Private Information Distributions in Securities Markets

Kerry Back, Kevin Crotty, Tao Li

Rice University & City University of Hong Kong

July 14, 2018

Motivation

- A fundamental question in market microstructure is how the (possible) presence of asymmetric information affects the price and trading process.

Motivation

- A fundamental question in market microstructure is how the (possible) presence of asymmetric information affects the price and trading process.
  - A problem: information asymmetry is generally unobservable.

Motivation

- A fundamental question in market microstructure is how the (possible) presence of asymmetric information affects the price and trading process.
  - A problem: information asymmetry is generally unobservable.

- How can we use theory and the price and/or order flow data to measure private information?

Motivation

- A fundamental question in market microstructure is how the (possible) presence of asymmetric information affects the price and trading process.
  - A problem: information asymmetry is generally unobservable.
- How can we use theory and the price and/or order flow data to measure private information?
- In Back, Crotty, Li (2018), we develop and estimate a structural model of informed trading to address this question.

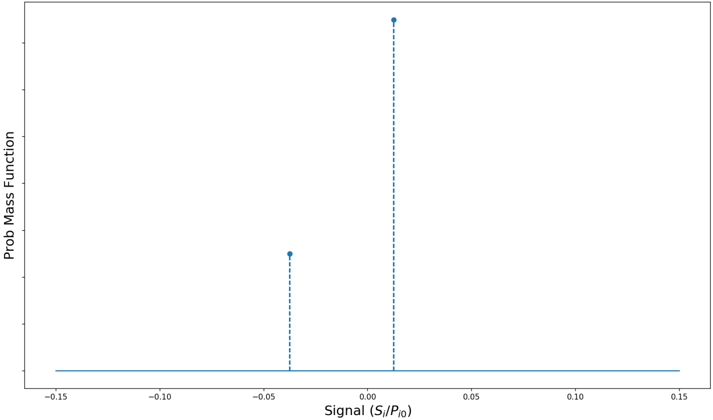
Motivation

- A fundamental question in market microstructure is how the (possible) presence of asymmetric information affects the price and trading process.
  - A problem: information asymmetry is generally unobservable.
- How can we use theory and the price and/or order flow data to measure private information?
- In Back, Crotty, Li (2018), we develop and estimate a structural model of informed trading to address this question.
- Unexplored question: **From what distribution is private information drawn?**

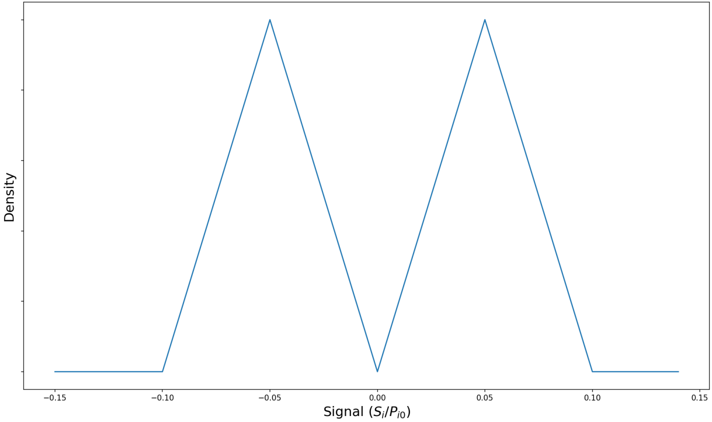
Back, Crotty, Li (2018)

- We develop a structural model of informed trading and an estimation procedure to identify information asymmetry.
    - Continuous-time Kyle model with uncertain information event and magnitude
    - Propose ML estimation that allows use of intraday observations
    - Estimation utilizes the joint distribution of returns and order flows
- For computational reasons, BCL (2018) makes simplifying assumption that signal is binary in order to facilitate estimation.

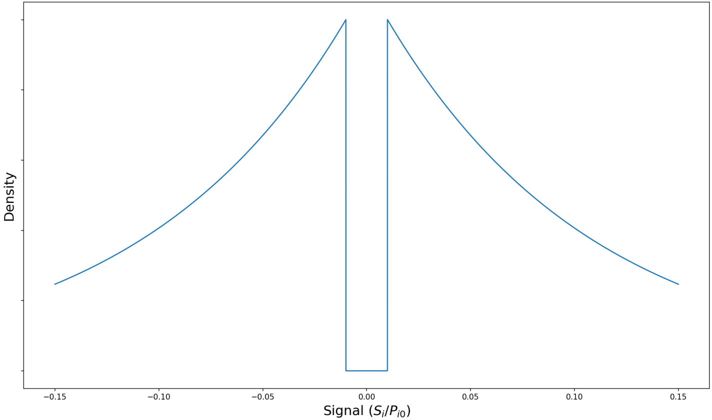
Some possible signal distributions



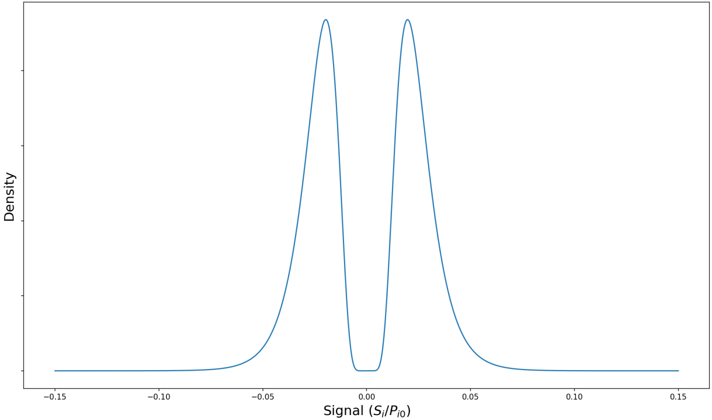
Some possible signal distributions



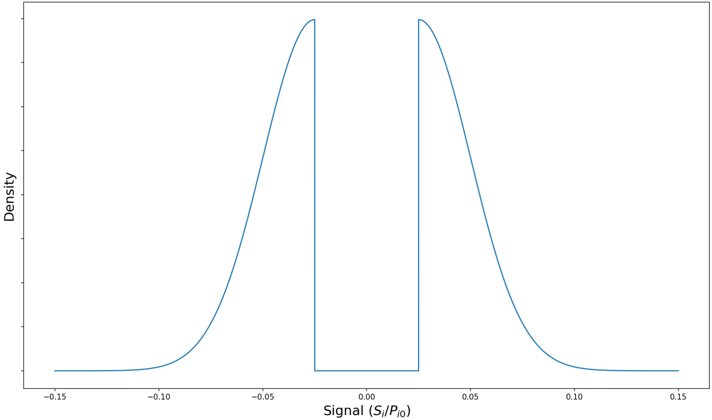
Some possible signal distributions



Some possible signal distributions

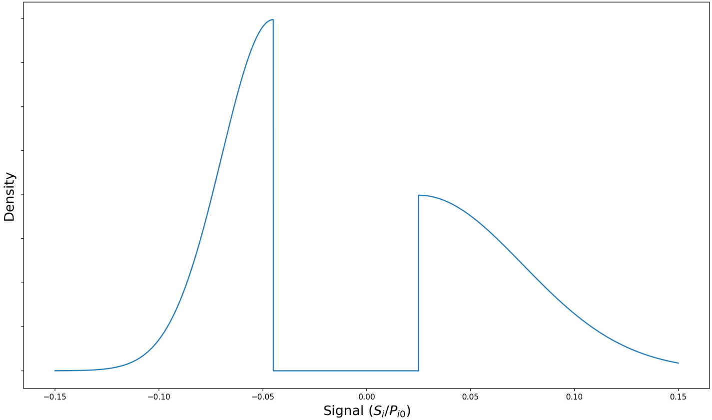


Some possible signal distributions



RICE | BUSINESS
Jones Graduate School of Business

Some possible signal distributions



Research questions

- What types of distributions are private signals drawn from?
- Are these distributions systematically linked to firm characteristics?
- What explains time-series variation in private signals?
- Are private signal distributions related to trading frictions (e.g., short-sale costs)?
- What are the asset-pricing implications of private information distributions?

Computational issues

- Estimating the model for non-binary signals involves optimizing a likelihood function containing numerical integration (in the pricing function).

	Computing Time in Python	
Signal Type	Single Firm-Year	25-yr panel of 2500 firms
Triangular	15 hrs	1,070 yrs
Exponential	7 hrs	500 yrs
Normal	4 hrs	285 yrs
Lognormal	22 hrs	1,570 yrs

Model

- Continuous-time Kyle model with asset traded on time interval $[0, 1]$
- Liquidity trades: Brownian motion Z with standard deviation σ
- Risk-neutral competitive market makers

Model

- Continuous-time Kyle model with asset traded on time interval $[0, 1]$
- Liquidity trades: Brownian motion Z with standard deviation σ
- Risk-neutral competitive market makers
- Information event at beginning with probability α
    - If there is an information event, a single trader sees a zero-mean signal S
    - If there is no information event, the trader is still present in the market as a value trader.
- Public information = martingale V

Definition of Equilibrium

- Let $\xi =$ indicator of information event (1 if yes, 0 if no).
- Let $Y = X + Z$ where $X =$ strategic trader's inventory.
- Market makers observe cumulative order imbalances Y and public information V.

Definition of Equilibrium

- Let $\xi =$ indicator of information event (1 if yes, 0 if no).
- Let $Y = X + Z$ where $X =$ strategic trader's inventory.
- Market makers observe cumulative order imbalances Y and public information V.

- Price must equal the expected value of the asset conditional on the market makers' information and given the trading strategy of the strategic trader:

$$P_t = V_t + \mathsf{E}\left[\xi S \mid \mathcal{F}_t^{V,Y}\right]$$

RICE | BUSINESS
Jones Graduate School of Business

Definition of Equilibrium

- Let $\xi =$ indicator of information event (1 if yes, 0 if no).
- Let $Y = X + Z$ where $X =$ strategic trader's inventory.
- Market makers observe cumulative order imbalances Y and public information V.

- Price must equal the expected value of the asset conditional on the market makers' information and given the trading strategy of the strategic trader:

$$P_t = V_t + \underbrace{\mathsf{E}\left[\xi S \mid \mathcal{F}_t^{V,Y}\right]}_{p(t,Y_t)}$$

RICE | BUSINESS
Jones Graduate School of Business

Definition of Equilibrium

- Let $\xi =$ indicator of information event (1 if yes, 0 if no).
- Let $Y = X + Z$ where $X =$ strategic trader's inventory.
- Market makers observe cumulative order imbalances Y and public information V.

- Price must equal the expected value of the asset conditional on the market makers' information and given the trading strategy of the strategic trader:

$$P_t = V_t + \underbrace{\mathsf{E}\left[\xi S \mid \mathcal{F}_t^{V,Y}\right]}_{p(t,Y_t)}$$

- Strategic trades must be optimal. The strategic trader chooses a rate of trade to maximize expected profits.

Order Imbalances in Equilibrium – Brownian Bridge

- Let F denote the distribution function of the normally distributed variable Z_1.
- Let G denote the continuous distribution function of the signal S.
- Set $y_L = F^{-1}(\alpha G(0))$ and $y_H = F^{-1}(1 - \alpha + \alpha G(0))$, so

$$\underbrace{\alpha\,\text{prob}(S \leq 0)}_{\text{Uncond. Prob. of Bad News}} = \text{prob}(Z_1 \leq y_L)\,,$$

and

$$\underbrace{\alpha\,\text{prob}(S > 0)}_{\text{Uncond. Prob. of Good News}} = \text{prob}(Z_1 > y_H)\,.$$

Order Imbalances in Equilibrium – Brownian Bridge

- Let F denote the distribution function of the normally distributed variable Z_1.
- Let G denote the continuous distribution function of the signal S.
- Set $y_L = F^{-1}(\alpha G(0))$ and $y_H = F^{-1}(1 - \alpha + \alpha G(0))$, so

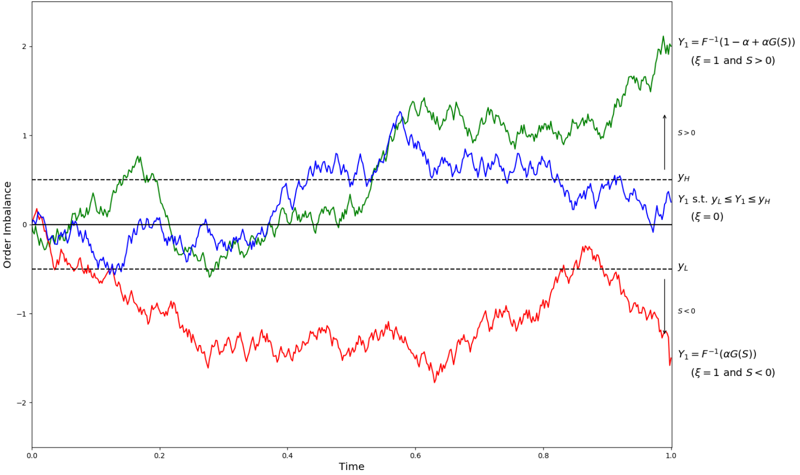
$$\underbrace{\alpha\, \text{prob}(S \leq 0)}_{\text{Uncond. Prob. of Bad News}} = \text{prob}(Z_1 \leq y_L)\,,$$

and

$$\underbrace{\alpha\, \text{prob}(S > 0)}_{\text{Uncond. Prob. of Good News}} = \text{prob}(Z_1 > y_H)\,.$$

- In equilibrium, final cumulative order flows (Y_1) satisfy:
  - $Y_1 = F^{-1}(\alpha G(S)) < y_L$ when there is a low signal ($\xi S < 0$),
  - $Y_1 = F^{-1}(1 - \alpha + \alpha G(S)) > y_H$ when there is a high signal ($\xi S > 0$),
  - $y_L \leq Y_1 \leq y_H$ when there is no info event ($\xi = 0$).

Order Imbalance Paths in Equilibrium



Equilibrium Trades

- Equilibrium rate of trade depends on t, Y_t, and whether $\xi S < 0$, $\xi S = 0$, or $\xi S > 0$:

$$\mathrm{E}[Z_1 - Z_t \mid Z_t = Y_t, \xi S] =$$

$$\begin{cases} F^{-1}(\alpha G(s)) - Y_t & \text{if } \xi S < 0\,, \\ \mathrm{E}[Z_1 \mid Z_t = Y_t, y_L \le Z_1 \le y_H] - Y_t & \text{if } \xi = 0\,, \\ F^{-1}(1 - \alpha + \alpha G(s)) - Y_t & \text{if } \xi S > 0\,. \end{cases} \tag{1}$$

divided by $1 - t$.

Equilibrium Trades

- Equilibrium rate of trade depends on t, Y_t, and whether $\xi S < 0$, $\xi S = 0$, or $\xi S > 0$:

$$\mathrm{E}[Z_1 - Z_t \mid Z_t = Y_t, \xi S] =$$

$$\begin{cases} F^{-1}(\alpha G(s)) - Y_t & \text{if } \xi S < 0\,, \\ \mathrm{E}[Z_1 \mid Z_t = Y_t, y_L \leq Z_1 \leq y_H] - Y_t & \text{if } \xi = 0\,, \\ F^{-1}(1 - \alpha + \alpha G(s)) - Y_t & \text{if } \xi S > 0\,. \end{cases} \tag{1}$$

divided by $1 - t$.

- Market makers perceive order flows Y as a Brownian motion with zero drift and std deviation σ.

Equilibrium Prices (and a computational challenge)

- Given the history of Y through time t, the equilibrium price is

$$p(t, Y_t) = \int_{-\infty}^{y_L} \overbrace{G^{-1}\left(\frac{F(z)}{\alpha}\right)}^{\text{Signal Value}} \overbrace{f(z \mid t, Y_t)}^{\text{Density of } Z_1 \text{ cond. on } Z_t = Y_t} \, \mathrm{d}z + \int_{y_H}^{\infty} G^{-1}\left(\frac{F(z) - 1 + \alpha}{\alpha}\right) f(z \mid t, Y_t) \, \mathrm{d}z \,.$$

Equilibrium Prices (and a computational challenge)

- Given the history of Y through time t, the equilibrium price is

$$p(t, Y_t) = \int_{-\infty}^{y_L} \overbrace{G^{-1}\left(\frac{F(z)}{\alpha}\right)}^{\text{Signal Value}} \overbrace{f(z \mid t, Y_t)}^{\text{Density of } Z_1 \text{ cond. on } Z_t = Y_t} \, \mathrm{d}z + \int_{y_H}^{\infty} G^{-1}\left(\frac{F(z) - 1 + \alpha}{\alpha}\right) f(z \mid t, Y_t) \, \mathrm{d}z \, .$$

- BCL (2018) makes a simplifying assumption about the signal distribution - signal is either high or low: $L < 0 < H$.
- Simplified pricing function:

$$p(t, Y_t) = L \times \overbrace{\mathrm{N}\left(\frac{y_L - Y_t}{\sigma\sqrt{1-t}}\right)}^{\Pr(\text{low info} \mid t, Y_t)} + H \times \overbrace{\mathrm{N}\left(\frac{Y_t - y_H}{\sigma\sqrt{1-t}}\right)}^{\Pr(\text{high info} \mid t, Y_t)} .$$

Estimation

- Use the joint distribution of intraday prices and order imbalances.
- Timing assumptions:
  - trading period corresponds to a day.
  - parameters are stable across year.
- The gross return through time t is

$$\frac{P_{it}}{P_{i0}} = \frac{V_{it}}{V_{i0}} + p(t, Y_{it})\,.$$

- V_{it} is geometric Brownian motion with volatility δ.

Estimation

- Use the joint distribution of intraday prices and order imbalances.
- Timing assumptions:
  - trading period corresponds to a day.
  - parameters are stable across year.
- The gross return through time t is

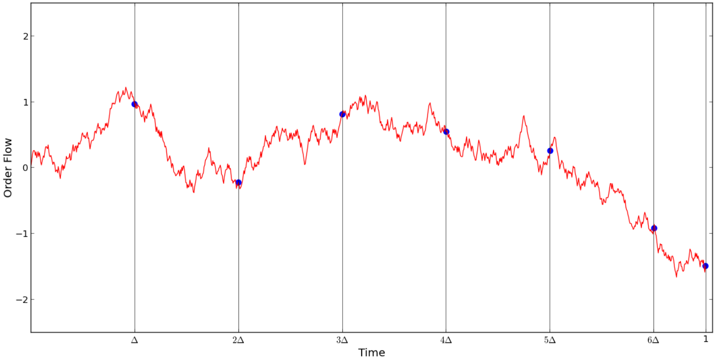
$$\frac{P_{it}}{P_{i0}} = \frac{V_{it}}{V_{i0}} + p(t, Y_{it}) .$$

- V_{it} is geometric Brownian motion with volatility δ.

- Some signal parametrization: BCL (2018) assumes a binary zero-mean signal with magnitude parameter κ such that

$$(H_i - L_i)/P_{i0} = 2\kappa$$

Timing of Observations



- Observe $k+1$ daily prices (P_{ij}) and order-flows (Y_{ij}) at $t_1, \ldots, t_{k+1}$
  - $t_{k+1} = 1$ being the close
  - Evenly spaced intraday observations: $t_j = j\Delta$ for $\Delta > 0$ and $j \leq k$.

Day-i likelihood

- The log-likelihood of observing day-i sample is:

$$\mathcal{L}_i = \log\left(f\left(\frac{P_{i1}}{P_{i0}}, \ldots, \frac{P_{i,k+1}}{P_{i0}} | Y_i\right) f(Y_i)\right)$$

- On each day i, the vector $Y_i = (Y_{i,t_1}, \ldots, Y_{i,t_{k+1}})'$ is normally distributed with mean 0 and covariance matrix $\sigma^2 \Delta \Sigma$.

$$\Sigma = \begin{pmatrix} 1 & 1 & \cdots & 1 & 1 \\ 1 & 2 & \cdots & 2 & 2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 2 & \cdots & k & k \\ 1 & 2 & \cdots & k & 1/\Delta \end{pmatrix}.$$

Day-i likelihood

- The density function of $(P_{i1}/P_{i0}, \ldots, P_{i,k+1}/P_{i0})$ conditional on Y_i is

$$f(U_{i1}, \ldots U_{i,k+1})\mathrm{e}^{-\sum_{j=1}^{k+1} U_{ij}},$$

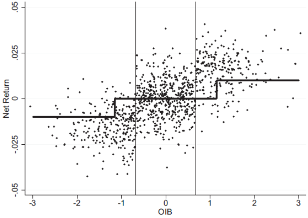
where f denotes the multivariate normal density function with mean vector $-(\delta^2\Delta/2)\Gamma$ and covariance matrix $\delta^2\Delta\Sigma$ and

$$\Gamma = \begin{pmatrix} 1 \\ 2 \\ \vdots \\ k \\ 1/\Delta \end{pmatrix}$$

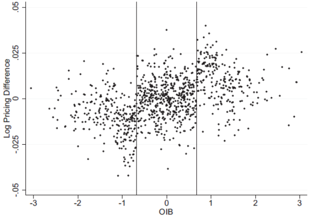
and

$$U_{ij} = \log\left(\frac{P_{ij}}{P_{i0}} - p(t_j, Y_{ij})\right) \tag{2}$$

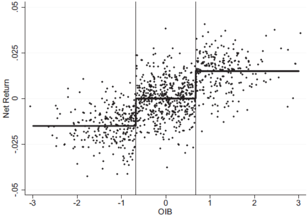
Intuition



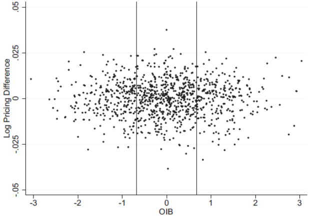
(a) $\hat{\alpha} = 0.25, \hat{\kappa} = 0.01, \hat{p}_L = 0.5$



(b) $\hat{\alpha} = 0.25, \hat{\kappa} = 0.01, \hat{p}_L = 0.5$

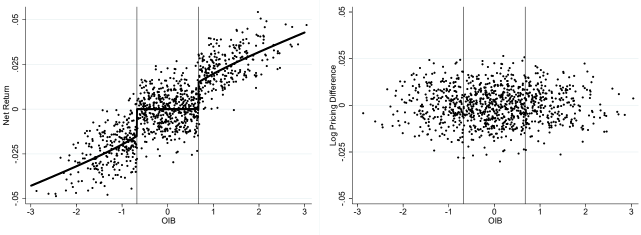


(c) $\hat{\alpha} = 0.50, \hat{\kappa} = 0.015, \hat{p}_L = 0.5$

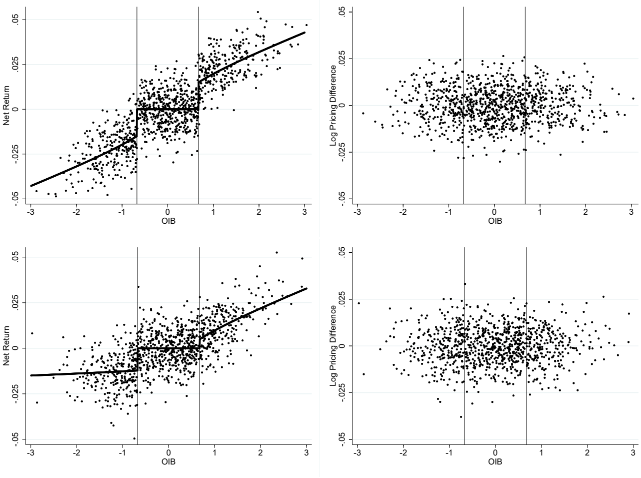


(d) $\hat{\alpha} = 0.50, \hat{\kappa} = 0.015, \hat{p}_L = 0.5$

More general distributions



More general distributions



- Estimating private information distributions from prices and trading data could provide insights into the information environment in financial markets.

Thanks for your feedback!