# Cardiff Economics Working Papers

Working Paper No. E2013/6

# Weak exogeneity in the financial point processes

*Yongdeng Xu*

April 2013

Cardiff Business School
Aberconway Building
Colum Drive
Cardiff CF10 3EU
United Kingdom
t: +44 (0)29 2087 4000
f: +44 (0)29 2087 4419
business.cardiff.ac.uk

# Weak exogeneity in the financial point processes

Yongdeng Xu*

**August 2011**

**Abstract**

This paper analyses issues related to weak exogeneity in a financial point process. We extend the Hausman test of weak exogeneity in a time series model and propose three cases in which weak exogeneity conditions will break down. The simulation study suggested that a failure of the exogeneity assumption implied biased estimators. The bias is very large in the third case non-weak exogeneity, which makes the econometric inferences on the parameters unreliable or even misleading. We then derive an LM test for weak exogeneity. The LM test is attractive because it only requires estimation of the restricted model. The empirical results indicate that the weak exogneity of duration is often rejected for frequently traded stocks, but is less likely to be rejected for infrequently traded stocks.

**Key words:** Weak exogeneity, ACD model, LM test, point process, market microstructure

* Economic Section, Cardiff Business School, Cardiff University, UK. Email: xuy16@cardiff.ac.uk

# 1 Introduction

High frequency data is widely used in analysis of market microstructure theory. Typical high frequency data, for example Trades and Quotes (TAQ) dataset in NYSE, consists trade time, trade volume, bid-ask price and other indicators. Such high variables are naturally irregularly spaced in time; they are usually considered as financial point processes. To capture their stochastic dynamics, the so-called Multiplicative Error Model (MEM) has been proposed (Engle 2000). The basic idea is to model the positive-valued indicator in terms of the product of a (conditional autoregressive) scale factor and an innovation process with nonnegative support (i.e. GARCH-like process). The best known univariate MEM model is the Autoregressive Conditional Duration (ACD) model of Engle and Russell (1998) for the financial durations. However, when modelling financial duration and other market indicators (marks) jointly, the directly use of multivariate MEM model is restricted since joint probability distributions for nonnegative valued random variables are often not available in the literature.

The commonly used strategy is to decompose the joint distribution of duration and market marks into the product of the marginal density of duration and the conditional density of marks given duration. In estimation, if the weak exogeneity of duration is valid, then the marginal density of duration the conditional density of marks can be estimated separated equation-by-equation. This approach simplifies the estimation procedure and is generally adopted in the empirical literature; see, for example, Engle (2000),Dufour and Engle (2000), Manganelli (2005), Engle and Sun (2007). However, if the parameters in the conditional density depend on some of the parameters of the marginal process (e.g., the weak exogeneity condition fails), the estimators would be inefficient or even inconsistent, leading to invalid inference; c.f White (1981,1982).

In this paper, we consider three cases in which the weak exogeneity condition will break down and we use a Monte Carlo simulation to study the consequences of the failure of weak exogeneity. The simulation study suggested that a failure of the exogeneity assumption implied biased estimators. The bias is very large in the third case non-weak exogeneity. In empirical analysis, we derive an LM test which is similar to Dolado, Rodriguez-Poo et al. (2004). However, we use a more fruitful specification of the conditional mean, which implies that the rejection of null is less likely due to the misspecification of conditional mean. Using two groups of high frequency data, we test both the weak exogeneity of duration and the joint weak exogeneity of duration and volume in empirical analysis and find that the assumption of weak exogeneity is often rejected.

The remainder of the paper is structured as follows. Section 2 reviews the related literature on weak exogeneity. Section 3 introduces the notion of weak exogeneity and methodology. Section 4 presents a simulation study to examine the consequences of ignoring non weak exogeneity. Section 5 derives an LM test for weak exogeneity. Section 6 contains an empirical application. And section 7 is the conclusion.

# 2 Relevant literature reviews

Different definitions of exogeneity are clarified by Engle, Hendry et al. (1983); for example, weak exogeneity, strong exogeneity, super exogeneity and invariance. Weak exogeneity is proposed as an answer to the question of under what conditions can one estimate the parameters of conditional density without loss of information from neglecting the marginal process. The idea of weak exogeneity is be expressed simply by saying that estimation and inference on the parameters of the marginal density and the conditional density can be undertaken separately, without loss of efficiency, if the endogenous variable in the marginal density is weakly exogenous for parameters in the conditional density. Engle and Hendry (1993) develop the different classes of tests of weak exogeneity. In particular, if the marginal processes are constant, the Wu-Hausman test is commonly used for testing weak exogeneity. The original Hausman test (Hausman 1978) contrasts two estimates obtained from different estimators (unconstrained and constrained parametric models). Under a null hypothesis, both of these estimators are consistent while only the second estimator is efficient. Under the alternative hypothesis of endogeneity, the first estimator is consistent while the second is not. This Hausman statistic has, under the null hypothesis, an asymptotically chi-squared distribution with the number of degrees of freedom equal to the number of endogenous regressors. An alternative to the Hausman contrast test is the two-stage Wald version test, originally derived by Wu (1973). In the first stage, by careful construction[1], a reduced form model (marginal

[1] See Terza, Basu and Rathouz (2008) for the conditions of choosing IV.

model) is specified for the endogenous variables which are estimated consistently. Then, the fitted values of the endogenous variables are computed and in the second stage, the conditional model is augmented by plugging in the fitted values as additional regressors. If the fitted values of the endogenous variables are jointly significant in the conditional model, the null hypothesis of weak exogeneity is rejected. A simple Wald statistic can be used to test the joint significance. Effectively, this two-stage Wald version test leads to a test which is asymptotically equivalent to the Hausman contrast test [an algebraic derivation of this result can be found in the Davidson and MacKinnon (2004, Section 8.7)]. Using Monte Carlo simulation, Chmelarova (2007) shows, under a series of different conditions, that the Wald version of the Hausman test often has better properties that the contrast version.

The Hausman test has been widely used in various areas, such as macroeconomics, health economics, and international trade. For example, Fischer (1993) and Boswijk and Urbain (1997) test the weak exogeneity of Swiss money Demand. Terza, Basu et al. (2008) address the endogeneity in an econometric model of health. Staub (2009) tests for the exogeneity of a binary explanatory variable in a count data regression model. Darrat, Hsu et al. (2000) test export exogeneity in Taiwan. However, Hausman tests suffer from three problems when applied to a market point processes and MEM models. Firstly, Hausman set is initially developed for a test of cross section model, whereas the MEM/ACD model is a time-series model and the dynamics of endogenous variables should also be considered. Secondly, the test is developed in a Gaussian/linear framework, whereas the market point process usually belongs to the exponential family. Thirdly, correct specification of the conditional mean is a fundamental assumption underlying the test, since the rejection of the null hypothesis could be due either to the absence of weak exogeneity or to the misspecification of the conditional mean. Dolado, Rodriguez-Poo et al. (2004) consider the latter two problems and propose a LM test for weak exogeneity under the pseudo-maximum likelihood condition. They analyze the relationship between trade size and trade duration and find that the hypothesis of weak exogeneity of trade duration if often rejected.

We address the first problem in this paper and discuss three cases of where the weak exogeneity condition may break down, for which we extend the Hausman test of exogeneity in a time series model. A Monte Carlo simulation study is used to examine the consequences of the failing of weak exogeniety. In empirical analysis, we derive an LM test, which is similar to Dolado, Rodriguez-Poo et al. (2004). However, we use a more powerful specification of the conditional mean and test both weak exogeneity of duration and jointly weak exogeneity of duration and volume.

# 3 Methodology

## *3.1 Formal Definition of Weak Exogeneity*

As in Engle, Hendry et al. (1983) and Engle and Hendry (1993), we start with a bivariate stochastic process $\{x_t, y_t\}$ and the joint density $f(x_t, y_t | \psi_t; \theta)$, where $\psi_t$ is the information set which includes lags and other important variables. Commonly, the joint density $(x_t, y_t)$ can be factorized into the product of the marginal density $x_t$ and conditional density of $y_t$ given $x_t$

$$f(x_t, y_t | \psi_t; \theta) = f_x(x_t | \psi_t; \theta^x) f_{y|x}(y_t | x_t, \psi_t; \theta^y) \quad (1)$$

where $(\theta^x, \theta^y) \in \Theta$. Let $\vartheta = f(\theta)$ be the parameters of interest, which are assumed to be present only in the conditional density. The key issue, addressed by Engle, Hendry et al. (1983),is to know under what conditions it is possible to estimate $\vartheta$ just as function of $\theta^y$ and without loss of information. In other words, that all the information needed for estimation of $\vartheta$ is $f_{y|x}$.

Engle, Hendry et al. (1983) define a variable of $x_i$ as weakly exogenous for a set of parameters of interest $\vartheta$ if:

i) $\vartheta = f(\theta)$, $\vartheta$ is a function of parameters $\theta^y$ alone, and

ii) $\theta^y$ and $\theta^x$ are variation free, i.e. $(\theta^x, \theta^y) \in \Theta^x \times \Theta^y$.

Consequently, if $x_i$ is weakly exogenous for $\vartheta$, there is no loss of information about $\vartheta$ from neglecting the process determining $x_i$. Otherwise, the estimation of $\theta^y$ would be inefficient or even inconsistent.

In econometric, the marginal density might also be interested. Weak exogeneity is also expressed simply by saying that estimation and inference on $\theta^x$ and $\theta^y$ can be undertaken separately without loss of information, if $x_i$ is weak exogenous for $\theta^y$. Engle, Hendry et al. (1983) further introduce the notation of "sequential cut" and "cross-restriction" to illustrate weak exogeneity, saying that $x_i$ is weak exogenous for $\theta^y$, if $[f_x(x_t|\psi_t;\theta^x), f_{y|x}(y_t|x_t,\psi_t;\theta^y)]$ operates a sequential cut on $f(x_t, y_t|\psi_t;\theta)$, or if $\theta^x$ and $\theta^y$ is not subject to "cross-restriction".

## *3.2 Different Types of Weak Exogeneity in Financial Point Processes*

Manganelli (2005) proposes a framework for the joint dynamics of trading duration, volume and price volatility. This model incorporates both causality and feedback effect among variables of interested and thereby can explain the various strategic models in the market microstructure literature. So we take Manganelli (2005)'s model for specification the dynamics of financial point process. To simplify, we only consider the jointly distribution of duration and volume. Define $\{d_t, v_t\}$, $t = 1,\cdots,T$ as the two-dimensional time series associated with intraday trading duration and trading volume. In particular, duration is defined as the time elapsing between consecutive trades, volume is the trade size associated with each transaction. The bivariate trading process- duration, volume - can be modelled as follows:

$$\{d_t, v_t\} \sim f(d_t, v_t \mid \Omega_t;\theta) \tag{2}$$

where $\Omega_t$ denotes the information available up to period $t$ and $\theta$ is a vector incorporating the parameters of interest.

In Manganelli (2005)'s framework, the joint distribution is decomposed into the product of marginal density of durations and the conditional density of volumes given durations:

$$\{d_t, v_t\} \sim g(d_t \mid \Omega_t;\theta_d)\cdot h(v_t \mid d_t, \Omega_t;\theta_v)\,. \tag{3}$$

Manganelli (2005) specifies the following univariate MEM model for duration and volume:

$$\begin{aligned} d_t &= \psi_t(\theta_d;\Omega_t)u_t,\ \ u_t \sim i.i.d.(1,\sigma_u^2). \\ v_t &= \varphi_t(\theta_v; d_t, \Omega_t)\eta_t,\ \ \eta_t \sim i.i.d.(1,\sigma_\eta^2). \end{aligned} \tag{4}$$

where $(\psi_t, \phi_t)$ are the conditional expectations of duration and volume, $\theta = (\theta_d, \theta_v)$ is a vector of s parameters of interest. The innovation terms are uncorrelated with each other by construction.

The log likelihood can be expressed as:

$$L(\theta^v, \theta^d) = \sum_{t=1}^{T} [\log g(d_t | \Omega_t; \theta^d) + \log h(v_t | d_t, \Omega_t; \theta^v)]. \tag{5}$$

Follows Manganelli (2005), the density of duration and the density of volume conditional on duration are expressed as:

$$\begin{aligned} g(d_t | \Omega_t; \theta^d) \quad &\sim \quad d_t = \psi_t(\theta_d; \Omega_t)\varepsilon_t, \\ &\quad\quad \psi_t = a_0 + a_1 d_{t-1} + a_2 v_{t-1} + \alpha_3 \psi_{t-1} + a_4 \phi_{t-1}. \\ h(v_t | d_t, \Omega_t; \theta^v) \quad &\sim \quad v_t = \varphi_t(\theta_v; d_t, \Omega_t)\eta_t, \\ &\quad\quad \phi_t = b_0 + b_1 d_{t-1} + b_2 v_{t-1} + b_3 \psi_{t-1} + b_4 \phi_{t-1} + b_5 d_t. \end{aligned} \tag{6}$$

It is well know that estimation and inference on the parameters characterising each density can be undertaken separately, without loss of efficiency, if two of following condition hold: a) weak exogeneity, and b) the respective densities are correctly specified. Consequently, failing of weak exogeneity would result in inefficient or even inconsistent estimators c.f White (1981,1982), leading to unreliable inferences.

In the econometrics literature, the Hausman specification is usually used to test weak exogeneity. As explained by Engle, Hendry et al. (1983) , if none of the parameters in the marginal model appear in the conditional model, then weak exogeneity is valid. Therefore, testing weak exogeneity implies testing the significance of the predictor from the marginal model, in the conditional model. However, Hausman set is initially developed for a test of cross section model, whereas the MEM/ACD model is a time-series model and the dynamics of endogenous variables should also be considered. In this section, we extend the Hausman test of weak exogeneity in a time series model and propose three cases in which the weak exogeneity condition will break down. We use the so called "non weak exogeneity" thereafter to express the notation that weak exogneiety condition breaks down.

***Case 1:*** The first case of non weak exogeneity is directly motivated by Manganelli (2005)'s model.

$$d_t = \psi_t(\theta_d;\Omega_t)\varepsilon_t,\ \ \varepsilon_t \sim i.i.d(1,\sigma_\varepsilon^2).$$
$$v_t = \varphi_t(\theta_v; d_t,\Omega_t)\eta_t,\ \ \eta_t \sim i.i.d(1,\sigma_\eta^2).$$
$$\psi_t = a_0 + a_1 d_{t-1} + a_2 v_{t-1} + \alpha_3\psi_{t-1} + a_4\phi_{t-1},$$
$$\phi_t = b_0 + b_1 d_{t-1} + b_2 v_{t-1} + b_3\psi_{t-1} + b_4\phi_{t-1} + b_5 d_t. \tag{7}$$

Writing it in matrix form as:

$$\begin{pmatrix}\psi_t\\ \varphi_t\end{pmatrix} = \begin{pmatrix}a_0\\ b_0\end{pmatrix} + \begin{pmatrix}a_1 & a_2\\ b_1 & b_2\end{pmatrix}\begin{pmatrix}d_{t-1}\\ v_{t-1}\end{pmatrix} + \begin{pmatrix}a_3 & a_4\\ b_3 & b_4\end{pmatrix}\begin{pmatrix}\psi_{t-1}\\ \varphi_{t-1}\end{pmatrix} + \begin{pmatrix}0 & 0\\ b_5 & 0\end{pmatrix}\begin{pmatrix}d_t\\ v_t\end{pmatrix}. \tag{8}$$

Both the parameters in marginal density and conditional density are interested. In order to optimize the two processes separately, the assumption of weak exogeneity has to be imposed. The condition for weak exogeneity is $\alpha_4 = 0,\ b_3 = 0$, since only under this condition can $[g(d_t|\Omega_t;\theta^d), h(v_t|d_t,\Omega_t;\theta^v)]$ operate a sequential cut on $f(d_t, v_t|\Omega_t;\theta)$ whereupon there is no cross-section restrictions between marginal and conditional density (Engle, Hendry et al. 1983).

If we look at it in another way and assume:

$$d_t - \psi_t = \varepsilon_{1t} \sim i.i.d(0,\tilde{\sigma}_\varepsilon^2)$$
$$v_t - \varphi_t = \varepsilon_{2t} \sim i.i.d(0,\tilde{\sigma}_\eta^2), \tag{9}$$

then the above becomes[2]

$$\begin{pmatrix}d_t\\ v_t\end{pmatrix} = \begin{pmatrix}a_0\\ b_0\end{pmatrix} + \begin{pmatrix}a_1+\alpha_3 & a_2+\alpha_4\\ b_1+b_3 & b_2+b_4\end{pmatrix}\begin{pmatrix}d_{t-1}\\ v_{t-1}\end{pmatrix} + \begin{pmatrix}0 & 0\\ b_5 & 0\end{pmatrix}\begin{pmatrix}d_t\\ v_t\end{pmatrix} + \begin{pmatrix}\varepsilon_{1t}\\ \varepsilon_{2t}\end{pmatrix} - \begin{pmatrix}a_3 & a_4\\ b_3 & b_4\end{pmatrix}\begin{pmatrix}\varepsilon_{1t-1}\\ \varepsilon_{2t-1}\end{pmatrix}. \tag{10}$$

Again, the condition of weak exogeneity is that $\alpha_4 = 0,\ b_3 = 0$, since only under such condition can $\theta_d$ and $\theta_v$ be variation free and subject to no cross equation restrictions. Generally, if any lagged expected (or fitted) value from marginal model is present in the conditional model, or any lagged expected (or fitted) value from conditional model is present in marginal model, the weak exogeneity condition will break down.

***Case 2:*** The second case of non weak exogeneity is based on the Hausman specification. As explained by Engle, Hendry et al. (1983), Hausman test for weak exogeneity implies testing the significance of the predicted variable from the marginal model, in the conditional model. In the MEM/ACD models, it is natural to use the

[2] See Appendix 1 for proof.

conditional expected value instead of predicted value, since the conditional expectation of duration is directly measured. Assuming the first case of weak exogeneity is satisfied, the joint distribution of duration and volume as specified in Eq. (6) has the following Hausman specification:

$$\begin{aligned} d_t &= \psi_t(\theta_d;\Omega_t)\varepsilon_t,\ \ \varepsilon_t \sim i.i.d(1,\sigma_\varepsilon^2) \\ v_t &= \varphi_t(\theta_v;d_t,\Omega_t)\eta_t,\ \ \eta_t \sim i.i.d(1,\sigma_\eta^2) \\ \psi_t &= a_0 + a_1 d_{t-1} + a_2 v_{t-1} + \alpha_3 \psi_{t-1} \\ \phi_t &= b_0 + b_1 d_{t-1} + b_2 v_{t-1} + b_3 \phi_{t-1} + b_4 d_t + b_5 \psi_t. \end{aligned} \tag{11}$$

Under assumption that the parameters of interest depend solely on the parameters of the conditional distribution, i.e. $\vartheta = f(b_0, b_1, b_2, b_3, b_4)$ and supposing that the expected duration $\psi_t$ is estimated from the marginal model, then in order to test weak exogeneity of duration, it suffices to test the significance of $\psi_t$ in the conditional distribution. In such a case, the parameters of interest are not subject to cross equation restrictions and $\vartheta$ are variation free with respect to the parameters of the duration process. And it is sufficient to test $b_5$ =0 in Eq.(11) in order to test weak exogeneity in this case.

If we look at in another way and assume

$$\begin{aligned} d_t - \psi_t &= \varepsilon_{1t} \sim i.i.d(0,\tilde{\sigma}_\varepsilon^2) \\ v_t - \phi_t &= \varepsilon_{2t} \sim i.i.d(0,\tilde{\sigma}_\eta^2) \end{aligned} \tag{12}$$

Eq.(11) then becomes[3]

$$\begin{aligned} d_t &= \alpha_0 + \alpha_1 d_{t-1} + a_2 v_{t-1} + \alpha_3 \psi_{t-1} + \varepsilon_{1t} \\ v_t &= b_o + b_1 d_{t-1} + b_2 v_{t-1} + b_3 \phi_{t-1} + (b_4 + b_5) d_t + \varepsilon'_{2t} \end{aligned} \tag{13}$$

where $\varepsilon'_{2t} = -b_5\varepsilon_{1t} + \varepsilon_{2t}$.Therefore, $Cov(\varepsilon_{1t},\varepsilon'_{2t}) = -b_5\,\mathrm{var}(\varepsilon_{1t}) = -b_5\tilde{\sigma}_\varepsilon^2$. The condition for weak exogeneity is that $Cov(\varepsilon_{1t},\varepsilon'_{2t}) = 0$ because in such a case, the parameters of interest $\theta_v$ are not subject to cross equation restrictions and are variation free with respect to the parameters from duration equation $\theta_d$.

The Hausman specification test might also take another form, see for example Dolado, Rodriguez-Poo et al. (2004). They specify the following functional form for testing weak exogeneity:

[3] See Appendix 2 for derivation.

$$v_t = \varphi_t(\theta_v; d_t, \Omega_t)\eta_t$$
$$\phi_t = b_0 + b_1 d_{t-1} + b_2 v_{t-1} + b_3\phi_{t-1} + (b_4 + b_5\psi_t)d_t. \quad (14)$$

Generally with the Hausman test for weak exogeneity, if any linear or nonlinear forms of expected duration enter the conditional model, the weak exogeneity of duration will break down[4].

***Case 3:*** Motivated by the case 2 of non weak exogeneity, we may consider a more restrictive case of non weak exogeneity. Let's look at the following model of duration and volume:

$$\begin{aligned} d_t &= \psi_t(\theta_d; \Omega_t)\varepsilon_t, \quad \begin{pmatrix}\varepsilon_t \\ \eta_t\end{pmatrix} \sim i.i.d.(\mathrm{I}, \Omega) \\ v_t &= \phi_t(\theta_v; d_t, \Omega_t)\eta_t, \\ \psi_t &= a_0 + a_1 d_{t-1} + a_2 v_{t-1} + \alpha_3\psi_{t-1} \\ \phi_t &= b_0 + b_1 d_{t-1} + b_2 v_{t-1} + b_3\phi_{t-1} + b_4 d_t \end{aligned} \quad (15)$$

where I is the unit vector, $corr(\varepsilon_t, \eta_t) = \rho, \rho \neq 0$.

If the error terms from the marginal and conditional model are correlated, the weak exogeneity condition will breaks down, since the parameters of volume equation $\theta_v$ are subject to cross equation restrictions and are not variation free with respect to parameters from the duration equation $\theta_v$. The condition of weak exogeneity in this case is that $\varepsilon_t$ and $\eta_t$ are uncorrelated. In empirical analysis, directly testing the correlation between $\varepsilon_t$ and $\eta_t$ is difficult. Instead, Hausman specification test of weak exogeneity is applied. As illustrated in case 2, Hausman specification (case 2 non weak exogeneity) can be viewed a special case of the case 3 non weak exogeneity.

**Summary of conditions for weak exogeneity**

a) Any lagged expected (or fitted) value from marginal model is not present in the conditional model, or any lagged expected (or fitted) value from conditional model is not present in marginal model.

b) The expected (or fitted) value from marginal model does not enter the conditional model

4 Since the nonlinear term can be linearized as $f(\psi_t) \approx (\psi_t - \overline{\psi}_t) f'(\overline{\psi}_t) + O_p(2)$. The proof thereafter is the same as is done in the Hausman test.

c) The errors from the marginal and conditional models are uncorrelated.

The violation of either one of the above conditions will result in non weak exogeneity.

# 4 Consequences of Incorrectly Assuming Weak Exogeneity- a Simulation Study

Based on the three cases of non weak exogeneity above, we will study the consequences of ignoring weak exogeneity in this section. We examine the consequences if one estimates the model under the assumption of weak exogeneity when there is none. To do so, we use a simulation study.

The experiments were designed as follows. The joint distribution of duration and volume is chosen as the benchmark model. The data is generated based on the fact that duration is not weak exogeneous. In particular, we generate the duration and volume data in accordance with each of the three cases of non weak exogeneity discussed in section 3.

**Case 1** The lagged expected (or fitted) value from the marginal model is present in the conditional model and the lagged expected (or fitted) value from the conditional model is present in the marginal model. (Eq. (7))

**Case 2** The expected (or fitted) value from the marginal model is present in the conditional model. (Eq. (11))

**Case 3** The errors from the marginal and conditional models are correlated. (Eq.(15))

We then estimate each model by two approaches. In the first approach, we assume the weak exogeneity condition is valid. The marginal process of duration and conditional process of volume given duration are estimated separately. We denote this estimation method the conditional MLE. In the second approach, we estimate the model under the fact of non weak exogeneity. The duration and volume processes are estimated jointly. We call this latter approach the full MLE. After estimation, we compare the estimation results of conditional MEL with those from the full MLE. In particular, we focus on a comparison of the bias/inconsistency and efficiency of the estimators.

We chose the sample sizes at N =2000, 5000, and 10000 respectively and the number of simulations S equals 2000. In the first two cases of non weak exogeneity, we use an exponential distribution with mean value 1 to generate the random disturbances $\varepsilon_t$ and $\eta_t$ individually. In the third case, we use a bivariate exponential distribution with correlations $\rho = 0.1$ and $\rho = 0.5$ to generate the random disturbances $\varepsilon_t$ and $\eta_t$ jointly. Table 1, Table 2, and Table 3 report the simulation results for the three cases.

Table 1: Case 1 simulation summary Statistics.    Estimated parameters

| | N=2000 | | | | N=5000 | | | | N=10000 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Conditional MLE | | Full MLE | | Conditional MLE | | Full MLE | | Conditional MLE | | Full MLE | |
| | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| $a_0$ | 0.132 | 0.048 | 0.106 | 0.032 | 0.125 | 0.026 | 0.102 | 0.020 | 0.124 | 0.018 | 0.101 | 0.013 |
| $a_1$ | 0.038 | 0.018 | 0.048 | 0.017 | 0.038 | 0.011 | 0.049 | 0.011 | 0.039 | 0.007 | 0.050 | 0.007 |
| $a_2$ | 0.054 | 0.013 | 0.051 | 0.012 | 0.053 | 0.008 | 0.051 | 0.007 | 0.053 | 0.006 | 0.051 | 0.005 |
| $a_3$ | 0.759 | 0.067 | 0.846 | 0.062 | 0.767 | 0.036 | 0.849 | 0.038 | 0.767 | 0.026 | 0.849 | 0.026 |
| $a_4$ | | | -0.050 | 0.035 | | | -0.051 | 0.020 | | | -0.050 | 0.014 |
| $b_0$ | 0.088 | 0.035 | 0.103 | 0.049 | 0.084 | 0.020 | 0.100 | 0.030 | 0.083 | 0.014 | 0.100 | 0.020 |
| $b_1$ | 0.057 | 0.039 | 0.052 | 0.036 | 0.056 | 0.024 | 0.051 | 0.022 | 0.055 | 0.017 | 0.050 | 0.015 |
| $b_2$ | 0.045 | 0.016 | 0.048 | 0.016 | 0.046 | 0.010 | 0.049 | 0.010 | 0.047 | 0.007 | 0.049 | 0.007 |
| $b_3$ | | | -0.046 | 0.098 | | | -0.047 | 0.061 | | | -0.048 | 0.041 |
| $b_4$ | 0.771 | 0.046 | 0.795 | 0.053 | 0.774 | 0.027 | 0.798 | 0.032 | 0.775 | 0.019 | 0.799 | 0.021 |
| $b_5$ | 0.100 | 0.033 | 0.101 | 0.032 | 0.100 | 0.021 | 0.100 | 0.020 | 0.100 | 0.015 | 0.100 | 0.014 |

Model:

$$d_t = \psi_t(\theta_d;\Omega_t)\varepsilon_t,\ \ \varepsilon_t \sim i.i.d(1,\sigma_\varepsilon^2)$$

$$v_t = \phi_t(\theta_v;d_t,\Omega_t)\eta_t,\ \ \eta_t \sim i.i.d(1,\sigma_\eta^2)$$

$$\psi_t = a_0 + a_1 d_{t-1} + a_2 v_{t-1} + a_3 \psi_{t-1} + a_4 \phi_{t-1}$$

$$\phi_t = b_0 + b_1 d_{t-1} + b_2 v_{t-1} + b_3 \psi_{t-1} + b_4 \phi_{t-1} +$$

The population parameter values;

$$a_0 = b_0 = 0.1,\ \ a_1 = b_1 = 0.05,\ a_2 = b_2 = 0.05$$

$$a_3 = 0.85,\ \ a_4 = -0.05,\ b_3 = -0.05,\ b_4 = 0.80,\ b_5 = 0.1$$

The parameters values are chosen partly from the empirical work of Manganelli (2005).

Table 2: Case 2 simulation summary statistics. Estimated parameters

| | N=2000 | | | | N=5000 | | | | N=10000 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Conditional MLE | | Full MLE | | Conditional MLE | | Full MLE | | Conditional MLE | | Full MLE | |
| | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| $a_0$ | 0.111 | 0.041 | 0.107 | 0.029 | 0.104 | 0.022 | 0.103 | 0.018 | 0.103 | 0.015 | 0.102 | 0.012 |
| $a_1$ | 0.049 | 0.014 | 0.049 | 0.013 | 0.049 | 0.009 | 0.049 | 0.008 | 0.050 | 0.006 | 0.049 | 0.005 |
| $a_2$ | 0.052 | 0.018 | 0.051 | 0.016 | 0.051 | 0.011 | 0.051 | 0.010 | 0.051 | 0.008 | 0.051 | 0.007 |
| $a_3$ | 0.842 | 0.038 | 0.846 | 0.026 | 0.847 | 0.022 | 0.848 | 0.017 | 0.848 | 0.015 | 0.849 | 0.012 |
| $b_0$ | 0.034 | 0.018 | 0.108 | 0.040 | 0.033 | 0.011 | 0.103 | 0.022 | 0.033 | 0.008 | 0.102 | 0.014 |
| $b_1$ | 0.062 | 0.022 | 0.051 | 0.024 | 0.061 | 0.014 | 0.051 | 0.014 | 0.061 | 0.010 | 0.050 | 0.010 |
| $b_2$ | 0.042 | 0.017 | 0.048 | 0.016 | 0.043 | 0.011 | 0.050 | 0.010 | 0.043 | 0.007 | 0.050 | 0.007 |
| $b_3$ | 0.715 | 0.033 | 0.802 | 0.049 | 0.716 | 0.020 | 0.799 | 0.028 | 0.716 | 0.014 | 0.800 | 0.019 |
| $b_4$ | 0.100 | 0.019 | 0.100 | 0.018 | 0.099 | 0.012 | 0.100 | 0.012 | 0.099 | 0.008 | 0.100 | 0.008 |
| $b_5$ | | | -0.107 | 0.048 | | | -0.101 | 0.027 | | | -0.101 | 0.017 |

Model:
$$d_t = \psi_t(\theta_d;\Omega_t)\varepsilon_t,\ \ \varepsilon_t \sim i.i.d(1,\sigma_\varepsilon^2)$$
$$v_t = \phi_t(\theta_v; d_t, \Omega_t)\eta_t,\ \ \eta_t \sim i.i.d(1,\sigma_\eta^2)$$

$$\psi_t = a_0 + a_1 d_{t-1} + a_2 v_{t-1} + a_3 \psi_{t-1}$$
$$\phi_t = b_0 + b_1 d_{t-1} + b_2 v_{t-1} + b_3 \phi_{t-1} + b_4 d + \psi$$

The population parameter values:

$$a_0 = b_0 = 0.1,\ \ a_1 = b_1 = 0.05,\ a_2 = b_2 = 0.05$$
$$a_3 = 0.85,\ b_3 = 0.80,\ b_4 = 0.1,\ b_5 = -0.1$$

The parameters values are chosen partly from the empirical work of Manganelli (2005).

Table 3: Case 3 simulation summary statistics . Estimated parameters (only conditional MLE is reported)

| | $\rho = 0.1$ | | | | | | $\rho = 0.5$ | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | N=2000 | | N=5000 | | N=10000 | | N=2000 | | N=5000 | | N=10000 | |
| | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| a0 | 0.108 | 0.033 | 0.104 | 0.021 | 0.102 | 0.014 | 0.110 | 0.030 | 0.103 | 0.018 | 0.102 | 0.012 |
| a1 | 0.048 | 0.014 | 0.049 | 0.009 | 0.050 | 0.006 | 0.048 | 0.016 | 0.049 | 0.010 | 0.050 | 0.007 |
| a2 | 0.051 | 0.012 | 0.050 | 0.007 | 0.050 | 0.005 | 0.051 | 0.012 | 0.051 | 0.008 | 0.050 | 0.005 |
| a3 | 0.847 | 0.025 | 0.848 | 0.015 | 0.849 | 0.011 | 0.847 | 0.022 | 0.849 | 0.014 | 0.849 | 0.009 |
| b0 | 0.090 | 0.041 | 0.086 | 0.025 | 0.084 | 0.018 | 0.050 | 0.067 | 0.041 | 0.044 | 0.036 | 0.018 |
| b1 | -0.098 | 0.042 | -0.101 | 0.026 | -0.102 | 0.019 | -0.515 | 0.096 | -0.528 | 0.059 | -0.534 | 0.036 |
| b2 | 0.040 | 0.015 | 0.041 | 0.009 | 0.041 | 0.007 | 0.021 | 0.018 | 0.019 | 0.012 | 0.019 | 0.008 |
| b3 | 0.790 | 0.035 | 0.792 | 0.021 | 0.793 | 0.015 | 0.749 | 0.088 | 0.765 | 0.053 | 0.772 | 0.028 |
| b4 | 0.279 | 0.037 | 0.279 | 0.023 | 0.280 | 0.017 | 0.796 | 0.049 | 0.794 | 0.031 | 0.794 | 0.022 |

Model:

$$d_t = \psi_t(\theta_d; \Omega_{t-1}) + \varepsilon_t, \quad v_t = \phi_t(\theta_v; d_t, \Omega_{t-1}) + \eta_t \qquad \begin{pmatrix} \varepsilon_t \\ \eta_t \end{pmatrix} \sim i.i.d.\, I(\Omega), \quad corr(\varepsilon_t, \eta_t) = \rho, \rho \neq 0$$

$$\psi_t = a_0 + a_1 d_{t-1} + a_2 v_{t-1} + a_3 \psi_{t-1}$$

$$\phi_t = b_0 + b_1 d_{t-1} + b_2 v_{t-1} + b_3 \phi_{t-1} + b_4\, l_t$$

The population parameter values:

$$a_0 = b_0 = 0.1, \quad a_1 = b_1 = 0.05, a_2 = b_2 = 0.05$$

$$a_3 = 0.85, b_3 = 0.80, b_4 = 0.1$$

The parameters values are chosen partly from the empirical work of Manganelli (2005).

From Table 1(the first case of non weak exogeneity), the means of the full MLE are all close to the population means. As the number of observations increases, the standard deviation of the full MLE gets smaller and the performance generally improves. In general, the conditional MELs are less efficient than those from the full MLEs, but the efficient loss is not very significant in most of cases. The full MLEs work well as a whole. On the other hand, the performance of the conditional MLE is somewhat different to that of the full MLE. In the duration process, both $a_1$ and $a_3$ are smaller than the population values. And the sum of $a_1$ and $a_3$ is downward biased towards smaller persistence when using the conditional MLE. The same result is also hold for the conditional distribution, where the sum of $b_1$ and $b_4$ is downward biased towards smaller persistence for volume. The conditional MLEs of $a_2$ and $b_1$ are larger than those from the full MLEs. It suggests the impact of duration on volume is over evaluated if case 1 of weak exogenetiy is ignored. However, the conditional MLE of $b_5$ is unbiased and consistent in this case, and the standard deviation is slightly bigger than that from full MEL. It can be seen that the same characteristics of the conditional MLEs continue to hold when N= 2000, 5000 and 10000. The poor performance of the conditional MLE seems to be due to the fact that the information from the marginal distribution contains some of the information of the conditional distribution.

From Table 2 (the second case of non weak exogeneity), we get similar results for the full MLE approach. The means of the full MLE are all close to the population values and the full MLE works well as a whole. The performance of the conditional MLE is different to that of the full MLE. For marginal distribution, the means of the conditional MLEs are unbiased and consistent in general. And the standard deviations of conditional MLEs are slightly larger than that of full MLEs, suggests an efficient gain when duration and volume are estimated jointly. For the conditional distribution, the sum of $b_2$ and $b_4$ is again downward biased towards smaller persistence for volume. However, the bias is even larger in this case. The conditional MLE of $b_1$ is greater than its population means, suggesting the impact of duration on volume is over estimated if case 2 of weak exogenetiy is ignored. And the conditional MLE of $b_5$ is unbiased and consistent. As the number of observations increases, the performance of the conditional MLEs generally improves. However, the same characteristics of conditional MLE continue to hold. It seems that when second weak exogeneity condition breaks down, the conditional MLEs for marginal distribution work fine, where the conditional MLEs for conditional distribution are biased.

Table 3(the third case of non exogeneity) only reports the results from conditional MLE. The full MLE requires the multivariate non-negative distribution, which is not directly available in the literature. We address this issue in our next Chapter. The conditional MLEs of the marginal distribution are unbiased and consistent in this case, even if the correlation between the marginal distribution and conditional distribution is high. The means of conditional MLEs are all close to the population means. In the conditional distribution, the conditional MLE of $b_3$ is unbiased and consistent when the correlation of errors between the marginal distribution and the conditional distribution is relatively small ($\rho = 0.1$), and it gets slightly biased and inconsistent when the correlation is relatively high ($\rho = 0.5$). The great differences are observed for the conditional MLEs of $b_1$ and $b_4$, which evaluate the impact of duration on volume. It can be seen that the conditional MLE of $b_1$ is negative in this case, and the negative size increases drastically as the correlation of the errors increases. The conditional MLE of $b_4$ is much larger that its population mean. As the correlation of the errors increases, the conditional MLE of $b_4$ gets larger. Thus, the incorrectly assuming case 3 weak exogeneity has severe consequences on the estimation results, which makes the inferences on the parameters unreliable or even misleading.

To summarize, incorrectly assuming weak exogeneity has particularly effects on the conditional distribution, where the persistence of volume will be biased and the impact of duration on volume will be over (or under) evaluated. The bias is very large in case 3 non-weak exogeneity, which makes the econometric inferences on the parameters unreliable or even misleading. In addition, the failure of the weak exogeneity implies that conditional MLEs are inefficient although the efficiency loss is relatively small. This result is consistent with White (1981, 1982). It is therefore necessary to conduct a test for weak exogeneity before estimation in empirical analysis.

# 5 An LM Test for Weak Exogeneity in Financial Point Processes

In this section, we will derive a Langrage-multiplier (LM) or efficient score test for weak exogeneity. It proves to be particularly useful since it only requires estimation of the restricted model.

In test of weak exogeneity, correct specification of the conditional mean is a fundamental assumption for the validity of the test since the rejection of the null hypothesis could be due to either the rejection of weak exogeneity or the result of misspecification of the conditional mean. To take account of this we introduce an Augmented ACD (AACD) model (Fernandes and Grammig 2006) for the specification of the conditional mean of duration and volume. The AACD model of Fernandes and Grammig (2006) is given by

$$d_t = \psi_t(\theta_d;\Omega_t)\varepsilon_t$$

where $\varepsilon_t$ is i.i.d with mean value 1, and

$$\frac{\psi_t^{\lambda}-1}{\lambda} = \omega^* + \alpha^* \varphi_{t-1}^{\lambda}[|\varepsilon_{t-1}-b| + c(\varepsilon_{t-1}-b)]^{v} + \beta\frac{\psi_{t-1}^{\lambda}-1}{\lambda}.$$

The AACD model then obtained by rewriting as

$$\psi_t^{\lambda} = \omega + \alpha\phi_{t-1}^{\lambda}[|\varepsilon_{t-1}-b| + c(\varepsilon_{t-1}-b)]^{v} + \beta\psi_{t-1}^{\lambda} \quad (16)$$

where $\omega = \lambda\omega^* - \beta + 1$ and $\alpha = \lambda\alpha^*$

The AACD model provides a flexible functional form and permits the conditional duration process $\{\psi_t\}$ to respond in distinct manners to small and large shocks. The shock impact curve $g(\varepsilon_i) = [|\varepsilon_{t-1}-b| + c(\varepsilon_{t-1}-b)]^{v}$ incorporates such asymmetric responses through the shift and rotation parameters b and c, respectively. The shape parameter $v$ plays a similar role to $\lambda$, which determines whether the Box-Cox transformation is concave ($\lambda \le 1$) or convex ($\lambda \ge 1$).

Appendix 3 summarizes the typology of ACD models which can be nested by the AACD model. Since the AACD model provides a flexible functional form and encompasses most of the current ACD models, the rejection of the null is less likely to be due to misspecification of the conditional mean. Because of the inherent complexity of the AACD model, the LM or efficient score testing principle proved to be particularly useful for this purpose, since it requires estimation under the null hypothesis only.

As illustrated in section 3, Hausman specification (second case of non weak exogeneity) is a special case of the case 3 non weak exogeneity. The rejection of case two weak exogeneity is also the rejection of case three weak exogeneity. To illustrate the test principle, only second case of non weak exogeneity is discussed in the LM test and in empirical analysis. The first case of non weak exogeneity can be derived in the same way.

Let us specify the duration and volume as represented by the AACD and Augmented ACV (AACV) models respectively with the errors belonging to the exponential distribution family (exponential, Burr or Weibull distribution); for example

$$\begin{aligned} d_t &= \psi_t(\theta_d;\Omega_t)\varepsilon_t \\ \psi_t^\lambda &= \omega_1 + \alpha_1\psi_{t-1}^\lambda[|\varepsilon_{t-1} - b_1| + c_1(\varepsilon_{t-1} - b_1)]^\nu + \beta_1\psi_{t-1}^\lambda \\ v_t &= \varphi_t(\theta_v; d_t, \Omega_t)\eta_t \\ \varphi_t^\lambda &= \omega_2 + \alpha_2\varphi_{t-1}^\lambda[|\eta_{t-1} - b_2| + c_2(\eta_{t-1} - b_2)]^\nu + \beta_2\varphi_{t-1}^\lambda + a_0 d_t + a_1\psi_t. \end{aligned} \tag{17}$$

As explained in section 3, it suffices to test $H_0 \;: a_1 = 0$ in order to test for weak exogeneity of duration. In such a case, the parameters of interest are not subject to cross equation restrictions and are variation free with parameters from marginal model.

As has been noted, the LM test is particularly useful for this purpose, since it requires estimation under the null hypothesis only.

Under $H_0$

$$\begin{aligned} v_t &= \varphi_t(\theta_v; d_t, \Omega_t)\eta_t \\ \varphi_t^\lambda &= \omega_2 + \alpha_2\varphi_{t-1}^\lambda[|\eta_{t-1} - b_2| + c_2(\eta_{t-1} - b_2)]^\nu + \beta_2\varphi_{t-1}^\lambda + a_0 d_t. \end{aligned} \tag{18}$$

Under $H_1$

$$\begin{aligned} v_t &= \varphi_t(\theta_v; d_t, \Omega_t)\eta_t \\ \varphi_t^\lambda &= \omega_2 + \alpha_2\varphi_{t-1}^\lambda[|\eta_{t-1} - b_2| + c_2(\eta_{t-1} - b_2)]^\nu + \beta_2\varphi_{t-1}^\lambda + a_0 d_t + a_1\psi_t. \end{aligned} \tag{19}$$

Assuming that the densities are correct, the general theory of ML leads to a simple score test for $a_1 = 0$.Given correctly specified duration and volume models, the quasi log-likelihood function is

$$L = -\sum_{t=1}^{T}(\log\varphi_t + v_t/\varphi_t + \log\psi_t + d_t/\psi_t). \tag{20}$$

The quasi log-likelihood MLE approach is most suitable since it allows for a wide range of different distributions capturing all possible supports of the point process.

Moreover, under $H_0$ of weak exogeneity, the AACD and AACV model can be estimated separately. Then, the score/LM test has the familiar form

$$S = -\hat{i}^{a_1}(\hat{\theta}_c)' \hat{I}^{\alpha_1}(\hat{\theta}_c)^{-1} \hat{i}^{a_1}(\hat{\theta}_c) \tag{21}$$

where $\hat{i}^{a_1}(\hat{\theta}_c) = \frac{\partial L}{\partial \theta}$ and $\hat{I}^{\alpha_1}(\hat{\theta}_c) = \frac{\partial L}{\partial \theta \partial \theta'}$ are the components corresponding to $a_1$ in the empirical score and Hessian from constrained model. Under mild regularity conditions it is well known that, the score test has an asymptotically $\chi^2(1)$ distribution under $H_0$.

### *5.1 Testing Joint Weak Exogeneity*

The above testing approach enables a test of weak exogeneity of duration for one market mark (volume or volatility). In market microstructure theory, sometimes there are more than two market variables (duration, volume, volatility, bid-ask spread, et.al) of interest which need to be modelled jointly. In such a case, a joint weak exogeneity test is necessary. We propose the joint weak exogeneity test principle in this section. We take the model from Manganelli (2005), where duration, volume and volatility are modelled jointly. But the methods can be extended to other multivariate frameworks.

$$(d_t, v_t, r_t) \sim f(d_t, v_t, r_t | \Omega_t; \theta) = g(d_t | \Omega_t; \theta^d).h(v_t | d_t, \Omega_t; \theta^v).k(r_t | d_t, v_t, \Omega_t; \theta^r). \tag{22}$$

The log likelihood can be expressed as:

$$L(\theta^d, \theta^v, \theta^r) = \sum_{t=1}^{T} [\log g(d_t | \Omega_t; \theta^d) + \log h(v_t | d_t, \Omega_t; \theta^v) + \log k(r_t | d_t, v_t, \Omega_t; \theta^r)]. \tag{23}$$

As illustrated before, we allow for a more flexible functional form for duration, volume and volatility process. This results in a LM score test. The conditional duration and volume are assumed to follow an AACD and AACV process and the conditional volatility are assumed to follow an Asymmetric Power APGARCH process (Ding, Granger et al. 1993), which is similar to the AACD specification. Then the volatility model has the following form:

$$\begin{aligned} \mu_t &= \sigma_t(\theta_v; d_t, v_t, \Omega_t)\xi_t, \ \ \xi_t \sim i.i.d(0,1) \\ \sigma_t^{\lambda} &= \omega_3 + \alpha_3 \sigma_{t-1}^{\lambda}[|\xi_{t-1} - b_3| + c_3(\xi_{t-1} - b_3)]^{v} + \beta_3 \sigma_{t-1}^{\lambda} + a_{01} d_t + a_{11}\psi_t + a_{02} v_t + a_{12}\varphi_t. \end{aligned} \tag{24}$$

For the same reason, it suffices to test that $\alpha_{11} = \alpha_{12} = 0$ in order to test jointly the weak exogeneity of duration and volume. So under the null hypothesis, the constrained model is:

Under $H_0$

$$\begin{aligned} &\mu_t = \sigma_t(\theta_v; d_t, v_t, \Omega_t)\xi_t,\ \ \xi_t \sim i.i.d(0,1) \\ &\sigma_t^{\lambda} = \omega_3 + \alpha_3\sigma_{t-1}^{\lambda}[\left|\xi_{t-1} - b_3\right| + c_3(\xi_{t-1} - b_3)]^{v} + \beta_3\sigma_{t-1}^{\lambda} + a_{01}d_t + a_{02}v_t. \end{aligned} \tag{25}$$

The density for both the AACD and AACV models is the exponential while for the APGARCH model it is a standard normal distribution. Assuming that the densities are correct, the general theory of ML leads to a simple Score test for $\alpha_{11} = \alpha_{12} = 0$. Given that the Augmented GARCH, AACV and AACD models are correctly specified, the quasi log-likelihood function is

$$L = -\sum_{t=1}^{T}(\log\varphi_t + v_t/\varphi_t + \log\psi_t + d_t/\psi_t + \log\sigma_t^2 + \mu_t^2/\sigma_t^2). \tag{26}$$

Moreover, under $H_0$ of weak exogeneity, the APGARCH and the AACD and AACV models can be estimated separately. The score/LM test has the familiar form

$$S = -\hat{i}^{a}(\widehat{\theta}_c)'\widehat{I}^{\alpha}(\widehat{\theta}_c)^{-1}\hat{i}^{a}(\widehat{\theta}_c) \tag{27}$$

where $\hat{i}^{a}(\widehat{\theta}_c) = \dfrac{\partial L}{\partial \theta}$ and $\widehat{I}^{\alpha}(\widehat{\theta}_c) = \dfrac{\partial L}{\partial\theta\partial\theta'}$ are, respectively, the components of the empirical score and Hessian from unconstrained model corresponding to $\alpha_{11}$ .and $\alpha_{12}$. Under mild regularity conditions it is well known that, the score test has an asymptotic $\chi^2(2)$ distribution under $H_0$.

## *5.2 Power of the test*

How powerful is this test for weak exogeneity? How many observations do we need to have for this test? To answer these questions, we need to conduct an investigation of the statistical power of the test. We begin with a simple model below:

$$\begin{aligned} &d_t = \psi_t(\theta_d; \Omega_t)\varepsilon_t,\ \ \varepsilon_t \sim i.i.d(1,\sigma_\varepsilon^2) \\ &v_t = \phi_t(\theta_v; d_t, \Omega_t)\eta_t,\ \ \eta_t \sim i.i.d(1,\sigma_\eta^2) \\ &\psi_t = a_0 + \alpha_1 d_{t-1} + a_2\psi_{t-1} \\ &\phi_t = b_0 + b_1 v_{t-1} + b_2\phi_{t-1} + b_3 d_t + b_4\psi_t \\ &H_0:\ b_4 = 0 \\ &H_1:\ b_4 \neq 0 \end{aligned} \tag{28}$$

Choosing a 5% significance level, the simulation results indicate that the empirical significance level is 6.7% for sample size n=10000, 7.2% for sample size n=5000 and 8.9% for sample size n=2000.

To explore the power of the test, we generate data under the alternative hypothesis and estimate the model under null hypothesis[5]. Under the alternative hypothesis the parameter $b_4$ varies between -0.2 to 0.2 with step 0.025. Given the sample size and empirical test size, the power of the test is the probability of rejecting a hypothesis when it is false. The results of the LM test for different sample sizes and empirical significance level are listed in Table 4.

Table 4: Percentage rejections of the LM tests at empirical significance level for testing $b_4 = 0$ against $b_4 \neq 0$

| $b_4$ | Power of test | | |
|---|---|---|---|
| | N=2000 | N=5000 | N=10000 |
| -0.200 | 0.825 | 0.989 | 1.000 |
| -0.175 | 0.736 | 0.977 | 1.000 |
| -0.150 | 0.614 | 0.933 | 0.999 |
| -0.125 | 0.457 | 0.803 | 0.972 |
| -0.100 | 0.354 | 0.610 | 0.878 |
| -0.075 | 0.232 | 0.403 | 0.600 |
| -0.050 | 0.149 | 0.210 | 0.347 |
| -0.025 | 0.084 | 0.093 | 0.141 |
| 0.000 | 0.052 | 0.051 | 0.051 |
| 0.025 | 0.041 | 0.057 | 0.098 |
| 0.050 | 0.054 | 0.120 | 0.236 |
| 0.075 | 0.105 | 0.277 | 0.502 |
| 0.100 | 0.164 | 0.458 | 0.781 |
| 0.125 | 0.260 | 0.684 | 0.938 |
| 0.150 | 0.414 | 0.848 | 0.987 |
| 0.175 | 0.547 | 0.924 | 0.997 |
| 0.200 | 0.642 | 0.970 | 1.000 |

As the sample size increases, the power of LM test increases. It can also be seen when $b_4$ decreases to 0, the power tends to be 5%. The test power grows quickly to 1 as $b_4$ move away from zero. They are plotted in Figure 1. They are appropriately symmetric. The simulation shows that the LM test has good power to test for weak exogeneity in a financial market point process.

[5] To avoid present of negative value of volume, we use logarithmic version of ACD model (see e.g., Bauwens and Giot 2000) for DGP process and estimation.

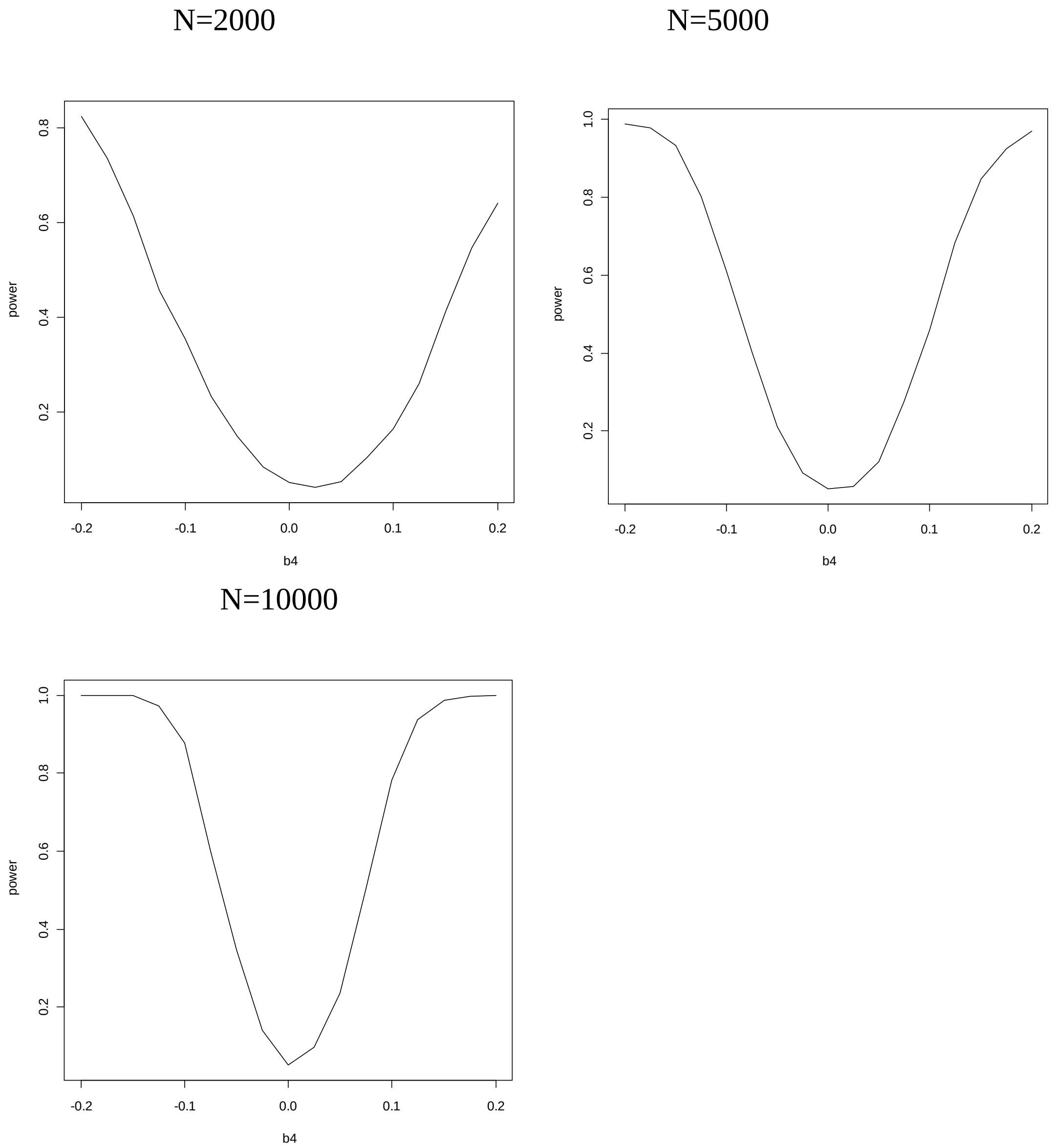


Figure 1: Power of the test

# 6 Empirical Analysis

In this section, we use the method discussed in section 5 to test weak exogeneity of duration for two groups of high frequency data. The empirical analysis starts with the joint distribution of the three variables: duration, volume and return volatility. These three variables are key factors in analysing market microstructure. Specifically, we will test weak exogeneity of duration for the conditional distribution of volume and volatility. We also test the joint weak exogeneity of duration and volume for the conditional distribution of volatility.

## *6.1 Data*

We use the data from the Trades and Quotes (TAQ) dataset at NYSE. The TAQ data consists of two parts: the first reports the trade data, while the second lists the quote data (bid and ask data) posted by the market maker. The data were kindly provided by Manganelli (2005). He constructed 10 deciles of stocks covering the period from Jan 1,1998 to June 30, 1999, on the basis of the 1997 total number of trades of all stocks quoted on the NYSE. We randomly selected 5 stocks from the eighth decile (frequently traded stocks) and 5 from the second decile (infrequently traded stocks) covering the period from Jan 1,1998 to June 30, 1999. The tickers and names of the ten stocks are reported in Table 5.

Before the analysis began, we adopted Manganelli (2005)'s strategy to prepare the data. First, all trades before 9:30 am or after 4:00 pm were discarded. Second, durations over night were computed as if the overnight periods did not exist. For example, the time elapsing between 15:59:50 and 9:30:05 of the following day is only 15 seconds. We keep overnight duration because our samples for infrequently traded stocks are very small. Eliminating this duration would cause the loss of important data for these stocks. Third, all transaction data with zero duration are eliminated. These transactions are treated as one single transaction, and the related volumes are summed. Fourth, to deal with the impact of dividend payments and trading halts, we simply deleted the first observation whose price incorporated the dividend payment or a trading halt. Fifth, to adjust the data for stock splits, we simply multiplied the price and volume by the stock split ratio. Sixth, the price of each transaction is calculated as the average of the prevailing bid and ask quote. To obtain the prevailing quotes, we use the 5 second rule used by Lee and Ready (1991) which liniks each trade to the quote posted at least 5 seconds before , since the quotes can be posted more quickly than trades are recorded. This procedure is standard in microstructure studies. Seventh, the returns were computed as the difference of the log of the prices. To obtain a return sequence that is free of the bid-ask bounce that affects prices (see Campbell et al., 1997, chapter 3), we follow Ghysels, Gourieroux et al. (1998) in using the residuals of an ARMA(1,1) model estimated on the return data.

Table 5: Stock used in this analysis

| A. Frequently traded | | B. Infrequently traded | |
|---|---|---|---|
| TRN | TRINITY INDUSTRIES | ABG | GROUPE AB S A ADS |
| R | RYDER SYSTEM INC | OFG | ORIENATAL FINL GRP HOLD CO |
| ARG | AIRGAS INC | LSB | LSB INDUSTRIES INC |
| FMO | FEDERAL-MOGUL CORP | HTD | HUNTINGDON LIFE S.G. |
| VTS | VERITAS DGC INC | HUN | HUNT CORP |

The second issue to be addressed prior to the analysis concerns the intraday pattern in the data. It is well known that duration, volume and volatility exhibit strong intraday periodic components, with a high trading activity at the beginning and end of the day. To adjust for this, we make use of a method used by Engle (2000). We regress the durations, volumes and returns squares on a piecewise cubic spline with knots at 9:30, 10:00, 11:00, 12:00, 13:00, 14:00, 15:00, 15:30 and 16:00. The original series are then divided by the spline forecast to obtain the adjusted series. **Error! Reference source not found.** depicts the nonparametric estimate of daily pattern of duration and return square for one typical stock ARG. Generally, less frequently traded stocks do not exhibit any regular intraday pattern. More frequently traded stocks typically show the inverted U pattern for duration, the L pattern for return squares, and no regular pattern for volume.

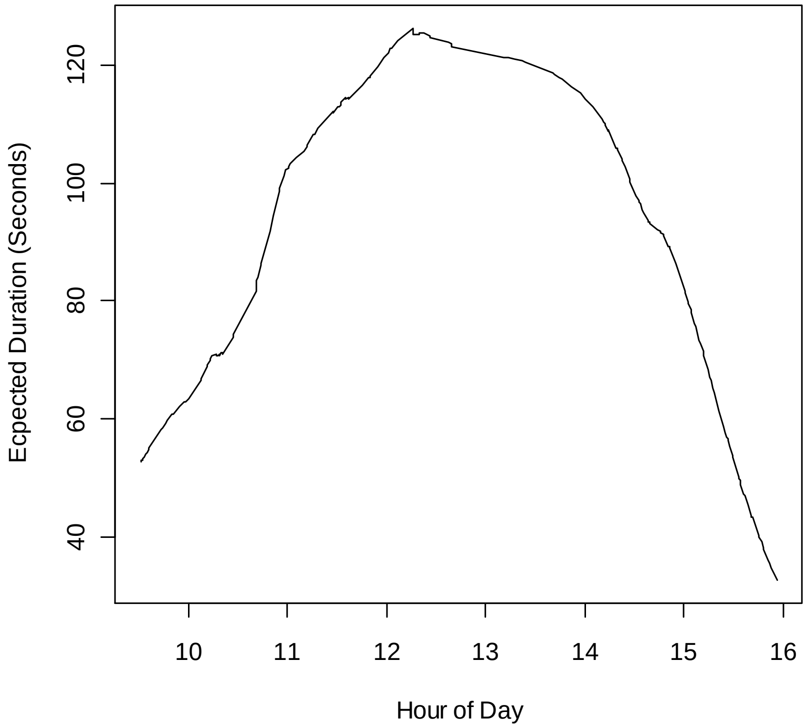


Figure 2: Nonparametric estimate of daily pattern of transaction durations.

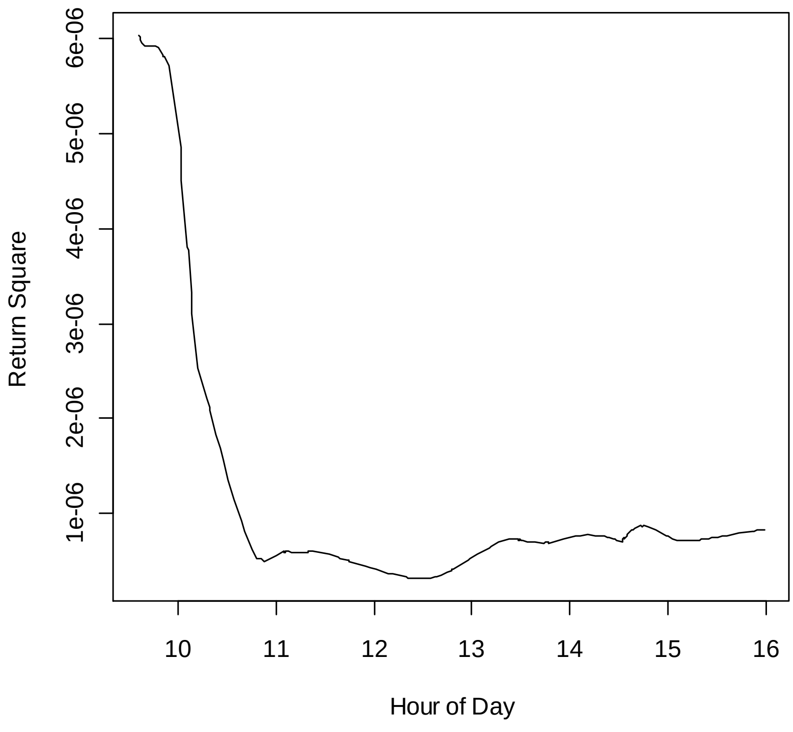


Figure 3: Nonparametric estimate of daily pattern of return square.

Table 6: Summary statistics for the 10 stocks.

| | Obs | Mean | | LB(20) | | |
|---|---|---|---|---|---|---|
| | | Duration | Volume | Duration | Volume | Variance |
| TRN | 55582 | 157.86 | 1369.43 | 3780.09 | 1383.35 | 3769.80 |
| GAS | 101332 | 86.54 | 3118.12 | 5951.85 | 2338.08 | 4073.09 |
| TCB | 55208 | 158.94 | 1855.20 | 4171.36 | 2644.11 | 2925.82 |
| R | 69702 | 125.67 | 2492.98 | 14072.3 | 7276.91 | 23685.7 |
| ARG | 33850 | 259.2 | 1280.70 | 3780.09 | 1383.35 | 3769.80 |
| ABG | 2074 | 4214.88 | 5259.05 | 120.28 | 225.07 | 146.00 |
| OFG | 7212 | 1214.58 | 833.86 | 523.16 | 1343.43 | 738.09 |
| LSB | 2962 | 2962.19 | 1971.61 | 481.41 | 435.69 | 523.58 |
| HUN | 5887 | 1483.73 | 1070.02 | 2431.00 | 660.60 | 788.81 |
| JNS | 3949 | 2215.94 | 2748.60 | 268.52 | 682.92 | 297.01 |

## *6.2 Testing for weak exogeneity - empirical results*

Table 7 reports the LM test statistics. The first and second rows are the LM statistics for weak exogeneity of duration in conditional distribution of volume and volatility respectively. The third row is the LM statistics for jointly weak exogeneity of duration and volume in conditional distribution of volatility.

Table 7: Weak Exogeneity Test -- LM Test Statistics

| | TRN | R | ARG | VTS | FMO | ABG | OFG | LSB | HUN | HTD |
|---|---|---|---|---|---|---|---|---|---|---|
| Volume | 2.78 | 20.2 | 15.8 | 74.5 | >100 | 0.72 | 2.00 | 4.20 | 8.91 | 3.02 |
| Volatility | >100 | >100 | 0.51 | 9.81 | >100 | >100 | >100 | 2.14 | 41.0 | 3.78 |
| Volatility-J | 31.9 | >100 | >100 | >100 | >100 | >100 | >100 | >100 | >100 | >100 |

Note: Critical values $\chi^2(1)_{0.05}$=3.84, $\chi^2(2)_{0.05}$=5.99

$\chi^2(1)_{0.01}$=6.63, $\chi^2(2)_{0.01}$=9.21

Let's look at the frequently traded stocks. In both the volume equation and the volatility equation, the null hypothesises that duration is weakly exogenous are rejected in 4 out of 5 cases, while the null hypothesis that duration and volume are jointly weakly exogenous is rejected for all the stocks in the volatility equation. A different picture emerges for infrequently traded stocks. In the volume equation, the null of weak exogeneity of duration is not rejected in 4 out of 5 cases (under 1% level). And in the volatility equation, the null is not rejected for 2 out of 5 stocks. The joint weak exogeneity of duration and volume, on the other hand, is again rejected in all the 5 cases. The different results found for frequently traded stocks compared to infrequently traded stocks are striking. In general, the null of weak exogeneity is rejected for frequently traded stocks, while it is less likely to be rejected for infrequently traded stocks. However, the joint weak exogeneity of duration and volume is rejected in both of the cases (Manganelli 2005).

Our LM test results indicate that that the empirical model of Engle (2000) and Manganelli (2005)on market microstructure analysis, in which duration and marks are estimated separately, may only be suitable for infrequently traded stocks. It is more efficient to estimate duration, volume, and price volatility jointly for frequently traded stocks.

# 7 Conclusion

A common practice when modelling several financial point processes jointly is to factor the joint density into the product marginal density of duration and conditional density of marks given duration. In estimation, the assumption of weak exogeniety of duration is made in order to estimate the marginal density and conditional density separately. This paper analyses the issues related to weak exogeneity in financial point processes. We propose three cases of non weak exogeneity, which extends the application of the Hausman test of weak exogeneity to a time series model. We then do a simulation to study the consequences of ignoring the weak exogeneity in estimation. We find that incorrectly assuming weak exogeneity implied biased estimators. Particularly, the persistence of volume will be biased and the impact of duration on volume will be over (or under) evaluated. The bias is very large in case 3 non-weak exogeneity, which makes the econometric inferences on the parameters unreliable or even misleading.

In empirical analysis, we derive a test for weak exogeneity based on LM test principles. The LM test is attractive because it only requires estimation of the restricted model. A simulation study suggests that the LM test has good power. We apply the method to two groups of high frequency data. The empirical results indicate that weak exogeneity of duration is often rejected for frequently traded stocks, but is less likely to be rejected for infrequently traded stocks.

# Reference

Bauwens, L. and P. Giot (2000). "The logarithmic ACD model: an application to the bid-ask quote process of three NYSE stocks." Annales d'Economie et de Statistique: 117-149.

Boswijk, H. P. and J. P. Urbain (1997). "Lagrance-multiplier tersts for weak exogeneity: a synthesis." Econometric Reviews **16**(1): 21-38.

Chmelarova, V. (2007). The Hausman test, and some alternatives, with heteroskedastic data, Louisiana State University, USA.

Darrat, A. F., M. K. Hsu, et al. (2000). "Testing export exogeneity in Taiwan: further evidence." Applied Economics Letters **7**(9): 563-567.

Davidson, R. and J. G. MacKinnon (2004). "Econometric Theory and Methods." Oxford University Press, New York.

Ding, Z., C. W. J. Granger, et al. (1993). "A long memory property of stock market returns and a new model." Journal of empirical finance **1**(1): 83-106.

Dolado, J., J. M. Rodriguez-Poo, et al. (2004). "Testing Weak Exogeneity in the Exponential Family: An Application to Financial Marked-Point Processes." CORE Discussion Paper No. 2004/49.

Dufour, A. and R. F. Engle (2000). "Time and the price impact of a trade." The Journal of Finance **55**(6): 2467-2498.

Engle, R. (2002). "New frontiers for ARCH models." Journal of Applied Econometrics **17**(5): 425-446.

Engle, R. F. (2000). "The Econometrics of Ultra‐high‐frequency Data." Econometrica **68**(1): 1-22.

Engle, R. F. and D. F. Hendry (1993). "Testing superexogeneity and invariance in regression models." Journal of econometrics **56**(1): 119-139.

Engle, R. F., D. F. Hendry, et al. (1983). "Exogeneity." Econometrica: Journal of the Econometric Society: 277-304.

Engle, R. F. and J. R. Russell (1998). "Autoregressive conditional duration: A new model for irregularly spaced transaction data." Econometrica: 1127-1162.

Engle, R. F. and Z. Sun (2007). "When is noise not noise–a microstructure estimate of realized volatility." Manuscript, Stern School, New York University **38**.

Fernandes, M. and J. Grammig (2006). "A family of autoregressive conditional duration models." Journal of econometrics **130**(1): 1-23.

Fischer, A. M. (1993). "Is money really exogenous? Testing for weak exogeneity in Swiss money demand." Journal of Money, Credit and Banking **25**(2): 248-258.

Ghysels, E., C. Gourieroux, et al. (1998). "Stochastic volatility duration models."

Hausman, J. A. (1978). "Specification tests in econometrics." Econometrica: Journal of the Econometric Society: 1251-1271.

Lee, C. M. C. and M. J. Ready (1991). "Inferring trade direction from intraday data." Journal of Finance: 733-746.

Manganelli, S. (2005). "Duration, volume and volatility impact of trades." Journal of Financial Markets **8**(4): 377-399.

Staub, K. E. (2009). "Simple tests for exogeneity of a binary explanatory variable in count data regression models." Communications in Statistics—Simulation and Computation® **38**(9): 1834-1855.

Terza, J. V., A. Basu, et al. (2008). "Two-stage residual inclusion estimation: addressing endogeneity in health econometric modeling." Journal of health economics **27**(3): 531-543.

White, H. (1981). "Consequences and detection of misspecified nonlinear regression models." Journal of the American Statistical Association: 419-433.

White, H. (1982). "Maximum likelihood estimation of misspecified models." Econometrica: Journal of the Econometric Society: 1-25.

Wu, D. M. (1973). "Alternative tests of independence between stochastic regressors and disturbances." Econometrica: Journal of the Econometric Society: 733-750.

Appendix

## *Appendix 1: Proof of Case 1 weak exogeneity*

Using matrix form

$$\begin{pmatrix}\psi_t\\ \varphi_t\end{pmatrix}=\begin{pmatrix}a_0\\ b_0\end{pmatrix}+\begin{pmatrix}a_1 & a_2\\ b_1 & b_2\end{pmatrix}\begin{pmatrix}d_{t-1}\\ v_{t-1}\end{pmatrix}+\begin{pmatrix}a_3 & a_4\\ b_3 & b_4\end{pmatrix}\begin{pmatrix}\psi_{t-1}\\ \varphi_{t-1}\end{pmatrix}+\begin{pmatrix}0 & 0\\ b_5 & 0\end{pmatrix}\begin{pmatrix}d_t\\ v_t\end{pmatrix}$$

Using the same method, the above model can be transformed into:

$$\begin{pmatrix}d_t\\ v_t\end{pmatrix}=\begin{pmatrix}a_0\\ b_0\end{pmatrix}+\begin{pmatrix}a_1 & a_2\\ b_1 & b_2\end{pmatrix}\begin{pmatrix}d_{t-1}\\ v_{t-1}\end{pmatrix}+\begin{pmatrix}a_3 & a_4\\ b_3 & b_4\end{pmatrix}\begin{pmatrix}\psi_{t-1}\\ \varphi_{t-1}\end{pmatrix}+\begin{pmatrix}0 & 0\\ b_5 & 0\end{pmatrix}\begin{pmatrix}d_t\\ v_t\end{pmatrix}+\begin{pmatrix}d_t-\psi_t\\ v_t-\varphi_t\end{pmatrix}$$

If we assume $\begin{matrix}d_t-\psi_t=\varepsilon_{1t}\sim i.i.d(0,\tilde{\sigma}_\varepsilon^2)\\ v_t-\phi_t=\varepsilon_{2t}\sim i.i.d(0,\tilde{\sigma}_\eta^2)\end{matrix}$, it becomes

$$\begin{pmatrix} d_t \\ v_t \end{pmatrix} = \begin{pmatrix} a_0 \\ b_0 \end{pmatrix} + \begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix} \begin{pmatrix} d_{t-1} \\ v_{t-1} \end{pmatrix} + \begin{pmatrix} a_3 & a_4 \\ b_3 & b_4 \end{pmatrix} \begin{pmatrix} d_{t-1} - \varepsilon_{1t-1} \\ v_{t-1} - \varepsilon_{2t-1} \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ b_5 & 0 \end{pmatrix} \begin{pmatrix} d_t \\ v_t \end{pmatrix} + + \begin{pmatrix} \varepsilon_{it} \\ \varepsilon_{2t} \end{pmatrix}$$

$$= \begin{pmatrix} a_0 \\ b_0 \end{pmatrix} + \begin{pmatrix} a_1 + \alpha_3 & a_2 + \alpha_4 \\ b_1 + b_3 & b_2 + b_4 \end{pmatrix} \begin{pmatrix} d_{t-1} \\ v_{t-1} \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ b_5 & 0 \end{pmatrix} \begin{pmatrix} d_t \\ v_t \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix} - \begin{pmatrix} a_3 & a_4 \\ b_3 & b_4 \end{pmatrix} \begin{pmatrix} \varepsilon_{1t-1} \\ \varepsilon_{2t-1} \end{pmatrix}$$

## *Appendix 2:* ***Proof of Hausman test***

$$d_t = \psi_t(\theta_d; \Omega_t)\varepsilon_t$$

$$\psi_t = a_0 + a_1 d_{t-1} + a_2 v_{t-1} + \alpha_3 \psi_{t-1}$$

$$v_t = \varphi_t(\theta_v; d_t, \Omega_t)\eta_t$$

$$\phi_t = b_0 + b_1 d_{t-1} + b_2 v_{t-1} + b_3 \phi_{t-1} + b_4 d_t + b_5 \psi_t$$

Implying

$$d_t = \alpha_0 + \alpha_1 d_{t-1} + a_2 v_{t-1} + \alpha_3 \psi_{t-1} + (d_t - \psi_t)$$

$$= \alpha_0 + \alpha_1 d_{t-1} + a_2 v_{t-1} + \alpha_3 \psi_{t-1} + \varepsilon_{1t}$$

$$v_t = b_o + b_1 d_{t-1} + b_2 v_{t-1} + b_3 \phi_{t-1} + b_4 d_t + b_5 \psi_t + (v_t - \phi_t)$$

$$= b_o + b_1 d_{t-1} + b_2 v_{t-1} + b_3 \phi_{t-1} + b_4 d_t + b_5 (d_t - \varepsilon_{1t}) + \varepsilon_{2t}$$

$$= b_o + b_1 d_{t-1} + b_2 v_{t-1} + b_3 \phi_{t-1} + (b_4 + b_5) d_t - b_5 \varepsilon_{1t} + \varepsilon_{2t}$$

$$= b_o + b_1 d_{t-1} + b_2 v_{t-1} + b_3 \phi_{t-1} + (b_4 + b_5) d_t + \varepsilon'_{2t}$$

Using matrix form for the two equation

$$\begin{pmatrix} d_t \\ v_t \end{pmatrix} = \begin{pmatrix} a_0 \\ b_0 \end{pmatrix} + \begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix} \begin{pmatrix} d_{t-1} \\ v_{t-1} \end{pmatrix} + \begin{pmatrix} a_3 & 0 \\ 0 & b_3 \end{pmatrix} \begin{pmatrix} \psi_{t-1} \\ \phi_{t-1} \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ b_4 + b_5 & 0 \end{pmatrix} \begin{pmatrix} d_t \\ v_t \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ -b_5 & 1 \end{pmatrix} \begin{pmatrix} \varepsilon_{it} \\ \varepsilon_{2t} \end{pmatrix}$$

$$= \begin{pmatrix} a_0 \\ b_0 \end{pmatrix} + \begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix} \begin{pmatrix} d_{t-1} \\ v_{t-1} \end{pmatrix} + \begin{pmatrix} a_3 & 0 \\ 0 & b_3 \end{pmatrix} \begin{pmatrix} \psi_{t-1} \\ \phi_{t-1} \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ b_4 + b_5 & 0 \end{pmatrix} \begin{pmatrix} d_t \\ v_t \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon'_{2t} \end{pmatrix}$$

*Appendix 3: Typology of ACD models*

| |
|---|
| Augmented ACD |
| $\psi_t^{\lambda} = \omega + \alpha\phi_{t-1}^{\lambda}[\lvert\varepsilon_{t-1} - b\rvert + c(\varepsilon_{t-1} - b)]^{v} + \beta\psi_{t-1}^{\lambda}$ |
| Asymmetric power ACD ( $\lambda = v$ ) |
| $\psi_t^{\lambda} = \omega + \alpha\phi_{t-1}^{\lambda}[\lvert\varepsilon_{t-1} - b\rvert + c(\varepsilon_{t-1} - b)]^{\lambda} + \beta\psi_{t-1}^{\lambda}$ |
| Asymmetric logarithmic ACD ( $\lambda \to 0$ and $v = 1$ ) |
| $\log\psi_t = \omega + \alpha[\lvert\varepsilon_{t-1} - b\rvert + c(\varepsilon_{t-1} - b)]^{\lambda} + \beta\log\psi_{t-1}$ |
| Asymmetric ACD ( $v = \lambda = 1$ ) |
| $\psi_t = \omega + \alpha[\lvert\varepsilon_{t-1} - b\rvert + c(\varepsilon_{t-1} - b)]^{\lambda} + \beta\psi_{t-1}$ |
| Power ACD ( $\lambda = v$ and b=c=0) |
| $\psi_t^{\lambda} = \omega + \alpha x_{t-1}^{\lambda} + \beta\psi_{t-1}^{\lambda}$ |
| Box-Cox ACD ( $\lambda \to 0$ and b=c=0) — Dufour and Engle(2000) |
| $\log\psi_t = \omega + \alpha\varepsilon_{t-1}^{\ \lambda} + \beta\log\psi_{t-1}$ |
| Logarithmic ACD type I ( $\lambda, v \to 0$ and b=c=0) — Bauwens and Giot's (2000) |
| $\log\psi_t = \omega + \alpha\log x_{t-1} + \beta\log\psi_{t-1}$ |
| Logarithmic ACD type II ( $\lambda \to 0, v = 1$ and b=c=0) — Bauwens and Giot's (2000) |
| $\log\psi_t = \omega + \alpha\varepsilon_{t-1} + \beta\log\psi_{t-1}$ |
| Linear ACD ( $\lambda = v = 1$ and b=c=0 ) — Engle and Russell(1998) |
| $\psi_t = \omega + \alpha x_{t-1} + \beta\psi_{t-1}$ |