1) is that each trader  $n$  assumes that the price  $p$  does not depend on the quantities he trades. Each trader maximizes the same expected exponential utility function of wealth  $\mathbb{E}^n[-e^{-AW_n}]$  using his own beliefs about  $\tau_H$  and  $\tau_L$  to calculate the expectation.

Trader  $n$  maximizes his expected utility, or equivalently he maximizes  $\mathbb{E}^n[W_n] - \frac{1}{2}A \text{Var}^n[W_n]$ . He chooses the quantity to trade  $x_n$  that solves the maximization problem

$$\max_{x_n} \left[ \frac{\tau_v^{1/2}}{\tau} \left( \tau_0^{1/2} i_0 + \tau_H^{1/2} i_n + (N-1)\tau_L^{1/2} i_{-n} \right) (S_n + x_n) - p x_n - \frac{A}{2\tau} (S_n + x_n)^2 \right]. \quad (\text{B2})$$

The first-order condition with respect to  $x_n$  yields

$$x_n^* = \frac{1}{A} \left( \tau_v^{1/2} \left( \tau_0^{1/2} i_0 + \tau_H^{1/2} i_n + (N-1)\tau_L^{1/2} i_{-n} \right) - p \tau \right) - S_n. \quad (\text{B3})$$

The market-clearing condition  $\sum_{n=1}^N x_n^* = 0$  implies

$$p^* = \frac{1}{N} \sum_{n=1}^N \mathbb{E}^n[v] = \frac{\tau_v^{1/2}}{\tau} \left( \tau_0^{1/2} i_0 + \frac{\tau_H^{1/2} + (N-1)\tau_L^{1/2}}{N} \sum_{n=1}^N i_n \right). \quad (\text{B4})$$

As in equation (13) for the case of imperfect competition, the equilibrium price  $p^*$  is equal to the average of traders’ valuations. Substituting (B4) into (B3) yields

$$x_n^* = \frac{1}{A} \left( 1 - \frac{1}{N} \right) \tau_v^{1/2} (\tau_H^{1/2} - \tau_L^{1/2}) (i_n - i_{-n}) - S_n. \quad (\text{B5})$$

Thus, each trader trades on the difference between his signal  $i_n$  and the average of all  $N$  signals and also trades out of his current inventory  $S_n$ .

Define the target inventory as

$$S_n^T = \frac{1}{A} \left( 1 - \frac{1}{N} \right) \tau_v^{1/2} (\tau_H^{1/2} - \tau_L^{1/2}) (i_n - i_{-n}). \quad (\text{B6})$$

Equation (B5) is similar to equation (11), except for the endogenous constant  $\delta = 1$ , implying each trader trades to his “target inventory”  $S_n^T$  immediately in the competitive model. Note that target inventories are identical to target inventories (14) in the model with imperfect competition.

To summarize, in the model of perfect competition, both the target inventories and equilibrium price are the same as in our smooth trading model with imperfect competition. The key difference is that in the model of perfect competition, traders trade to their target inventories fully ( $\delta = 1$ ) instead of partially ( $0 < \delta < 1$ ).

### B.2. A continuous-time model of perfect competition

For the competitive equilibrium, we use the same notation and information structure as in our smooth trading model of imperfect competition. The only difference from our smooth trading model is that traders do not take into account price impact when solving for their optimal demand. For all dates  $t > -\infty$ , the optimal strategies  $S_n^*$  and  $C_n^*$  solve trader  $n$ ’s

maximization problem

$$\max_{[C_n, S_n]} \mathbb{E}_t^n \left[ \int_{s=t}^{\infty} e^{-\rho(s-t)} U(c_n(s)) ds \right], \quad (\text{B7})$$

where wealth  $W_n(t)$  follows the process

$$dW_n(t) = rW_n(t)dt + S_n(t)(dP(t) + D(t)dt - rP(t)dt) - c_n(t)dt. \quad (\text{B8})$$

These two equations are similar to equations (22) and (23), but there are several differences. First, traders take prices in equation (B8) as given. Second, in the model with perfect competition traders can costlessly transfer funds from their money account to stock account. It is therefore sufficient to keep track only of aggregate wealth dynamics, rather than to keep track of a money account and a stock account separately.

Traders use the history of the dividend process, the history of their own private signals, and the average of all signals, as inferred from prices, to obtain their estimates of the growth rate. The inference problem is identical to the one in the smooth trading model.

To solve the equilibrium, we conjecture that price is a linear function of  $D(t)$  and  $\bar{G}(t)$ , specifically,

$$P(t) = \frac{D(t)}{r + \alpha_D} + C_G \frac{\bar{G}(t)}{(r + \alpha_D)(r + \alpha_G)}. \quad (\text{B9})$$

It can be shown that

$$\begin{aligned} dP(t) = & -\frac{1}{r + \alpha_D} \left( \alpha_D D(t) - \sigma_G \Omega^{1/2} \left( \tau_H^{1/2} \hat{H}_n(t) + (N-1) \tau_L^{1/2} \hat{H}_{-n}(t) \right) \right) dt \\ & + \frac{C_G \sigma_G \Omega^{1/2} (\tau_H^{1/2} + (N-1) \tau_L^{1/2})}{N(r + \alpha_D)(r + \alpha_G)} ((a_1 + (N-1)a_4) \hat{H}_n(t) + (a_3 + (N-1)a_2) \hat{H}_{-n}(t)) dt \\ & + \frac{1}{r + \alpha_D} ((G^*(t) - G_n(t)) dt + \sigma_D dB_D) \\ & + \frac{C_G \sigma_G \Omega^{1/2} (\tau_H^{1/2} + (N-1) \tau_L^{1/2})}{N(r + \alpha_D)(r + \alpha_G)} \left( N \hat{a} dB_0^n(t) + dB_n^n(t) + \sum_{\substack{m=1 \\ m \neq n}}^N dB_m^n(t) \right), \end{aligned} \quad (\text{B10})$$

where the constants  $a_1$ ,  $a_2$ ,  $a_3$ , and  $a_4$  are defined as

$$\begin{aligned} a_1 &:= -\alpha_G - \tau + \tau_H^{1/2} (\tau_H^{1/2} + \hat{a} \tau_0^{1/2}), \\ a_2 &:= -\alpha_G - \tau + (N-1) \tau_L^{1/2} (\tau_L^{1/2} + \hat{a} \tau_0^{1/2}), \\ a_3 &:= (\tau_H^{1/2} + \hat{a} \tau_0^{1/2})(N-1) \tau_L^{1/2}, \\ a_4 &:= (\tau_L^{1/2} + \hat{a} \tau_0^{1/2}) \tau_H^{1/2}. \end{aligned} \quad (\text{B11})$$

We conjecture and verify that the value function  $V(W_n, \hat{H}_n, \hat{H}_{-n})$  has the specific quadratic exponential form

$$V(W_n, \hat{H}_n, \hat{H}_{-n}) = -\exp \left( \psi_0 + \psi_W W_n + \frac{1}{2} \psi_{nn} \hat{H}_n^2 + \frac{1}{2} \psi_{xx} \hat{H}_{-n}^2 + \psi_{nx} \hat{H}_n \hat{H}_{-n} \right). \quad (\text{B12})$$

As in our smooth trading model, the five constants  $\psi_0$ ,  $\psi_W$ ,  $\psi_{nn}$ ,  $\psi_{xx}$ , and  $\psi_{nx}$  have values consistent with a steady-state equilibrium. The terms  $\psi_{nn}$ ,  $\psi_{xx}$ , and  $\psi_{nx}$  capture the value of future trading opportunities based on current public and private information. The value of trading on innovations to future information is built into the constant term  $\psi_0$ . Equation (B12) is similar to equation (A37), except that it has a simpler form because the five terms  $M_n$ ,  $S_n^2$ ,  $S_n D$ ,  $S_n \hat{H}_n$ , and  $S_n \hat{H}_{-n}$  are effectively replaced by one term,  $W_n$ .

The Hamilton–Jacobi–Bellman (HJB) equation corresponding to the conjectured value function  $V(W_n, \hat{H}_n, \hat{H}_{-n})$  in equation (B12) is

$$\begin{aligned}
0 = & \min_{c_n, S_n} - \frac{e^{-Ac_n}}{V} - \rho + \psi_W \left( rW_n + S_n D(t) - c_n - rP(t)S_n(t) - \frac{\alpha_D}{r + \alpha_D} D(t)S_n \right. \\
& + \frac{\sigma_G \Omega^{1/2}}{r + \alpha_D} \left( \tau_H^{1/2} \hat{H}_n(t) + (N-1)\tau_L^{1/2} \hat{H}_{-n}(t) \right) S_n \\
& + \frac{C_G \sigma_G \Omega^{1/2} (\tau_H^{1/2} + (N-1)\tau_L^{1/2})}{N(r + \alpha_D)(r + \alpha_G)} \left( (a_1 + (N-1)a_4) \hat{H}_n(t) + (a_3 + (N-1)a_2) \hat{H}_{-n}(t) \right) S_n \Big) \\
& + \left( \psi_{nn} \hat{H}_n(t) + \psi_{nx} \hat{H}_{-n}(t) \right) \left( -(\alpha_G + \tau) \hat{H}_n + (\tau_H^{1/2} + \hat{a} \tau_0^{1/2}) (\tau_H^{1/2} \hat{H}_n + (N-1)\tau_L^{1/2} \hat{H}_{-n}) \right) \\
& + \left( \psi_{xx} \hat{H}_{-n}(t) + \psi_{nx} \hat{H}_n(t) \right) \left( -(\alpha_G + \tau) \hat{H}_{-n} + (\tau_L^{1/2} + \hat{a} \tau_0^{1/2}) (\tau_H^{1/2} \hat{H}_n + (N-1)\tau_L^{1/2} \hat{H}_{-n}) \right) \\
& + \frac{1}{2} \psi_W^2 S_n^2 \left( \frac{C_G^2 \sigma_G^2 \Omega (N\hat{a}^2 + 1) (\tau_H^{1/2} + (N-1)\tau_L^{1/2})^2}{N(r + \alpha_D)^2 (r + \alpha_G)^2} + \frac{\sigma_D^2}{(r + \alpha_D)^2} + \frac{2C_G \sigma_G \sigma_D \Omega^{1/2} \tau_0^{1/2}}{(r + \alpha_D)^2 (r + \alpha_G)} \right) \\
& + \frac{1}{2} \left( (\psi_{nn} \hat{H}_n(t) + \psi_{nx} \hat{H}_{-n}(t))^2 + \psi_{nn} \right) (1 + \hat{a}^2) \\
& + \frac{1}{2} \left( (\psi_{xx} \hat{H}_{-n}(t) + \psi_{nx} \hat{H}_n(t))^2 + \psi_{xx} \right) \left( \frac{1}{N-1} + \hat{a}^2 \right) \\
& + \psi_W S_n \left( (\psi_{nn} + \psi_{nx}) \hat{H}_n(t) + (\psi_{xx} + \psi_{nx}) \hat{H}_{-n}(t) \right) \\
& \cdot \left( \frac{C_G \sigma_G \Omega^{1/2}}{N(r + \alpha_D)(r + \alpha_G)} (\tau_H^{1/2} + (N-1)\tau_L^{1/2}) (N\hat{a}^2 + 1) + \frac{\sigma_D \hat{a}}{r + \alpha_D} \right) \\
& + \left( (\psi_{nn} \hat{H}_n(t) + \psi_{nx} \hat{H}_{-n}(t)) (\psi_{xx} \hat{H}_{-n}(t) + \psi_{nx} \hat{H}_n(t)) + \psi_{nx} \right) \hat{a}^2.
\end{aligned} \tag{B13}$$

As in the smooth trading model, the solution for optimal consumption is

$$c_n^*(t) = -\frac{1}{A} \log \left( \frac{\psi_W V(t)}{A} \right). \tag{B14}$$

Plugging optimal consumption and  $P(t)$  from equation (B9) into the HJB equation yields a quadratic function of  $S_n$ . The second-order condition is always satisfied because the coefficient on the  $S_n^2$ -term is positive. It can be shown that the optimal trading strategy is a linear function of the state variables  $\hat{H}_n(t)$  and  $\hat{H}_{-n}(t)$ ,

$$\begin{aligned}
S_n^*(t) = & C \left( C_G \sigma_G \Omega^{1/2} (\tau_H^{1/2} + (N-1)\tau_L^{1/2}) \left( (r - a_1 - (N-1)a_4) \hat{H}_n(t) \right. \right. \\
& \left. \left. + ((N-1)(r - a_2) - a_3) \hat{H}_{-n}(t) \right) \right. \\
& - \sigma_G \Omega^{1/2} (r + \alpha_G) N \left( \tau_H^{1/2} \hat{H}_n(t) + (N-1)\tau_L^{1/2} \hat{H}_{-n}(t) \right) \\
& - \left( (\psi_{nn} + \psi_{nx}) \hat{H}_n(t) + (\psi_{xx} + \psi_{nx}) \hat{H}_{-n}(t) \right) \\
& \cdot \left( C_G \sigma_G \Omega^{1/2} (\tau_H^{1/2} + (N-1)\tau_L^{1/2}) (N\hat{a}^2 + 1) + \sigma_D \hat{a} N (r + \alpha_G) \right) \Big),
\end{aligned} \tag{B15}$$

where

$$C := \frac{(r + \alpha_D)(r + \alpha_G) / \psi_W}{C_G^2 \sigma_G^2 \Omega (\tau_H^{1/2} + (N-1)\tau_L^{1/2})^2 (N\hat{a}^2 + 1) + N\sigma_D^2 (r + \alpha_G)^2 + 2N(r + \alpha_G) \sigma_D C_G \sigma_G \Omega^{1/2} \tau_0^{1/2}}. \tag{B16}$$

Market clearing,  $\sum_{n=1}^N S_n^*(t) = 0$ , implies

$$C_G = \frac{N(r + \alpha_G) \left( \sigma_G \Omega^{1/2} + \sigma_D \hat{a} (\psi_{nn} + \psi_{xx} + 2\psi_{nx}) \left( \tau_H^{1/2} + (N-1)\tau_L^{1/2} \right)^{-1} \right)}{\sigma_G \Omega^{1/2} \left( N(r + \alpha_G) + (N-1)(\tau_H^{1/2} - \tau_L^{1/2})^2 - (1 + N\hat{a}^2)(\psi_{nn} + \psi_{xx} + 2\psi_{nx}) \right)}. \tag{B17}$$

Combining equations (B15) and (B17) yields

$$S_n^*(t) = C_L (\hat{H}_n - \hat{H}_{-n}), \tag{B18}$$

where the constant  $C_L$  is defined as

$$C_L := C \left( \sigma_G \Omega^{1/2} \left( C_G (\tau_H^{1/2} + (N-1)\tau_L^{1/2})(r - a_1 - (N-1)a_4) - N\tau_H^{1/2}(r + \alpha_G) \right) \right. \\ \left. - (\psi_{nm} + \psi_{nx}) \left( C_G \sigma_G \Omega^{1/2} (\tau_H^{1/2} + (N-1)\tau_L^{1/2})(1 + N\hat{a}^2) + \sigma_D \hat{a} N(r + \alpha_G) \right) \right). \quad (\text{B19})$$

Plugging (B14) and (B18) back into the Bellman equation and setting the constant term and the coefficients of  $W_n$ ,  $\hat{H}_n^2$ ,  $\hat{H}_{-n}^2$ , and  $\hat{H}_n \hat{H}_{-n}$  to be zero yields five equations, which can be solved for the five unknown parameters  $\psi_0$ ,  $\psi_W$ ,  $\psi_{nm}$ ,  $\psi_{nx}$ , and  $\psi_{xx}$ .

Equating the constant term and the coefficient of  $W_n$  to zero yields

$$\psi_W = -rA, \quad (\text{B20})$$

$$\psi_0 = 1 - \log(r) + \frac{1}{r} \left( -\rho + \frac{1}{2}(1 + \hat{a}^2)\psi_{nm} + \frac{1}{2} \left( \frac{1}{N-1} + \hat{a}^2 \right) \psi_{xx} + \hat{a}^2 \psi_{nx} \right). \quad (\text{B21})$$

Equating the coefficients of  $\hat{H}_n^2$ ,  $\hat{H}_{-n}^2$ , and  $\hat{H}_n \hat{H}_{-n}$  to zero results in three polynomial equations in the three unknowns  $\psi_{nm}$ ,  $\psi_{xx}$ , and  $\psi_{nx}$ . Defining  $c_1$ ,  $c_2$ ,  $c_3$ , and  $c_4$  by

$$c_1 := \frac{C_G^2 \sigma_G^2 \Omega (N\hat{a}^2 + 1) (\tau_H^{1/2} + (N-1)\tau_L^{1/2})^2}{N(r + \alpha_D)^2 (r + \alpha_G)^2} + \frac{\sigma_D^2}{(r + \alpha_D)^2} + \frac{2C_G \sigma_G \sigma_D \Omega^{1/2} \tau_0^{1/2}}{(r + \alpha_D)^2 (r + \alpha_G)}, \quad (\text{B22})$$

$$c_2 := \frac{C_G \sigma_G \Omega^{1/2}}{N(r + \alpha_D)(r + \alpha_G)} (\tau_H^{1/2} + (N-1)\tau_L^{1/2})(N\hat{a}^2 + 1) + \frac{\sigma_D \hat{a}}{r + \alpha_D}, \quad (\text{B23})$$

$$c_3 := \frac{rA\sigma_G\Omega^{1/2}C_L}{r + \alpha_D} \left( \frac{C_G(\tau_H^{1/2} + (N-1)\tau_L^{1/2})(r - a_1 - (N-1)a_4)}{N(r + \alpha_G)} - \tau_H^{1/2} \right), \quad (\text{B24})$$

$$c_4 := \frac{rA\sigma_G\Omega^{1/2}C_L}{r + \alpha_D} \left( \frac{C_G(\tau_H^{1/2} + (N-1)\tau_L^{1/2})(r - a_2 - \frac{a_3}{N-1})}{N(r + \alpha_G)} - \tau_L^{1/2} \right), \quad (\text{B25})$$

these three equations in three unknowns can be written as follows:

$$0 = -\frac{r}{2}\psi_{nm} + a_1\psi_{nm} + a_4\psi_{nx} - rAC_Lc_2(\psi_{nm} + \psi_{nx}) + \frac{1}{2}(1 + \hat{a}^2)\psi_{nm}^2 \\ + \frac{1}{2} \left( \frac{1}{N-1} + \hat{a}^2 \right) \psi_{nx}^2 + \hat{a}^2 \psi_{nm}\psi_{nx} + c_3 + \frac{1}{2}r^2A^2c_1C_L^2, \quad (\text{B26})$$

$$0 = -\frac{r}{2}\psi_{xx} + a_2\psi_{xx} + a_3\psi_{nx} + rAC_Lc_2(\psi_{xx} + \psi_{nx}) + \frac{1 + \hat{a}^2}{2}\psi_{nx}^2 \\ + \frac{1}{2} \left( \frac{1}{N-1} + \hat{a}^2 \right) \psi_{xx}^2 + \hat{a}^2 \psi_{xx}\psi_{nx} - (N-1)c_4 + \frac{1}{2}r^2A^2c_1C_L^2, \quad (\text{B27})$$

$$0 = -r\psi_{nx} + (a_1 + a_2)\psi_{nx} + a_3\psi_{nm} + a_4\psi_{xx} + rAC_Lc_2(\psi_{nm} - \psi_{xx}) \\ + (1 + \hat{a}^2)\psi_{nm}\psi_{nx} + \left( \frac{1}{N-1} + \hat{a}^2 \right) \psi_{xx}\psi_{nx} + \hat{a}^2 (\psi_{nm}\psi_{xx} + \psi_{nx}^2) \\ + (N-1)c_4 - c_3 - r^2A^2c_1C_L^2. \quad (\text{B28})$$

To summarize, optimal consumption is defined in (B14), the optimal trading strategy is defined in (B18), and the endogenous coefficient  $C_L$  is defined in (B19). The equilibrium price is defined in (B9), and the endogenous coefficient  $C_G$  is defined in (B17). The parameters  $\psi_W$  and  $\psi_0$  are presented in (B20) and (B21). The parameters  $\psi_{nm}$ ,  $\psi_{nx}$ , and  $\psi_{xx}$  are obtained from numerical solution of the system of the three equations (B26)–(B28). These results are stated in Theorem 5.

Information has no value if there is no trading, so that  $\psi_{nm} = \psi_{nx} = \psi_{xx} = 0$  solves the three equations (B26)–(B28) when there is no liquidity. This implies

$$c_3 + \frac{1}{2}r^2A^2c_1C_L^2 = 0, \quad -(N-1)c_4 + \frac{1}{2}r^2A^2c_1C_L^2 = 0, \quad (N-1)c_4 - c_3 - r^2A^2c_1C_L^2 = 0. \quad (\text{B29})$$

These equations imply that liquidity vanishes when  $\tau_H = \tau_L$  and  $C_G = 1$ . This is different from our smooth trading model of disagreement with imperfect competition, in which market liquidity vanishes when  $\tau_H^{1/2}/\tau_L^{1/2} = 2 + 2/(N-2)$  and  $C_G = 1$ .



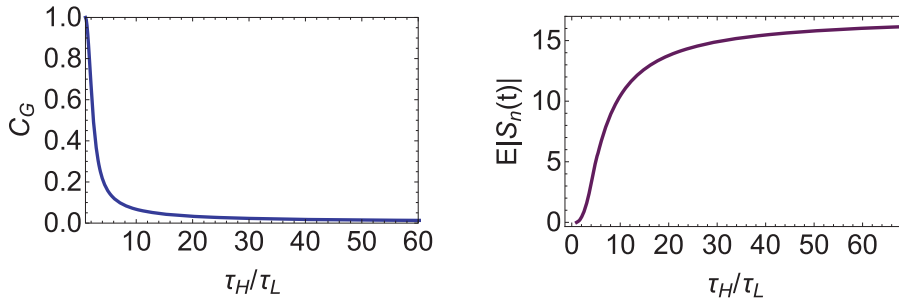


FIGURE 10

Coefficients  $C_G$  and  $\mathbb{E}[|S_n(t)|]$  against  $\tau_H/\tau_L$  while fixing  $\tau = 7.4$ .

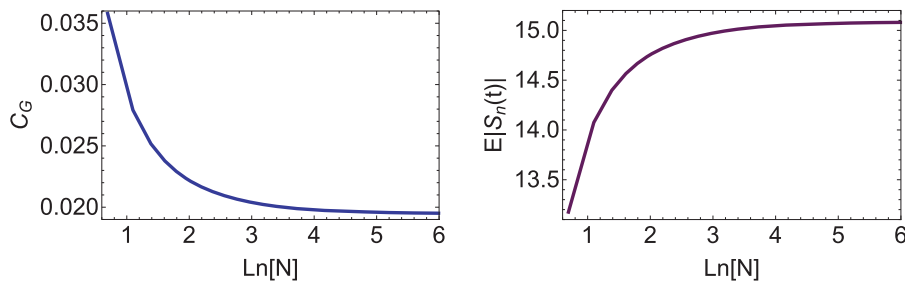


FIGURE 11

Coefficients  $C_G$  and  $\mathbb{E}[|S_n(t)|]$  against  $\ln(N)$  while fixing  $\tau = 1.4$  and  $\tau_L = 0$ .

Figure 10 shows the effect of changes in the degree of overconfidence  $\tau_H/\tau_L$  on the endogenous parameters  $C_G$  and  $\mathbb{E}[|S_n(t)|]$ . To compare the results with our smooth trading model, we use the same exogenous parameter values as in Figure 1 and panel (a) of Figure 8. The horizontal axis shows the ratio  $\tau_H/\tau_L$ . As this ratio increases,  $\tau_H$  is increasing and  $\tau_L$  is decreasing so that the total precision  $\tau$  is fixed (and other exogenous parameters are also fixed). Higher values of the ratio  $\tau_H/\tau_L$  correspond to higher degrees of overconfidence. As disagreement  $\tau_H/\tau_L$  increases, the left panel shows that the parameter  $C_G$  declines monotonically, while the right panel shows that the expected size of inventories  $\mathbb{E}[|S_n(t)|]$  increases monotonically.

For finite  $N$ , Figure 11 shows the effect of changes in the number of traders  $N$  on  $C_G$  and  $\mathbb{E}[|S_n(t)|]$ , using the same exogenous parameter values as in Figure 2 and Figure 8b. As  $N$  increases, the left panel shows that  $C_G$  decreases monotonically towards a constant asymptote, and the right panel shows that  $\mathbb{E}[|S_n(t)|]$  increases monotonically towards a constant asymptote. When  $N$  is large, our numerical results show that our smooth trading model of imperfect competition converges to the equilibrium of the competitive model.

As in the smooth trading model, we find a closed-form solution when we set  $\tau_L = 0$ , and then we evaluate the solution in the limit as  $N \rightarrow \infty$  and  $\hat{a} \rightarrow 0$ . We conjecture and verify that  $\psi_{nn} = \bar{\psi}_{nn}$ ,  $\psi_{nx} = \bar{\psi}_{nx}$ , and  $\psi_{xx} = \bar{\psi}_{xx}$ , where  $\bar{\psi}_{nn}$ ,  $\bar{\psi}_{nx}$ , and  $\bar{\psi}_{xx}$  are constants that do not depend on  $N$ .

Solving the system of equations (B26)–(B28) yields

$$\bar{\psi}_{nn} = \frac{1}{2} \left( r + 2(\alpha_G + \tau - \tau_H) - \left( (r + 2(\alpha_G + \tau - \tau_H))^2 + \frac{4\Omega\sigma_G^2\tau_H}{\sigma_D^2} \right)^{1/2} \right), \quad (\text{B30})$$

$$\bar{\psi}_{nx} = \frac{\Omega\sigma_G^2\tau_H/\sigma_D^2}{r + 2(\alpha_G + \tau) - \tau_H - \bar{\psi}_{nn}}, \quad (\text{B31})$$

$$\bar{\psi}_{xx} = \frac{1}{r + 2\alpha_G + 2\tau} \left( \bar{\psi}_{nx}^2 - \frac{\Omega\sigma_G^2\tau_H}{\sigma_D^2} \right). \quad (\text{B32})$$

Equations (B17) and (B19) imply

$$C_G \rightarrow \frac{r + \alpha_G}{r + \alpha_G + \tau} < 1, \quad C_L = \frac{\Omega^{1/2} \sigma_G \tau_H^{1/2} (r + \alpha_D)}{A r \sigma_D^2}. \quad (\text{B33})$$

These results are exactly the same as the limiting case when  $N \rightarrow \infty$  and  $\hat{a} \rightarrow 0$  in the smooth trading model. This confirms that our smooth trading model of imperfect competition converges to the competitive model when  $N \rightarrow \infty$ .

*Acknowledgments.* We thank Bradyn Breon-Drish, Vincent Fardeau, Thierry Foucault, Jungsuk Han, Jeongmin Lee, Wei Li, Hong Liu, Mark Loewenstein, Natalie Popovic, Ioanid Roşu, Marzena Rostek, Dimitri Vayanos, Bin Wei, Liyan Yang, and three anonymous referees for their comments.

### Supplementary Data

Supplementary data are available at *Review of Economic Studies* online.

### REFERENCES

- ALMGREN, R. and CHRISS, N. (2000), "Optimal Execution of Portfolio Transactions", *Journal of Risk*, **3**, 5–39.
- BACK, K., CAO, H. C. and WILLARD, G. A. (2000), "Imperfect Competition Among Informed Traders", *Journal of Finance*, **55**, 2117–2155.
- BANERJEE, S. and KREMER, I. (2010), "Disagreement and Learning: Dynamic Patterns of Trade", *Journal of Finance*, **65**, 1269–1302.
- BLACK, F. (1971), "Towards a Fully Automated Exchange, Part I", *Financial Analysts Journal*, **27**, 29–34.
- BLACK, F. (1982), "The Trouble with Econometric Models", *Financial Analysts Journal*, **38**, 29–37.
- BLACK, F. (1986), "Noise", *Journal of Finance*, **41**, 529–543.
- BLACK, F. (1995), "Equilibrium Exchanges", *Financial Analysts Journal*, **51**, 23–29.
- BRUNNERMEIER, M. K. and PEDERSEN, L. H. (2005), "Predatory Trading", *Journal of Finance*, **60**, 1825–1863.
- BUDISH, E., CRAMTON, P. and SHIM, J. (2015), "The High-frequency Trading Arms Race: Frequent Batch Auctions as a Market Design Response", *Quarterly Journal of Economics*, **130**, 1547–1621.
- CALDENTEY, R. and STACCHETTI, E. (2010), "Insider Trading with a Random Deadline", *Econometrica*, **78**, 245–283.
- CARLIN, B., LOBO, M. and VISWANATHAN, S. (2007), "Episodic Liquidity Crises: Cooperative and Predatory Trading", *Journal of Finance*, **62**, 2235–2274.
- CHAKRABARTY, B., MOULTON, P. C. and TRZCINKA, C. (2013), "Institutional Holding Periods" (Technical report, Cornell University, The Center for Real Estate and Finance Working Paper Series, 2013-007), 1–44.
- CHAN, L. and LAKONISHOK, J. (1995), "The Behavior of Stock Prices Around Institutional Trades", *Journal of Finance*, **50**, 1147–1174.
- CHAU, M. and VAYANOS, D. (2008), "Strong-form Efficiency with Monopolistic Insiders", *Review of Financial Studies*, **21**, 2275–2306.
- CREMERS, M. and PAREEK, A. (2015), "Short-term Trading and Stock Return Anomalies: Momentum, Reversal, and Share Issuance", *Review of Finance*, **19**, 1649–1701.
- DU, S. and ZHU, H. (2017), "What is the Optimal Trading Frequency in Financial Markets?", *Review of Economic Studies* (forthcoming).
- DUFFIE, D. (2010), "Presidential Address: Asset Price Dynamics with Slow-moving Capital", *Journal of Finance*, **65**, 1237–1267.
- DUFOUR, A. and ENGLE, R. (2000), "Time and the Price Impact of a Trade", *Journal of Finance*, **55**, 2469–2498.
- DUGAST, J. and FOUCAULT, T. (2014), "False News, Informational Efficiency, and Price Reversals" (Technical report, Banque de France, Working Paper No. 513), <https://ideas.repec.org/p/bfr/banfra/513.html>.
- EYSTER, E., RABIN, M. and VAYANOS, D. (2015), "Financial Markets Where Traders Neglect the Informational Content of Prices" (Technical report, National Bureau of Economic Research, Working Paper No. 21224), <http://www.nber.org/papers/w21224>.
- FOSTER, D. F. and VISWANATHAN, S. (1996), "Strategic Trading when Agents Forecast the Forecasts of Others", *Journal of Finance*, **51**, 1437–1478.
- FOSTER, D. F. and VISWANATHAN, S. (1994), "Strategic Trading with Asymmetrically Informed Traders and Long-lived Information", *Journal of Financial and Quantitative Analysis*, **29**, 499–518.
- GARLEANU, N. and PEDERSEN, L. H. (2013), "Dynamic Trading with Predictable Returns and Transaction Costs", *Journal of Finance*, **68**, 2309–2340.
- GARLEANU, N. and L. H. PEDERSEN (2016), "Dynamic Portfolio Choice with Frictions", *Journal of Economic Theory*, **165**, 487–516.
- GRINOLD, R. C. and KAHN, R. N. (1994), *Active Portfolio Management: Quantitative Theory and Applications* (New York: Irwin Professional Publishing).
- HAN, J. and KYLE, A. S. (2017), "Speculative Equilibrium with Differences in Higher Order Beliefs", *Management Science*, (accepted for publication), Swedish House of Finance Research Paper No. 14–01, <https://ssrn.com/abstract=2400136> or <http://dx.doi.org/10.2139/ssrn.2400136>.

- HARSANYI, J. C. (1976), "Games with Incomplete Information Played by 'Bayesian' Players. Part I: The Basic Model", *Management Science*, **14**, 159–182.
- HASBROUCK, J. and SOFIANOS, G. (1993), "The Trades of Market Makers: An Empirical Analysis of NYSE Specialists", *Journal of Finance*, **48**, 1565–1593.
- HE, H. and WANG, J. (1995), "Differential Information and Dynamic Behavior of Stock Trading Volume", *Review of Financial Studies*, **8**, 919–972.
- HENDERSHOTT, T. and MENKVELD, A. J. (2014), "Price Pressures", *Journal of Financial Economics*, **114**, 405–423.
- HOLDEN, C. W. and SUBRAHMANYAM, A. (1992), "Long-lived Private Information and Imperfect Competition", *Journal of Finance*, **47**, 247–270.
- HOLTHAUSEN, R. W., LEFTWICH, R. W. and MAYERS, D. (1990), "Large-block Transactions, the Speed of Response, and Temporary and Permanent Stock-price Effects", *Journal of Finance*, **26**, 71–95.
- KEIM, D. B. and MADHAVAN, A. (1997), "Transaction Costs and Investment Style: An Inter-exchange Analysis of Institutional Equity Trades", *Journal of Financial Economics*, **46**, 265–292.
- KEYNES, J. M. (1936), *The General Theory of Employment, Interest and Money* (Palgrave Macmillan, United Kingdom).
- KYLE, A. S. (1985), "Continuous Auctions and Insider Trading", *Econometrica*, **53**, 1315–1335.
- KYLE, A. S. (1989), "Informed Speculation with Imperfect Competition", *Review of Economic Studies*, **56**, 317–355.
- KYLE, A. S. and LIN, T. (2001), "Continuous Speculation with Overconfident Competitors" (Technical report, Duke University).
- KYLE, A. S. and OBIZHAEVA, A. A. (2016), "Large Bets and Stock Market Crashes" (Technical report), <https://ssrn.com/abstract=2023776> or <http://dx.doi.org/10.2139/ssrn.2023776>.
- KYLE, A. S., OBIZHAEVA, A. A. and WANG, Y. (2016), "Beliefs Aggregation and Return Predictability" (Technical report, Centre for Economic and Financial Research at New Economic School, Working Paper No. 231), <http://econpapers.repec.org/RePEc:cfr:cefir:w0231> or <http://www.cefir.ru/papers/WP231.pdf>.
- LEE, C. and SWAMINATHAN, B. (2000), "Transactions Costs and Holding Periods for Common Stocks", *Journal of Finance*, **55**, 2017–2069.
- LONGSTAFF, F. (2001), "Optimal Portfolio Choice and the Valuation of Illiquid Securities", *Review of Financial Studies*, **14**, 407–431.
- MADHAVAN, A. and SMIDT, S. (1993), "An Analysis of Changes in Specialist Inventories and Quotations", *Journal of Finance*, **48**, 1595–1628.
- MENKVELD, A. J. and YUESHEN, B. Z. (2016), "The Flash Crash: A Cautionary Tale about Highly Fragmented Markets" (Technical report), <https://ssrn.com/abstract=2243520> or <http://dx.doi.org/10.2139/ssrn.2243520>.
- MORRIS, S. (1995), "The Common Prior Assumption in Economic Theory", *Economics and Philosophy*, **11**, 227–253.
- MOSKOWITZ, T. J., OOI, Y. H. and PEDERSEN, L. H. (2012), "Time Series Momentum", *Journal of Financial Economics*, **104**, 228–250.
- OBIZHAEVA, A. A. and WANG, J. (2013), "Optimal Trading Strategy and Supply/Demand Dynamics", *Journal of Financial Markets*, **16**, 1–32.
- PUCKETT, A. and YAN, X. (2011), "The Interim Trading Skills of Institutional Investors", *Journal of Finance*, **66**, 601–633.
- ROSTEK, M. and WERETKA, M. (2012), "Price Inference in Small Markets", *Econometrica*, **80**, 687–711.
- STAFFS OF THE CFTC AND SEC (2010a), *Preliminary Findings Regarding the Market Events of May 6, 2010*. Report of the Staffs of the CFTC and SEC to the Joint Advisory Committee on Emerging Regulatory Issues. May 18, 2010. Preliminary Report.
- STAFFS OF THE CFTC AND SEC (2010b), *Findings Regarding the Market Events of May 6, 2010*. Report of the Staffs of the CFTC and SEC to the Joint Advisory Committee on Emerging Regulatory Issues. September 30, 2010. Final Report.
- TOWNSEND, R. M. (1983), "Forecasting the Forecasts of Others", *Journal of Political Economy*, **91**, 546–588.
- TREYNOR, J. (1995), "The Only Game in Town", *Financial Analysts Journal*, **51**, 81–83 (reprinted from *Financial Analysts Journal*, 1971(22):12–14, 22).
- VAYANOS, D. (1999), "Strategic Trading and Welfare in a Dynamic Market", *Review of Economic Studies*, **66**, 219–254.
- WANG, J. (1993), "A Model of Intertemporal Asset Prices under Asymmetric Information", *Review of Economic Studies*, **60**, 249–282.
- WANG, J. (1994), "A Model of Competitive Stock Trading Volume", *Journal of Political Economy*, **102**, 127–167.

Copyright of Review of Economic Studies is the property of Oxford University Press / USA and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.