Journal of Econometrics 200 (2017) 312–325

Contents lists available at ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

Semiparametric identification of the bid–ask spread in extended Roll models

Xiaohong Chen [a], Oliver Linton [b], Yanping Yi [c,*]

[a] Cowles Foundation for Research in Economics, Yale University, PO Box 208281, New Haven, CT 06520-8281, USA
[b] Department of Economics, University of Cambridge, Austin Robinson Building, Sidgwick Avenue, Cambridge, CB3 9DD, UK
[c] School of Economics, Shanghai University of Finance and Economics, 777 Guoding Road, Shanghai, 200433, China

ARTICLE INFO

Article history:
Available online 28 June 2017

JEL classification:
C12
C13
C14

Keywords:
Bid–ask spread
Roll model
Semiparametric identification
Latent variables

ABSTRACT

This paper provides new identification results for the bid–ask spread and the nonparametric distribution of the latent fundamental price increments (ε_t) from the observed transaction prices alone. The results are established via the characteristic function approach, and hence allow for discrete or continuous ε_t and the observed price increments do not need to have any finite moments. Constructive identification (and overidentification) results are established first in the basic Roll (1984) model, and then in various extended Roll models, including general unbalanced order flow, serially dependent latent trade direction indicators, adverse selection, random spread and a multivariate Roll model.

1. Introduction

The (quoted) bid–ask spread of a financial asset is the difference between the best quoted prices for an immediate purchase and an immediate sale of that asset. The spread represents a potential profit for the market maker handling the transaction, and is a major part of the transaction cost facing investors, especially since the elimination of commissions and the reduction in exchange fees that has happened in the last twenty years; see for example Jones (2002) and Angel et al. (2011). Measuring the bid–ask spread in practice can be quite time consuming (since it requires reconstruction of the limit order book) and may be subject to a number of potential accuracy issues due to the quoting strategies of High Frequency Traders, for example.

The seminal paper Roll (1984) provides a simple market microstructure model that allows one to estimate the bid–ask spread from observed transaction prices alone, without information on the underlying bid–ask price quotes and the order flow (i.e., whether a trade was buyer- or seller-induced). This is particularly useful for long historical data sets, which are often limited in their scope. For instance, Hasbrouck (2009) notes that "investigations into the role of liquidity and transaction costs in asset pricing must generally confront the fact that while many asset pricing tests make use of US equity returns from 1926 onward, the high-frequency data used to estimate trading costs are usually not available prior to 1983. Accordingly, most studies either limit the sample to the post-1983 period of common coverage or use the longer historical sample with liquidity proxies estimated from daily data". Another area where the available data is limited is open-outcry markets (like the CME), in which bid and ask quotes by traders expire (if not filled) without recording (see, e.g., Hasbrouck (2004) for more details).

In the famous Roll (1984) model, an observed (log) asset price p_t evolves according to

$$p_t = p_t^* + I_t \frac{s_0}{2}, \quad p_t^* = p_{t-1}^* + \varepsilon_t. \tag{1}$$

$$\Delta p_t := p_t - p_{t-1} = \varepsilon_t + (I_t - I_{t-1})\frac{s_0}{2}, \tag{2}$$

where p_t^* is the underlying fundamental (log) price with innovations ε_t, and the trade direction indicators $\{I_t\}$ are i.i.d. and take the values ± 1 with probability $q_0 := \Pr(I_t = 1) = 1/2$. $I_t = 1$ indicates that the transaction is a purchase, and $I_t = -1$ denotes a sale. The price p_t is observed, whereas *all other variables in* Eq. (1) *are unobserved*. The parameter of interest is the effective bid–ask spread s_0.[1] Roll (1984) assumes that $\{\varepsilon_t\}$ is serially uncorrelated and uncorrelated with the trade direction indicators $\{I_t\}$, and that

[1] The bid–ask spread in Eq. (1) is called effective bid–ask spread because it is based on the effective (average) price p_t that is paid to fill an order, and not necessarily on the quoted bid or ask price, since it might be the case that the order cannot be filled at the latter price (e.g., due to insufficient depth of the market).

* Corresponding author.
E-mail addresses: xiaohong.chen@yale.edu (X. Chen), obl20@cam.ac.uk (O. Linton), yi.yanping@mail.shufe.edu.cn (Y. Yi).

http://dx.doi.org/10.1016/j.jeconom.2017.06.013

the one period returns (i.e., the price increments) $\{\Delta p_t\}$ have finite second moments. Under these assumptions, s_0 is identified in a closed form as

$$s_0 = 2\sqrt{-\text{Cov}\,(\Delta p_t, \Delta p_{t-1})}. \tag{3}$$

Roll (1984) proposes to estimate s_0 from (3) by replacing the theoretical covariance by its empirical counterpart, i.e.,

$$\widehat{s}_{Roll} := 2\sqrt{-\widehat{\text{Cov}}\,(\Delta p_t, \Delta p_{t-1})}. \tag{4}$$

In practice, this estimator is not satisfactory, since the empirical first-order autocovariance of price changes is often positive, in which case (4) is not well-defined. Another problem is that the nonparametric distribution of the latent true one period returns (i.e., the latent fundamental price increment), $\Delta p_t^* = \varepsilon_t$, is not identifiable in the original Roll model.

In a well-known alternative, Hasbrouck (2004) proposes to strengthen Roll's modelling assumptions by assuming that $\{\varepsilon_t\}$ is i.i.d. with a known parametric distribution, and is independent of $\{I_t\}$.[2] He then uses a Bayesian Gibbs sampling methodology to estimate the spread parameter subject to a non-negativity constraint. Specifically, Hasbrouck (2004) assumes that $\varepsilon_t \sim$ i.i.d.$N(0, \sigma_\varepsilon^2)$, where the parameter σ_ε is estimated jointly with the spread s_0. Unfortunately the spread estimator of Hasbrouck (2004) performs poorly or is not well defined when ε_t is discrete or continuous but fat-tailed and/or asymmetric. Basically the spread estimator of Hasbrouck (2004) is very sensitive to departures from the assumption that $\varepsilon_t \sim$ i.i.d.$N(0, \sigma_\varepsilon^2)$. Moreover, it is difficult to justify a specific parametric distribution such as Gaussian for the latent ε_t.

The more recent empirical finance literature emphasizes several additional issues with the Roll model: (a) It assumes balanced market order flow, i.e., $q_0 = 1/2$, which may be accurate on average, but may be inaccurate for certain episodes of trading. (b) It assumes no serial correlation in trade direction indicators, i.e., I_t is uncorrelated with I_{t-j} for any $j \geq 1$. (c) Market orders are assumed not to bring any news into the fundamental prices (i.e., no adverse selection), so that I_t is uncorrelated with Δp_{t+j}^* for $j \geq 0$. (d) Spreads are constant within the sample period. Admitting any one of these effects in the model will lead to the undesired consequence that the spread estimators of Roll (1984) and Hasbrouck (2004) become *inconsistent* (i.e., biased even as sample size goes to infinity). Furthermore, without additional model assumptions, or additional observed information (such as trade volume data in addition to $\{p_t\}$), it may not be possible to identify the spread jointly with parameters describing order flow imbalance or adverse selection, for example. See, e.g., Bleaney and Li (2015) for a very recent discussion of all the above and additional problems with the original Roll model.

In this paper we propose new methods for identifying the bid–ask spread s_0 and the unknown distribution of $\{\varepsilon_t\}$ jointly from the observed time series transaction prices alone. The observed prices $\{p_t\}$ could be daily or weekly closing prices, or high-frequency intra-day prices. Our methods are based on the characteristic function approach, and hence do not require the existence of any finite moments of $\{\Delta p_t\}$, and allow the latent $\{\varepsilon_t\}$ to be discrete or continuous, symmetric or asymmetric. Under the assumption of strict stationarity of the latent process $\{\varepsilon_t, I_t\}_{t=1}^\infty$, our identification results do not require the full independence between $\{\varepsilon_t\}$ and $\{I_t\}$, and mainly impose some restrictions on the dependence structure of ε_t, ε_{t-1}, I_t, I_{t-1} and I_{t-2}. Constructive identification results for s_0 and the characteristic function (φ_ε) of ε_t or/and parameters in various extended Roll models are established based on the joint characteristic function of consecutive one period returns

$$\varphi_{\Delta p,2}(u, u') := \mathbb{E}\left[\exp\left(iu\Delta p_t + iu'\Delta p_{t-1}\right)\right] \quad \text{for any } (u, u') \in \mathbb{R}^2, \tag{5}$$

which is nonparametrically identified from the observed price increment time series $\{\Delta p_t\}$.

We first provide a closed-form solution of $(s_0, \varphi_\varepsilon)$ in the basic Roll (1984) model under a mild sub-independence assumption, which is only slightly stronger than the uncorrelatedness condition in Roll (1984) but is much weaker than the full independence between $\{\varepsilon_t\}$ and $\{I_t\}$ assumption in Hasbrouck (2004). In addition, we do not impose finite second moment of Δp_t as in Roll (1984) and Gaussian error of ε_t as in Hasbrouck (2004). We then propose solutions to the four problems (a)–(d) with the Roll model listed above. We show how to identify $(s_0, \varphi_\varepsilon)$ and other parameters associated with unbalanced order flow and/or general asymmetric supported $\{I_t\}$, or those for serially correlated $\{I_t\}$, or those capturing adverse selection effects, or the random spread. We also extend the basic Roll model to the multivariate case and derive the identification results. Again, all these are accomplished without requiring additional data.

In principle, both the basic Roll (1984) model and the various extended Roll models could fit into the vast measurement error literature (see, e.g., Li and Vuong, 1998; Carroll et al., 2006; Hu, 2008; Hu and Schennach, 2008; Chen et al., 2011; Evdokimov and White, 2012; Bonhomme et al., 2016; Hu, forthcoming, and the references therein). However, to the best of our knowledge, our identification results are not direct consequences of any existing published results. This is because the Roll model and its various extensions contain some special structures, and our identification results utilize these special features and are constructive under conditions reasonable for financial applications.

Our constructive identification results for $(s_0, \varphi_\varepsilon)$ or/and parameters in extended Roll models are derived under conditions much weaker than those in the existing literature and more realistic for financial applications when $\{p_t\}$ is the only information available. All our identification results are essentially based on solving the unknown model parameters by matching the nonparametrically identified characteristic function $\varphi_{\Delta p,2}(u, u')$ to its model-implied semiparametric counterpart. This approach actually leads to Hansen (1982) style overidentification.[3] Therefore, one could easily compute consistent estimators of s_0, the distribution of ε_t or/and other model parameters via minimum distance procedures based on empirical characteristic functions. And the overidentification restrictions allow for model specification tests. As a natural follow-up to this identification paper, Chen et al. (2017) studies in detail the estimation and testing aspects of these models and presents an interesting empirical application. In particular, based on our constructive identification results, Chen et al. (2017) provides simple sample analog estimation of the spread s_0, the characteristic function of ε_t or/and other parameters in various extended Roll models (such as order flow imbalance, adverse selections). In the simulation studies, their sample analog spread estimator does not suffer the pitfalls of the spread estimators of Roll (1984) and Hasbrouck (2004).

The rest of the paper is organized as follows: Section 2 presents the basic Roll model and identification of both the spread s_0 and the characteristic function of ε_t in closed form, allowing for $\{\Delta p_t\}$ to have infinite first absolute moments. Section 3 considers extensions to models that allow for unbalanced order flow and more general asymmetric supported $\{I_t\}$. Section 4 studies identification

[2] Hasbrouck (2004) presents an extension that relaxes the independence between $\{\varepsilon_t\}$ and $\{I_t\}$ assumption but uses additional trade volume data.

[3] See Chen and Santos (2015) for a notion of overidentification in semiparametric and nonparametric models.