How Well Does the Weighted Price Contribution Measure Price Discovery?

Jianxin Wang and Minxian Yang

August 2012

Abstract

The weighted price contribution (WPC) is a popular measure for price discovery. This paper examines the theoretical properties and empirical performance of the WPC. The benchmark measure for the WPC is the information share (IS) based on the variation of the efficient price. We derive the asymptotic value of the WPC under the assumption of normality. We show that the WPC converges to the IS when the returns follow independent normal distributions with zero mean, and it diverges from the IS when cross-period returns are correlated or the cross-period variance ratio is high. Our theoretical predictions based on normality hold well in the empirical analyses of the overnight price discovery for the S&P 100 index and its constituent stocks. As the correlation between overnight and daytime returns increases in recent years, the deviation between the WPC and the IS becomes large.

JEL classification: G14; G15; C32

Keywords: price discovery, the weighted price contribution, the information share, information flow, the efficient price, overnight return, daytime return, the S&P 100 index.

Jianxin Wang (jianxin.wang@uts.edu.au) is from UTS Business School, University of Technology Sydney, Sydney, Australia, 2007. Minxian Yang (m.yang@unsw.edu.au) is from the Australian School of Business, University of New South Wales, Sydney, Australia 2052. Jianxin Wang is the corresponding author. We thank Doug Foster, Raymond Liu, David Michayluk, and Michael Theobald, and participants of the UTS microstructure conference and the Finance and Corporate Governance Conference of La Trobe University for helpful comments. Research support from the Australian Research Council and the Commonwealth Bank of Australia is gratefully acknowledged. All errors are our own.

I. Introduction

A core function of financial markets is price discovery, the process that incorporates new information into asset prices. The price discovery process is affected by a wide range of factors, including rules and regulations, characteristics of assets traded and market participants, features of the trading platform, etc. How well a trading venue performs the price discovery function has a significant impact on the valuation, the volatility, and the liquidity of the traded assets. As new technologies and trading venues replace the traditional exchange-based market structure, their impact on price discovery and market efficiency will be an important issue for investors, regulators, as well as researchers.

Early studies of price discovery focus on parallel markets where the same asset (or highly correlated assets) is traded simultaneously in different venues, e.g. NYSE versus regional exchanges in the United States. Hasbrouck (1995) and Harris, et al. (2002) are the two dominant models for parallel markets and have been adopted by numerous studies. They are analysed and compared in a special issue of the Journal of Financial Markets in 2002. Yan and Zivot (2010) use a structural cointegration model to bring new insights to the comparison. Lien and Shrestha (2009) and de Jong and Schotman (2010) introduce additional structures to improve the Hasbrouck model. Note that these models are designed for parallel markets and are inappropriate for non-overlapping sequential markets (e.g. Tokyo and New York) or time periods (e.g. the pre-opening period and exchange trading hours). Recently Wang and Yang (2011) propose a model for measuring price discovery in sequential markets in the spirit of Hasbrouck (1995).

A popular measure of price discovery in sequential markets is a non-parametric measure called the weighted price contribution (WPC). It was originally proposed by Barclay and Warner (1993) to measure price movements associated with different transaction sizes. Cao, et al,

(2000) is the first to adopt it as a price discovery measure for sequential time periods and gives it the current name. The WPC has been used to measure price discovery during the pre-opening period (Cao, et al. 2000), during overnight trading (Barclay and Hendershott, 2003, 2008), and during opening and closing call auctions (Ellul, et al. 2005). The need to empirically measure price discovery in sequential periods and the simplicity of the WPC greatly enhances its popularity, particularly for supplementing and supporting the core methodology and findings, e.g. Owens and Steigerwald (2005) and Agarwal, et al. (2007).

Although the WPC is widely used, no one has explored its theoretical validity as a price discovery measure.[1] Consider a trading day t that is divided into n consecutive periods. Let $p_{i,t}$ be the log price of an asset at the end period i on day t. Let $r_{i,t} = p_{i,t}-p_{i-1,t}$ be the return in the i^{th} period and $r_t = \sum_{i=1}^{n} r_{i,t}$ be the daily return. The WPC of the i^{th} period is defined as

$$(1) \qquad WPC_i = \sum_{t=1}^{T} \frac{r_{i,t}}{r_t}\left(\frac{|r_t|}{\sum_{s=1}^{T}|r_s|}\right),\ i = 1,\dots,n.$$

The validity of the WPC as a price discovery measure seems to come from its definition: the contribution of the i^{th} period to price discovery is measured by the cross-day weighted average return ratio $r_{i,t}/r_t$, with the weight for day t being $|r_t|/\sum_{s=1}^{T}|r_s|$. Is $r_{i,t}/r_t$ a valid measure for price discovery in the i^{th} period? Is it consistent with the definition of price discovery as the process of incorporating new information into asset prices? This paper aims to answer these questions. We first derive the asymptotic expression for the WPC in (1) under the assumption of normality. We show that it is primarily a measure of the volatility ratio across periods, not the return ratio as it appears. We explore the theoretical relationship between the WPC and the characteristics of the return series: its mean, variance, and serial correlation. We then draw theoretical comparison

[1] Van Bommel (2011) explores the statistical properties of the WPC. We relate our study to that of van Bommel (2011) in section III.

between the WPC in (1) and a benchmark measure of price discovery. Our benchmark is the information share (IS) measure proposed by Wang and Yang (2011), which is based on the variation in the efficient price as in Hasbrouck (1995). We show that the WPC becomes a consistent estimator of the IS only when returns are uncorrelated and have zero means. The difference between the IS and the WPC crucially depends on return characteristics, especially return serial correlations.

We support our theoretical findings by drawing empirical comparisons between the WPC and the IS in the context of estimating the overnight and daytime price discovery for the S&P 100 index and its current constituent stocks. Several studies have documented significant overnight or pre-opening price discovery when the organized exchanges are closed, e.g. Cao, et al. (2000), Barclay and Hendershott (2003, 2004, and 2008), and Moulton and Wei (2005). Tompkins and Wiener (2008) and Cliff, et al. (2008) document positive overnight returns and negative daytime returns across major international markets. The overnight price discovery is reflected in the price change between today's market close and next day's market open. We use the WPC and the IS measure of Wang and Yang (2011) to estimate overnight price discovery. The main empirical findings are the following:

- For the S&P 100 index, the annual time-series analyses indicate that the overnight WPC is indeed largely determined by the standard deviation ratio of overnight and daytime returns, consistent with the theoretical analyses. The asymptotic values of the overnight WPC are very close to the estimated WPC. The difference between the overnight WPC and IS is mainly driven by the correlation between overnight and daytime returns and the standard deviation ratio.

- The cross-sectional analyses based on the S&P 100 stocks confirms that the overnight WPC is determined by the standard deviation ratio of overnight and daytime returns. Furthermore, the correlation between overnight and daytime returns has strong effects on the overnight WPC and its deviation from the overnight IS. Both effects are consistent with the theoretical predictions. Other return characteristics (such as skewness and kurtosis) do not have strong effect on the WPC and its deviation from the IS.
- In recent years, the high correlations between the overnight and daytime returns of the S&P 100 Index have resulted large deviations between the estimated WPC and IS.

In summary, our theoretical analyses show that when returns have near zero mean and near zero correlation, the WPC is a valid measure for price discovery. We show theoretically and empirically that the value of the WPC and its deviation from the IS are very sensitive to return serial correlations. Because of the presence of serial correlations in most return series, the conceptually cleaner and empirically safer choice is the IS measure of Wang and Yang (2011) when measuring price discovery in sequential markets or time periods.

This paper is organized as the following: section II defines sequential markets and motivates the IS of Wang and Yang (2011) as a benchmark measure for price discovery. Section III explores the relationship between the WPC and return characteristics, and draws theoretical comparison between the IS and the WPC. Section IV presents the structural VAR estimation of the IS and the empirical comparisons based on the overnight and daytime returns of the S&P 100 index. Section V concludes.

II. Information Flow and Price Discovery

Price discovery is commonly defined as the incorporation of new information into asset prices, e.g. Garbade and Silber (1983), Hasbrouck (1995), and O'Hara (2003). By definition, new information leads price changes that are uncorrelated to its past. The random-walk component of price changes represents changes in the efficient price and reflects the market's ability to collect and process information. Therefore a natural measure for information flow or price discovery is the variance of the efficient price change.

In this section, we define the structure of sequential markets and outline the IS measure for sequential markets proposed by Wang and Yang (2011). The sequential markets are a natural and appropriate setting for analysing the WPC. The definition in section I shows that the WPC is based on sequential price changes over time. Most empirical analyses of the WPC are based on sequential periods, e.g. the pre- versus post-opening periods, overnight versus daytime trading, etc.[2] Van Bommel (2011) examines the statistical properties of the WPC under the same setting.

To define sequential markets, we again consider a trading day with n sequential periods or markets. These periods or markets do not have to be of equal length. They can be based on trading hours in Asia, Europe, and the United States, e.g. global currency trading in Wang and Yang (2011). They can also be daytime and overnight periods at a single trading venue, e.g. Cao, et al. (2000) and Barclay and Hendershott (2003, 2004, and 2008). The log price at the end of the i^{th} period on day t can be written as $p_{i,t} = m_{i,t} + u_{i,t}$, where $m_{i,t}$ is the efficient price reflecting information on economic fundamentals, and $u_{i,t}$ is a noise term resulting from short-term mispricing or changes in liquidity and microstructure factors, e.g. bid-ask bounce or

[2] When applied to parallel markets, as in Huang (2002), the WPC is based on the price change around a single trade at a given time. It cannot be calculated when there are simultaneous trades in different markets.

inventory control. The changes in the efficient price $\Delta m_{i,t} = m_{i,t}-m_{i-1,t}$, i=1,…,n, are serially uncorrelated and capture the permanent or information components in price innovations. The information flow in the i^{th} period is measured by $var(\Delta m_{i,t})$. The change of the efficient price over day t is $\Delta m_t = \sum_{i=1}^{n} \Delta m_{i,t}$. The information share of period i on day t is defined as

$$(2) \qquad IS_i = \frac{var(\Delta m_{i,t})}{var(\Delta m_t)} = \frac{var(\Delta m_{i,t})}{\sum_{i=1}^{n} var(\Delta m_{i,t})}, \ i=1,\ldots,n.$$

The above measure is in the same spirit of Hasbrouck (1995), and is used as the benchmark for analysing the characteristics and performance of the WPC.

For the empirical analysis, a trading day t is defined from the market close on day t-1 to the market close on day t.[3] It is divided into overnight and daytime periods: n=2. Let $p_{o,t}$ and $p_{c,t}$ be the log opening and closing values of the S&P 100 index (or individual stock price) respectively. The overnight return is $r_{N,t} = p_{o,t} - p_{c,t-1}$ and the daytime return is $r_{D,t} = p_{c,t} - p_{o,t}$. Wang and Yang (2011) model the return vector, $R_t = [r_{N,t}, r_{D,t}]'$, as a structural VAR process:[4]

$$(3) \qquad B_0 R_t = a + \sum_{k=1}^{K} B_k R_{t-k} + \eta_t,$$

where $\eta_t = [\eta_{N,t}, \eta_{D,t}]'$ is the vector of structural shocks and $a=[a_N, a_D]'$ is the vector of intercepts. Here $\eta_{N,t}$ and $\eta_{D,t}$ are serially uncorrelated and reflect respectively the night-specific and day-specific shocks to the return process. Their variances are normalized to one. Therefore $E(\eta_t)=0$; $E(\eta_t \eta_{t-k}') = 0$ for $k \neq 0$; $E(\eta_t \eta_t') = I$, a 2×2 identity matrix.[5] Any serial correlation between $r_{N,t}$ and $r_{D,t}$, from short-term mispricing or microstructure factors, is captured by the structural coefficient B_0, which is a lower triangular matrix: within the same trading day t, $r_{N,t}$ affects $r_{D,t}$

[3] Our definition of a trading day implies that the overnight period precedes the daytime trading period. As shown by Wang and Yang (2011), rotating the periods does not affect the structural VAR estimation.
[4] Note that R_t differs from r_t, which is the daily return $p_{c,t} - p_{c,t-1}$ defined in section II.
[5] An alternative and equivalent parameterization is to normalize the diagonal elements of B_0 as unity and leave the variance of η_t as a positive diagonal matrix.

but not vice versa. The impact of daytime trading on overnight returns is captured by the lagged returns on the right hand of (3).

The corresponding reduced form VAR is given by $A(L)R_t = \alpha + \varepsilon_t$, where $A(L) = I - A_1L - ... - A_KL^K$, $A_k = B_0^{-1}B_k$, and $\alpha = B_0^{-1}a$. The vector of reduced-form shocks is given by $\varepsilon_t = B_0^{-1}\eta_t$. As discussed above, the daily closing price $p_{c,t}$ can be viewed as a combination between an efficient price m_t that follows a random walk, and a serially-correlated noise component. Although the efficient price is not observable, Wang and Yang (2011) show that the daily change of the efficient price m_t is given by

(4) $$\Delta m_t = \mu + \iota' A(1)^{-1}B_0^{-1}\eta_t = \mu + h'\eta_t = \mu + h_N\eta_{N,t} + h_D\eta_{D,t},$$

where $\mu = \iota' A(1)^{-1}B_0^{-1}a$, $\iota = [1,1]'$, and $h' \equiv [h_N, h_D] = \iota' A(1)^{-1}B_0^{-1}$. Since $E(\eta_t\eta_t') = I$, therefore $var(\Delta m_t) = h_N^2 + h_D^2$, and the IS defined in (2) becomes

(5) $$IS_i = \frac{h_i^2}{h_N^2 + h_D^2}, \; i = N \text{ or } D.$$

Note that A(1) in the reduced-form VAR is easily estimated by OLS and the B_0^{-1} matrix is the lower triangle Cholesky factor of the estimated variance matrix of ε_t. Hence the IS is almost as easy to compute as the WPC.

III. Understanding the WPC

In this section, we explore the asymptotic properties of the WPC in (1) and compare them with the information share measure in (2). We can rewrite the WPC in (1) as

$$WPC_i = \frac{1}{\sum_{s=1}^{T}|r_s|}\sum_{t=1}^{T}\frac{r_{i,t}|r_t|}{r_t} = \frac{\sum_{t=1}^{T}\text{sign}(r_t)r_{i,t}}{\sum_{t=1}^{T}|r_t|} = \frac{\sum_{t=1}^{T}\text{sign}(r_t)r_{i,t}}{\sum_{t=1}^{T}\text{sign}(r_t)r_t}, \; i = 1,\ldots,n.$$

where sign(x) is the sign of x, being 1 for positive x and -1 for non-positive x. The WPC then can be interpreted as the ratio of the weighted average returns, where the weight is sign(r_t). By the law of large numbers,

$$\text{(6)} \qquad \text{WPC}_i \rightarrow \frac{E[\text{sign}(r_t)r_{i,t}]}{E(|r_t|)} = \frac{E[\text{sign}(r_t)r_{i,t}]}{E[\text{sign}(r_t)r_t]}, \; i = 1,\ldots,n.$$

in probability as T → ∞. Equation (6) gives the large-sample WPC. Define $r_t = r_{i,t} + r_{-i,t}$, we have the following theorem (the proof is given in the Appendix).

Theorem on the Large-Sample WPC:

Assume that the returns ($r_{i,t}$, $r_{-i,t}$) are jointly normally distributed with means (μ_i, μ_{-i}), variances (σ_i^2,σ_{-i}^2) respectively and correlation ρ. Define $\mu = E(r_t) = \mu_i + \mu_{-i}$ and $\sigma^2 = var(r_t) = \sigma_i^2 + \sigma_{-i}^2 + 2\rho\sigma_i\sigma_{-i}$. The large-sample WPC is

$$\text{(7)} \qquad \frac{E[\text{Sign}(r_t)r_{i,t}]}{E(|r_t|)} = \frac{2\mu_i\left[0.5-\Phi\left(-\frac{\mu}{\sigma}\right)\right]+\sqrt{\frac{2}{\pi}}\exp\left[-\frac{\mu^2}{2\sigma^2}\right](\sigma_i^2+\rho\sigma_i\sigma_{-i})/\sigma}{2\mu\left[0.5-\Phi\left(-\frac{\mu}{\sigma}\right)\right]+\sqrt{\frac{2}{\pi}}\exp\left[-\frac{\mu^2}{2\sigma^2}\right]\sigma}, \quad i = 1,\ldots,n.$$

where Φ is the standard normal cumulative distribution function.

The Theorem reveals the determinants of the large-sample WPC under the normality assumption. It allows us to further explore the WPC's relationship with return parameters (μ, σ, and ρ) and the IS in (1). The corollaries below can be easily seen from equation (7).

Corollary 1: *In the absence of autocorrelations (with respect to t),when μ =0 and ρ = 0, WPC_i → IS_i in probability.*

When there is no autocorrelation within and across t, $p_{i,t}$ follows a random walk and is the efficient price. Therefore $r_{i,t} = \Delta m_{i,t}$, $\sigma_i^2 = var(\Delta m_{i,t})$ in equation (1), $IS_i = \sigma_i^2/\sigma^2$. From (7), $WPC_i \rightarrow \sigma_i^2/\sigma^2$ when $\mu \rightarrow 0$ and $\rho \rightarrow 0$. That is, when price follows a drift-less Gaussian martingale, WPC and IS are identical.[6] Corollary 1 points to the importance of μ and ρ in determining the relationship between the WPC and the IS. Figure 1 depicts the surface of (7) as functions of μ and ρ with $\mu_i = 0.2\mu$ ($\mu_{-i} = 0.8\mu$), $\sigma_i = 1$ and $\sigma_{-i} = 2$. It shows that the large-sample WPC is not very sensitive to changes in μ. On the other hand, it is very sensitive to the return serial correlation ρ.

While the mean return is generally very small ($\mu \approx 0$), especially at daily or higher frequencies, returns are generally serially correlated ($\rho \neq 0$). We further explore the determinants of the WPC when $\mu \to 0$. From (7),

(8) $$\mathrm{WPC}_i \to \frac{\sigma_i^2 + \rho\sigma_i\sigma_{-i}}{\sigma^2} \text{ in probability when } \mu \to 0 \text{ but } \rho \neq 0,$$

The presence of the noise term makes $\rho \neq 0$, resulting in the deviation between the WPC and the IS. The result in (8) leads to the following corollaries.

Corollary 2: $\frac{\partial WPC_i}{\partial(\sigma_i/\sigma_{-i})} > 0$ *when* $\rho > \frac{-2\sigma_i\sigma_{-i}}{\sigma_i^2+\sigma_{-i}^2}$ *and* $\mu \to 0$.

Let $\lambda_i = \frac{\sigma_i}{\sigma_{-i}}$. From (8) we have $\mathrm{WPC}_i \to \frac{\sigma_i^2+\rho\sigma_i\sigma_{-i}}{\sigma_i^2+\sigma_{-i}^2+2\rho\sigma_i\sigma_{-i}} = \frac{\lambda_i+\rho}{\lambda_i+\lambda_i^{-1}+2\rho}$. Then it is easy to see that $\frac{\partial WPC_i}{\partial \lambda_i} = \frac{2\lambda_i+\rho\lambda_i^2+\rho}{\lambda_i^2(\lambda_i+\lambda_i^{-1}+2\rho)^2} > 0$ when $\rho > \frac{-2\lambda_i}{1+\lambda_i^2} = \frac{-2\sigma_i\sigma_{-i}}{\sigma_i^2+\sigma_{-i}^2}$. The condition holds when $\rho \geq 0$ and may not hold when $\rho \ll 0$, e.g. when $\rho = -1$. Corollary 2 shows that WPC_i is an increasing function of the relative volatility σ_i/σ_{-i} when the mean return is small and the serial correlation ρ is not severely negative. Figure 2 depicts the surface of equation (9) as functions of ρ and the

[6] Von Bommel (2011) compares the WPC with the benchmark measure $\theta_i = 1 - \frac{\mathrm{var}(r_t|r_{i,t})}{\mathrm{var}(r_t)}$ which is the population R^2 for the unbiasedness regression $r_t = \alpha + \beta r_{i,t} + \varepsilon_t$. One can easily show that $\theta_i = \frac{\sigma_i^2+\rho^2\sigma_{-i}^2+2\rho\sigma_i\sigma_{-i}}{\sigma^2}$, therefore $\theta_i = \mathrm{IS}_i$ when μ=0 and ρ=0. When $\rho \neq 0$, which usually is the case for the observed returns, θ_i depends on price movements in other periods σ_{-i}^2. This contradicts the definition of price discovery as the process of incorporating *new* information into prices. Therefore we believe that θ_i is not the appropriate benchmark for price discovery.

relative volatility $\lambda_i = \sigma_i/\sigma_{-i}$. WPC_i is increasing in σ_i/σ_{-i} when $\rho \geq 0$. The opposite is true when $\rho < 0$ and λ_i is very small, or when ρ is close to -1.

Corollary 3: $\frac{\partial WPC_i}{\partial \rho} = \frac{\sigma_i \sigma_{-i}}{\sigma^4}(\sigma_{-i}^2 - \sigma_i^2)$ *when* $\mu \to 0$.

Because the denominator in (8) σ^2 is also a function of ρ, the impact of ρ on WPC_i depends on the relative values of σ_i and σ_{-i}. Figure 2 shows that WPC_i is increasing in ρ when $\sigma_i/\sigma_{-i} < 1$ (except the extreme case of $\sigma_i/\sigma_{-i} = 0$) and is decreasing in ρ when $\sigma_i/\sigma_{-i} > 1$. To illustrate the intuition behind the result, consider a market with two sequential trading periods (n=2). Corollary 3 states that $\frac{\partial WPC_1}{\partial \rho} > 0$ if $\sigma_1 < \sigma_2$. Note that $\sigma_1 < \sigma_2$ indicates volatility spill-over from yesterday's period 2 to today's period 1. A higher ρ leads to greater spill-over, increasing period 1's proportional variance σ_1/σ. This in turn increases WPC_1 as shown in Figure 1.

Corollary 4: *When* $\frac{\mu^2}{\sigma^2} \to \infty$, $WPC_i \to \mu_i/\mu$ *in probability.*

Corollary 4 sets the condition when the WPC in (1) converges to the ratio of mean returns. Since the condition is never satisfied in real financial data, the WPC in (1) does not measure the ratio of the average returns μ_i/μ.

The above analyses provide several insights to the properties of the WPC (1) and its relationships with the IS. The IS in (2) and (5) have a clear economic interpretation as the relative changes in the efficient price in different markets/periods. It is difficult to give such an interpretation to the WPC in (1) and (7) because of the correlation ρ. Note that IS in (5) is independent of ρ. From Corollary 3, $\frac{\partial (WPC_i - IS_i)}{\partial \rho} = \frac{\sigma_i \sigma_{-i}}{\sigma^4}(\sigma_{-i}^2 - \sigma_i^2)$ when $\mu \to 0$. Conditional on $\sigma_{-i}^2 \neq \sigma_i^2$, the difference between the WPC and the IS widens as ρ deviates from zero.

IV. Empirical Comparison between the IS and the WPC

The analyses in the previous section are based on the assumption that returns are normally distributed. In this section, we empirically explore the determinants of the WPC and its deviation from the IS. The aims for the empirical analyses are two folds. First, we assess the theoretical predictions of section III against S&P 100 index returns that may not be normally distributed and examine how the WPC is affected by non-normality in returns, i.e. skewness and kurtosis. Second, we want to explore the determinants of the deviations between the WPC and the IS.

Many studies have documented significant overnight or pre-opening price discovery. The overnight and daytime periods can be viewed as two sequential markets, even though there may or may not be overnight trading. This overnight-daytime sequential markets provide the platform for the empirical analyses. While the daily closing values of the S&P 100 index are available for many decades, the daily opening value of the index is available from DataStream only from March 5, 1999. Our sample period is from March 5, 1999 to April 20, 2010. Figure 3 shows the index value over the sample period. It has three distinctive trends: a bear market due to the burst of the technology bubble from mid-2000 to early 2003, a bull market from early 2003 to late 2007, and the crash associated with the recent global financial crisis. We will examine the WPC in these three sub-periods.

Table 1 reports the overnight and daytime returns and volatility of the S&P 100 index over the sample period. The magnitudes of the daytime return and volatility are much larger than overnight return and volatility. The "bad-day and good-night" return pattern generally does not hold for the S&P 100 index. Only four of the eleven years have the average overnight return higher than the average daytime return. The result is consistent with Tompkins and Wiener (2008) but in contrast with Cliff, et al. (2008).

The annual estimates of the overnight IS (IS_N) and the overnight WPC (WPC_N) are reported in Table 2. IS_N is estimated from the structural VAR in (2) and WPC_N is based on equation (5). The lag length of the structural VAR is based on the Schwarz criterion. We also report the ratio of overnight and daily return volatility (σ_N/σ), the correlation between overnight and daytime returns $Cor(r_N,r_D)$, and the large-sample WPC_N (WPC_N^0) calculated from equation (8) using the sample statistics. There are several features in Table 2:

- Confirming Corollary 1, when the correlation between night return and day return is small, as in 1999 and 2001 – 2006, the overall similarity between the IS and the WPC is evident. The differences between the WPC and the IS become large (more negative) when the overnight and daytime return correlations are large in 2000 and 2007-2009. Given the large empirical differences in recent years, using the IS or the WPC may lead to different economic conclusions.
- WPC_N has a correlation of 0.94 with volatility ratio σ_N/σ. While not reported in Table 2, the correlation between WPC_N and return ratio $\bar{r}_N/\bar{r}$ is -0.03. These findings support Equation (9) and Corollaries 2 & 4. The WPC is a measure for variance ratio, not return ratio.
- The theoretical WPC_N^0 from equation (7) is very similar to the estimated sample WPC_N from equation (5), indicating that the normality assumption holds reasonably well for the S&P 100 index returns. While not reported in Table 2, the correlation coefficient between WPC_N and WPC_N^0 is 0.98.
- The estimated overnight information shares vary from 2.7% to 38%. What drives the variation remains unclear in the literature.

While the above analyses generally support the theoretical predictions in section III, they are based on unconditional correlations without controlling the effects of other variables. For example, the year 2004 has a relatively low $Cor(r_N,r_D) = -0.082$ but a large difference between the IS and the WPC. This suggests that high $Cor(r_N,r_D)$ is not the only condition for large deviation between WPC_N and IS_N. Other parameters (e.g. σ_N/σ) and the normality assumption may have a joint effect on WPC and WPC_N-IS_N. Similarly, the correlation between WPC_N and $Cor(r_N,r_D)$ is negative (-0.56) even though $\sigma_D > \sigma_N$. It appears to violate Corollary 3 which predicts a positive effect from $Cor(r_N,r_D)$ when $\sigma_D > \sigma_N$. However Corollary 3 requires σ_N and σ_D to be held constant while $Cor(r_N,r_D)$ changes. Therefore a direct test of Corollary 3 beyond normality requires a regression analysis where the influence of other variables can be controlled. Another concern is the calculation of the opening value of the S&P 100 index which may lead to artificially high overnight and daytime return correlation. If a stock is not traded at the opening, the previous closing price is used to calculate the index's opening value. This may lead to spurious autocorrelation between daytime and overnight returns of the index.[7] This potential problem can be avoided by using individual stocks where the opening price is taken from the first trade of the day.

The above issues motivate our cross-sectional regression analyses based on the stocks in the S&P 100 index. The aim is to directly test how the theoretical predictions obtained under normality are affected by non-normality characteristics in return series (such as skewness and kurtosis). The regression analyses also allow us to examine the effect of a factor of interest, e.g. $Cor(r_N,r_D)$, while controlling the effects of other factors. We start by looking at the cross-sectional summary statistics of return characteristics and the estimated WPC and IS for S&P 100

---

[7] We thank Raymond Liu for pointing out this potential problem.

stocks in Panel A of Table 3. The average overnight return is higher than the average daytime return, although the difference is not statistically significant. Daytime volatility is much higher than overnight volatility. Overnight returns are more skewed and have greater kurtosis. The average daytime and overnight return correlation ranges from -0.24 to 0.12. The average overnight information share is 25%, higher than the average overnight WPC which is 21%[8]. The overnight WPC does not capture the wide cross-sectional range in information share revealed by IS. For most stocks, overnight and daytime returns are not autocorrelated across trading days; and the number of lags in the SVAR model is mostly zero. This is not surprising given the sample of large and actively traded stocks. Figure 4 shows the scatter graph of WPC_N against IS_N. There is a positive cross-sectional correlation between WPC_N and IS_N, consistent with their positive time-series correlation (0.66) for the S&P 100 index in Table 2.

Panel B of Table 3 reports WPC_N - IS_N in relation to the different quartiles of the variance ratio σ_N/σ_D and the overnight-daytime correlation $Cor(r_N,r_D)$. When σ_N/σ_D and $Cor(r_N,r_D)$ are both in their bottom quartiles, i.e. $\sigma_N/\sigma_D < 0.484$ and $Cor(r_N,r_D) < -0.08$, the average deviation between WPC_N and IS_N is 2.6%. When σ_N/σ_D and $Cor(r_N,r_D)$ are both in their top quartiles, i.e. $\sigma_N/\sigma_D \geq 0.562$ and $Cor(r_N,r_D) \geq 0.004$, the average difference becomes -11.3%. Out of the 100 stocks, 35 have the deviation between WPC_N and IS_N greater than 5%, with a mean value of 9.5%. The deviation between WPC_N and IS_N can be substantial depending on σ_N/σ_D and $Cor(r_N,r_D)$.

We explore the empirical relationship between the WPC and return characteristics using the following cross-sectional regression:

[8] We note that the average WPC_N and IS_N for individual stocks are higher than those of the S&P 100 index reported in Table 2. As discussed in footnote 8, when a stock is not traded at opening, its previous closing price is taken as its opening price when calculating the index opening value. This results in a zero overnight return, reducing the overnight return variance and overnight price discovery for the index.

(9) $$WPC_{N,i} = \beta_0 + \beta_1(\bar{r}_{N,i}/\bar{r}_{D,i}) + \beta_2(\sigma_{N,i}/\sigma_{D,i}) + \beta_3 Cor(r_{N,i},r_{D,i}) + \beta_4 Skew_{N,i} + \beta_5 Skew_{D,i} + \beta_6 \ln(Kurt_{N,i}/3) + \beta_7 \ln(Kurt_{D,i}/3) + \beta_8 SVARLag_i + \varepsilon_i$$

for i = 1,…,100. The variables $\bar{r}_N/\bar{r}_D$ and σ_N/σ_D allow us to examine the theoretical predictions in Corollaries 2 and 4. Given that $\sigma_N < \sigma_D$, Corollary 3 requires $\beta_3 > 0$ while holding σ_N / σ_D constant. Non-normal return characteristics, i.e. skewness and kurtosis, are included to examine the sensitivities of WPC to non-normalities. $SVARLag_i$ is the number of lags in the SVAR model (2) for estimating IS_N. Even though SVARLag does not directly enter the calculation of the WPC, it is added to capture any effect of inter-daily correlation, as oppose to correlation between $r_{N,t}$ and $r_{D,t}$, on the WPC. We estimate equation (9) for the full sample and for three trend-based sub-periods. The results are reported in Table 4.[9] The t-statistics are based on the heteroskedastic-consistent standard errors. The findings are summarized as the following:

- Consistent with Corollaries 3 and 4, σ_N/σ_D has a strong positive effect on WPC_N with all t-statistics above 11. The coefficients of $\bar{r}_N/\bar{r}_D$ switch signs across sub-periods and are extremely small when they are significant.
- After controlling the effect of other variables, especially σ_N/σ_D, there is a strong positive relationship between WPC_N and $Cor(r_N,r_D)$. This is in contrast to the negative unconditional correlation in Table 2, and confirms Corollary 3 and the asymptotic relationship in Figure 2 when normality does not hold. The positive relationship is robust in all sub-periods, with a slight upward trend over time.
- Skewness does not have any effect on WPC_N. The overnight kurtosis $Kurt_N$ has a small negative effect, while the daytime kurtosis does not affect WPC_N in the last two sub-periods.

[9] The regression results with log[WPC/(1-WPC)] as the dependent variable are similar to those in Table 4. We report the results for WPC in Table 4 to draw comparison with the results for WPC – IS reported in Table 5.

- The inter-daily return correlation, represented by SVARLag, has a small negative effect on WPC_N in the second sub-period.

We analyse the cross-sectional determinants of the deviations between WPC_N and IS_N using the same explanatory variables as in equation (9). Table 5 reports the estimation of equation (9) with $WPC_N - IS_N$ as the dependent variable. The main findings are:

- Table 5 shows that the deviation between WPC_N and IS_N is large (negative) when $Cor(r_N,r_D)$ is large (positive). This is consistent with the results of the SP500 index in Table 2 and provides empirical support for Corollary 1.
- The results show that σ_N/σ_D has a negative effect on the deviation between WPC_N and IS_N: WPC_N tends to severely underestimate IS_N when σ_N/σ_D is large. This explains the anomaly in Table 2: the year 2004 has the highest σ_N/σ hence a large negative WPC_N - IS_N, even though $Cor(r_N,r_D)$ is relatively small.
- The daytime skewness has a small positive effect on WPC_N - IS_N in the first two sub-periods but not in the full sample. The overnight skewness has little effect.
- Table 5 shows that the overnight kurtosis has a small negative effect on WPC_N - IS_N, with similar magnitude as in Table 4. The daytime kurtosis has a small positive effect.
- $WPC_N - IS_N$ is affected by SVARLag, representing the inter-daily return autocorrelation, because the IS takes into account the possible autocorrelations in returns whereas WPC does not. We find that the SVARLag reduces $WPC_N - IS_N$ in up markets, and increases $WPC_N - IS_N$ in down markets.

In summary, our empirical analyses show that the theoretical predictions obtained under normality reasonably hold for the S&P100 stock returns that are not exactly normal. The WPC from equation (1) is sensitive to the relative volatility (Corollary 2) and return serial correlation

(Corollary 3). It is not sensitive to the ratio of mean returns (Corollary 4). Large serial correlation leads to large deviations between the WPC and the IS (Corollary 1). In addition, we find that the WPC tends to severely underestimate the IS when the relative volatility ratio is large. We find that return features related to non-normality, i.e. skewness and kurtosis, have relatively little effect on the WPC or its deviation from the IS.

V. Conclusion

Price discovery is a central function of financial markets and a central theme in market microstructure literature. We theoretically and empirically analysed the properties of the WPC, a popular measure for price discovery in sequential markets. We adopt the information share measure of Wang and Yang (2011) as a comparison benchmark and argue that price discovery measures should be based on changes in the efficient price. We show that the deviation between the WPC and the IS can be substantial depending on the cross-period variance ratio and return correlation. Therefore the IS should be the preferred measure for price discovery. While our analysis is based on sequential trading periods or markets, future research should explore the performance of the WPC when trading takes place simultaneously in parallel markets.

Appendix: Proof of Equation (7)

Following the definitions and notations in sections II and III, we aim to find expressions for $E(|r_t|)$ and $E[\text{sign}(r_t)r_{i,t}]$, where sign(·) is the sign function. We can write $E(|r_t|)$ as

$$
\begin{aligned}
E(|r_t|) &= E[r_t I(r_t > 0)] - E[r_t I(r_t \leq 0)] \\
&= E(r_t) - 2E[r_t I(r_t \leq 0)] \\
&= 2\mu\left[.5 - \Phi\left(-\frac{\mu}{\sigma}\right)\right] + \sqrt{2/\pi}\exp(-\frac{\mu^2}{2\sigma^2})\sigma,
\end{aligned}
$$

where I(·) is the indicator function. Similarly

$$
\begin{aligned}
E\left[\text{sign}(r_t)r_{i,t}\right] &= E\left[\text{sign}(r_t)r_{i,t}I(r_t > 0)\right] + E\left[\text{sign}(r_t)r_{i,t}I(r_t \leq 0)\right] \\
&= E\left[r_{i,t}I(r_t > 0)\right] - E\left[r_{i,t}I(r_t \leq 0)\right] \\
&= E\left(r_{i,t}\right) - 2E\left[r_{i,t}I(r_t \leq 0)\right].
\end{aligned}
$$

Define $\mu_{i|-i} = \mu_i + (\rho\sigma_i/\sigma_{-i})\left(r_{-i,t} - \mu_{-i}\right)$ as the conditional mean of $r_{i,t}$ given $r_{-i,t}$. Using the identity $r_{i,t} = \mu_i + \left[\left(r_{i,t} - \mu_{i|-i}\right) + (\rho\sigma_i/\sigma_{-i})(r_t - \mu)\right]/(1 + \rho\sigma_i/\sigma_{-i})$, we find

$$
E\left[r_{i,t}I(r_t \leq 0)\right] = \mu_i E[I(r_t \leq 0)] + \frac{E\left[\left(r_{i,t}-\mu_{i|-i}\right)I(r_t\leq 0)\right]+(\rho\sigma_i/\sigma_{-i})E[(r_t-\mu)I(r_t\leq 0)]}{[1+(\rho\sigma_i/\sigma_{-i})]}
$$

Therefore we have

$$
\begin{aligned}
&E[I(r_t \leq 0)] = \Phi(-\mu/\sigma), \\
&E\left[\left(r_{i,t} - \mu_{i|-i}\right)I(r_t \leq 0) \mid r_{-i,t}\right] = -\left[\frac{(1-\rho^2)\sigma_i^2}{2\pi}\right]^{1/2}\exp(-\frac{(r_{-i,t}+\mu_{i|-i})^2}{2(1-\rho^2)\sigma_i^2}), \\
&E\left[\left(r_{i,t} - \mu_{i|-i}\right)I(r_t \leq 0)\right] = -\frac{(1-\rho^2)\sigma_i^2}{\sqrt{2\pi}\,\sigma}\exp(-\frac{\mu^2}{2\sigma^2}), \\
&E[(r_t - \mu)I(r_t \leq 0)] = -(2\pi\sigma^2)^{-1/2}\sigma^2\int_{-\infty}^{0} d\exp(-\frac{(r_t-\mu)^2}{2\sigma^2}) = -\frac{\sigma}{\sqrt{2\pi}}\exp(-\frac{\mu^2}{2\sigma^2}),
\end{aligned}
$$

where the last expression illustrates how these expectations are evaluated. Finally, putting the above together, we obtain

$$
\begin{aligned}
&E\left[r_{i,t}I(r_t \leq 0)\right] = \mu_i\Phi\left(-\frac{\mu}{\sigma}\right) - \frac{\sigma_i^2+\rho\sigma_i\sigma_{-i}}{\sqrt{2\pi}\,\sigma}\exp(-\frac{\mu^2}{2\sigma^2}), \\
&E[\text{sign}(r_t)r_{i,t}] = 2\mu_i\left[.5 - \Phi\left(-\frac{\mu}{\sigma}\right)\right] + \sqrt{2/\pi}\exp(-\frac{\mu^2}{2\sigma^2})\frac{\sigma_i^2+\rho\sigma_i\sigma_{-i}}{\sigma}.
\end{aligned}
$$

Equation (8) is given by $\frac{E[\text{sign}(r_t)r_{i,t}]}{E(|r_t|)}$.

Reference

Agarwal, S., C. Liu, and S. G. Rhee, 2007, "Where Does Price Discovery Occur for Stocks Traded in Multiple Markets? Evidence from Hong Kong and London" *Journal of International Money and Finance* 26, 46-63.

Andersen, T. G., and L. Benzoni, 2008, "Realized Volatility," chapter prepared for the Handbook of Financial Time Series, Springer Berlin Heidelberg.

Barclay, M. J., and Jerold B. Warner, 1993, Stealth trading and volatility: Which trades move prices? *Journal of Financial Economics* 34, 281-305.

Barclay, M., and T. Hendershott, 2003, "Price Discovery and Trading after Hours," *Review of Financial Studies* 16, 1041-1073.

Barclay, M., and T. Hendershott, 2004, "Liquidity externalities and adverse selection: evidence from trading after hours", *Journal of Finance* 59, 681–710.

Barclay, M., and T. Hendershott, 2008, "A comparison of trading and non-trading mechanisms for price discovery," *Journal of Empirical Finance* 15, 839-849.

Cao, C., E. Ghysels, and F. Hatheway, 2000, "Price discovery without trading: Evidence from the Nasdaq preopening," *Journal of Finance* 55, 1339-1365.

Cliff, M., and M. J. Cooper, and H. Gulen, 2008, "Return Differences between Trading and Non-trading Hours: Like Night and Day," manuscript, Virginia Tech.

Covrig, V., and M. Melvin, 2002, "Asymmetric information and price discovery in the FX market: does Tokyo know more about the yen?" *Journal of Empirical Finance* 9, p271-285.

De Jong, F., and P. C. Schotman, 2010, "Price Discovery in Fragmented Markets," *Journal of Financial Econometrics* 8, 1-28.

Ellul, A., Shin, H., Tonks, I., 2005, "Opening and closing the market: Evidence from the London Stock Exchange," *Journal of Financial Quantitative Analysis* 40, 779–801.

Harris, F., T.H. McInish and R.A. Wood, 2002, "Security Price Adjustment across Exchanges: An Investigation of Common Factor Components for Dow Stocks," *Journal of Financial Markets* 5, 277-308.

Hasbrouck, J., 1995, "One Security, Many Markets: Determining the Contributions to Price Discovery," *Journal of Finance*, 50(4), 1175-1199.

Hasbrouck, J., 2002, "Stalking the "Efficient Price" in Market Microstructure Specifications: An Overview," *Journal of Financial Markets* 5, 329-339.

Huang, R. D., 2002, "The quality of ECN and Nasdaq market maker quotes," *Journal of Finance* 57, 1285-1319.

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

Lehmann, B., 2002. "Some Desiderata for the Measurement of Price Discovery across Markets," *Journal of Financial Markets* 5, 259-276.

Lien, D. and K. Shrestha, 2009, "A new information share measure," *Journal of Futures Markets* 29, p377 -395.

Menkveld, A., S. Koopman, and A. Lucas, 2007, "Modeling Round-the-Clock Price Discovery for Cross-Listed Stocks using State Space Methods," *Journal of Business & Economic Statistics* 25, 213-225.

Moulton, P., and L. Wei, 2005, "A Tale of Two Time Zones: Cross-Listed Stock Liquidity and the Availability of Substitutes," manuscript, NYSE.

Owens, J. and Steigerwald, D. G., 2005, "Inferring information frequency and quality," *Journal of Financial Econometrics* 3, 500–524.

Tompkins, R. G., and Z. Wiener, 2008, "Bad Days and Good Nights: A re-examination of Non-Traded and Traded Period Returns," manuscript, Hebrew University of Jerusalem.

Van Bommel, J., 2011, "Measuring Price Discovery: The Variance Ratio, the R^2, and the Weighted Price Contribution," *Finance Research Letters* 8, 112-119.

Wang, J. and M. Yang, 2011, "Housewives of Tokyo versus the Gnomes of Zurich: Measuring Price Discovery in Sequential Markets," *Journal of Financial Markets* 14, 82-108.

Yan, B. and E. Zivot, 2007, "The Dynamics of Price Discovery," working paper, University of Washington.

Yan, B. and E. Zivot, 2010, "A Structural Analysis of Price Discovery Measures," *Journal of Financial Markets* 13, 1-19.

Table 1: Overnight and Daytime Returns of the S&P100 Index

In this table, "$\bar{r}_N$", "$\bar{r}_D$", and "$\bar{r}$" are the average overnight, daytime, and daily return respectively; "σ_N", "σ_D", and "σ" are the overnight, daytime, and daily return volatility respectively; "r_{year}" is the annual index return.

Year	$\bar{r}_N$ (%)	σ_N (%)	$\bar{r}_D$ (%)	σ_D (%)	$\bar{r}$ (%)	σ (%)	r_{year} (%)
1999	0.016	0.296	0.086	1.117	0.101	1.135	21.6
2000	0.018	0.302	-0.075	1.389	-0.057	1.485	-14.4
2001	-0.015	0.460	-0.050	1.390	-0.065	1.474	-16.1
2002	-0.085	0.881	-0.024	1.551	-0.108	1.721	-27.3
2003	0.038	0.563	0.047	1.002	0.085	1.103	21.4
2004	-0.003	0.375	0.020	0.622	0.017	0.700	4.4
2005	0.030	0.257	-0.034	0.567	-0.004	0.620	-0.9
2006	-0.007	0.056	0.066	0.592	0.059	0.596	14.7
2007	0.000	0.138	0.014	0.943	0.015	0.993	3.8
2008	-0.020	0.308	-0.163	2.403	-0.183	2.527	-46.3
2009	-0.023	0.278	0.092	1.478	0.069	1.606	17.5
Full	-0.005	0.411	-0.001	1.286	-0.005	1.375	-14.8

Table 2: Overnight Price Discovery for the S&P100 Index

In this table, "σ_N/σ" is the ratio of overnight and daily return volatility; "$Cor(r_N,r_D)$" is the correlation between overnight and daytime returns; "SVar Lags" is the number of lags in the structural VAR model based on the Schwarz criterion; "IS_N" and "WPC_N" are the overnight information share (equation 12) and the overnight weighted price contribution (equation 6) respectively. WPC_N^0 is the large-sample WPC_N, i.e. equation (8), based on the sample statistics of a given year.

Year	σ_N/σ	$Cor(r_N,r_D)$	WPC_N (%)	IS_N (%)	WPC_N-IS_N (%)	WPC_N^0 (%)	SVar Lags
1999	0.260	-0.070	5.1	3.8	1.3	5.0	0
2000	0.204	0.221	7.7	16.6	-8.9	8.3	0
2001	0.312	0.024	13.1	13.9	-0.8	10.4	2
2002	0.512	-0.080	23.8	19.6	4.2	22.5	0
2003	0.510	-0.092	18.0	19.4	-1.4	21.7	1
2004	0.537	-0.082	25.6	38.0	-12.4	24.8	1
2005	0.414	-0.011	14.0	17.8	-3.8	16.7	1
2006	0.093	0.020	1.1	2.7	-1.6	1.0	1
2007	0.139	0.298	6.1	17.8	-11.7	5.9	0
2008	0.122	0.350	5.3	13.6	-8.3	5.5	2
2009	0.173	0.386	8.5	31.5	-23	9.1	1
Full	0.300	0.064	11.5	14.6	-3.1	10.7	2
Cross-Correlation							
$Cor(r_N,r_D)$	-0.75						
WPC_N	0.94	-0.56					
IS_N	0.48	0.12	0.66				
WPC_N-IS_N	0.34	-0.73	0.15	-0.64			

Table 3: Summary Statistics for S&P100 Stocks

Panel A of this table reports time-series and cross-sectional summary statistics for S&P100 stocks. The columns report various time-series statistics for individual stocks. The rows report cross-sectional statistics of 100 stocks. Subscript N (D) indicates overnight (daytime) statistics. "SVAR Lags" is the number of lags in the SVAR model based on Schwarz criterion. Panel B reports the average value of WPC-IS in relation to the quartiles of σ_N/σ_D and $Cor(r_N,r_D)$. The bottom quartiles of σ_N/σ_D and $Cor(r_N,r_D)$ are 0.484 and -0.08 respectively. The top quartiles of σ_N/σ_D and $Cor(r_N,r_D)$ are 0.562 and 0.004 respectively.

Panel A: Summary statistics of returns, WPC_N, and IS_N

	Mean	St Dev	Min	Median	Max
$\bar{r}_N$	0.014	0.066	-0.124	0.009	0.254
$\bar{r}_D$	0.002	0.063	-0.285	0.005	0.127
σ_N	1.29	0.399	0.664	1.22	2.43
σ_D	2.12	0.564	1.22	1.98	3.62
$Skew_N$	-1.71	3.62	-21.5	-0.872	6.08
$Skew_D$	-0.058	0.595	-3.44	0.019	2.33
$Kurt_N$	2.59	0.945	1.15	2.36	5.63
$Kurt_D$	1.14	0.496	0.475	1.02	3.08
$Cor(r_N,r_D)$	-0.037	0.068	-0.240	-0.028	0.120
IS_N	0.25	0.07	0.08	0.25	0.45
WPC_N	0.21	0.03	0.14	0.21	0.27
SVAR Lags	0.57	1.56	0	0	9

Panel B: Difference between WPC_N and IS_N

	$Cor(r_N,r_D) < -0.08$		$-0.08 \leq Cor(r_N,r_D) < 0.004$		$Cor(r_N,r_D) \geq 0.004$	
	WPC_N-IS_N	#Stocks	WPC_N-IS_N	#Stocks	WPC_N-IS_N	#Stocks
$\sigma_N/\sigma_D < 0.484$	2.6%	3	-0.5%	8	-3.2%	13
$0.484 \leq \sigma_N/\sigma_D < 0.562$	2.2%	13	-3.6%	27	-5.7%	10
$\sigma_N/\sigma_D \geq 0.562$	-1.2%	8	-9.1%	15	-11.3%	3

Table 4: Cross-Sectional Determinants of the WPC

This table presents the results of the following cross-sectional regressions:

$$WPC_{N,i} = \beta_0 + \beta_1(\bar{r}_{N,i}/\bar{r}_{D,i}) + \beta_2(\sigma_{N,i}/\sigma_{D,i}) + \beta_3 Cor(r_{N,i},r_{D,i}) + \beta_4 Skew_{N,i} + \beta_5 Skew_{D,i} + \beta_6 \ln(Kurt_{N,i}/3) + \beta_7 \ln(Kurt_{D,i}/3) + \beta_8 SVARLag_i + \varepsilon_i$$

where $SVARLag_i$ is the number of lags in the SVAR model for estimating IS_N, i = 1,…,100. Subscript N (D) indicates overnight (daytime) statistics. The t-statistics below the coefficients are based on the heteroskedastic-consistent standard errors. The asterisks ** and * indicate statistical significance at 1% and 5% respectively.

	Full Sample 1999/3/5 – 2010/4/20	Down Trend 2000/7/1 – 2003/1/31	Up Trend 2003/2/1 – 2007/9/30	Down Trend 2007/10/1 – 2009/1/31
Constant	-0.061**	-0.076**	-0.063**	-0.086**
	-3.21	*-5.16*	*-2.81*	*-4.38*
$\bar{r}_N/\bar{r}_D$	7×10^{-6}**	-6×10^{-5}	4×10^{-5}	-0.001*
	3.67	*-0.33*	*0.09*	*-2.01*
σ_N/σ_D	0.601**	0.630**	0.608**	0.671**
	16.0	*19.0*	*11.5*	*12.9*
$Cor(r_N,r_D)$	0.238**	0.229**	0.274**	0.318**
	9.34	*9.08*	*5.66*	*10.3*
$Skew_N$	0.0004	-0.0001	0.001	-0.001
	0.68	*-0.14*	*0.88*	*-0.83*
$Skew_D$	-0.001	0.007*	-0.006	-0.002
	-0.27	*1.79*	*-0.69*	*-0.29*
$Kurt_N$	-0.018**	-0.026**	-0.021**	-0.024**
	-7.48	*-8.66*	*-6.68*	*-3.40*
$Kurt_D$	0.013**	0.028**	0.003	0.007
	3.22	*4.43*	*0.37*	*0.88*
SVARLag	-0.002	0.001	-0.011*	0.005
	-1.16	*0.46*	*-2.32*	*1.34*
$\bar{R}^2$	0.75	0.81	0.69	0.66

Table 5: Cross-Sectional Determinants of WPC-IS

This table presents the results of the following cross-sectional regressions:

$$WPC_{N,i} - IS_{N,i} = \beta_0 + \beta_1(\bar{r}_{N,i}/\bar{r}_{D,i}) + \beta_2(\sigma_{N,i}/\sigma_{D,i}) + \beta_3 Cor(r_{N,i},r_{D,i}) + \beta_4 Skew_{N,i} + \beta_5 Skew_{D,i} + \beta_6 \ln(Kurt_{N,i}/3) + \beta_7 \ln(Kurt_{D,i}/3) + \beta_8 SVARLag_i + \varepsilon_i$$

where $SVARLag_i$ is the number of lags in the SVAR model for estimating IS_N, i = 1,…,100. Subscript N (D) indicates overnight (daytime) statistics. The t-statistics below the coefficients are based on the heteroskedastic-consistent standard errors. The asterisks ** and * indicate statistical significance at 1% and 5% respectively.

	Full Sample 1999/3/5 – 2010/4/20	Down Trend 2000/7/1 – 2003/1/31	Up Trend 2003/2/1 – 2007/9/30	Down Trend 2007/10/1 – 2009/1/31
Constant	0.232**	0.140**	0.176**	0.111**
	5.60	*7.72*	*5.46*	*4.89*
$\bar{r}_N/\bar{r}_D$	9.6×10^{-6}**	-4×10^{-5}	0.001	-0.0003
	2.80	*-0.27*	*1.07*	*-0.37*
σ_N/σ_D	-0.487**	-0.301**	-0.405**	-0.230**
	-5.80	*-7.83*	*-5.15*	*-4.17*
$Cor(r_N,r_D)$	-0.548**	-0.611**	-0.640**	-0.516**
	-10.3	*-22.7*	*-10.2*	*-13.7*
$Skew_N$	0.0005	0.001	0.002	-0.004*
	0.25	*0.78*	*1.40*	*-1.97*
$Skew_D$	0.006	0.008*	0.019*	0.004
	0.77	*1.99*	*2.16*	*0.68*
$Kurt_N$	-0.023**	-0.030**	-0.017**	-0.031**
	-4.71	*-8.82*	*-2.86*	*-4.06*
$Kurt_D$	0.023**	0.036**	0.009	0.019*
	2.52	*5.08*	*1.08*	*2.30*
SVARLag	0.005	0.044**	-0.072**	0.061**
	0.91	*10.8*	*-3.19*	*5.23*
$\bar{R}^2$	0.72	0.89	0.84	0.81

Figure 1: The WPC as a function of μ and ρ

Under the assumption of normally distributed returns, the figure depicts the large-sample WPC, i.e. equation (8), as a function of daily mean return μ and the cross-period return correlation ρ. It assumes that $\mu_i = 0.2\mu$ ($\mu_{-i} = 0.8\mu$) with $\sigma_i = 1$ and $\sigma_{-i} = 2$.

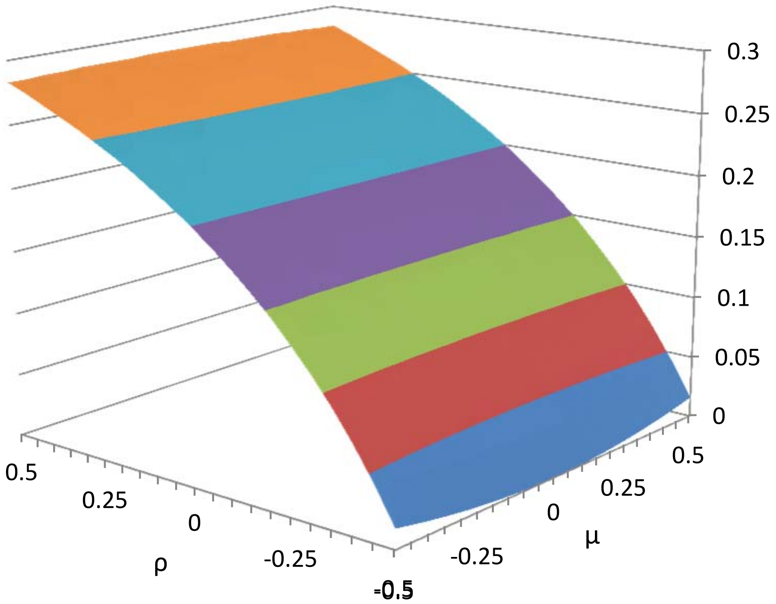

Figure 2: The WPC as a function of ρ and σ_i/σ_{-i}

Under the assumption of normally distributed returns, the figure depicts the large-sample WPC, i.e. equation (8), as a function of return serial correlation ρ and the relative volatility σ_i/σ_{-i}, assuming the mean return is zero $\mu = 0$.

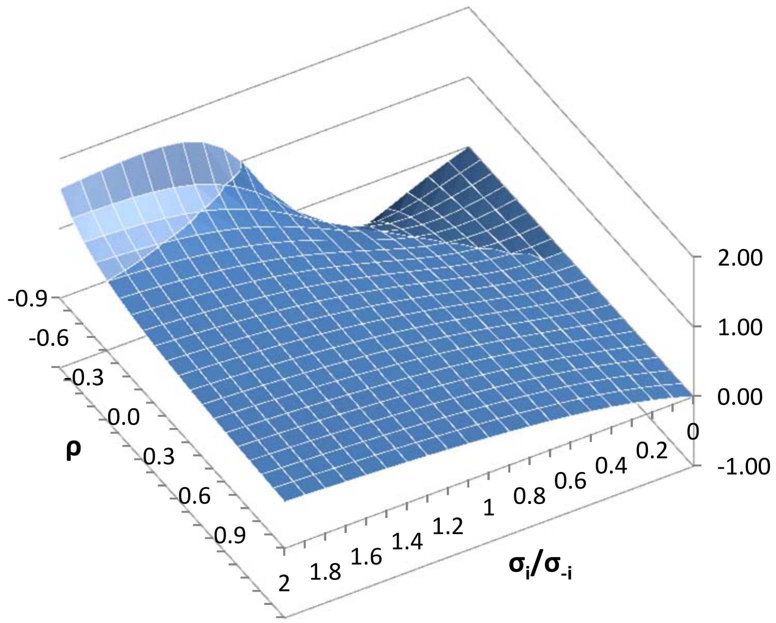

Figure 3: S&P100 Index

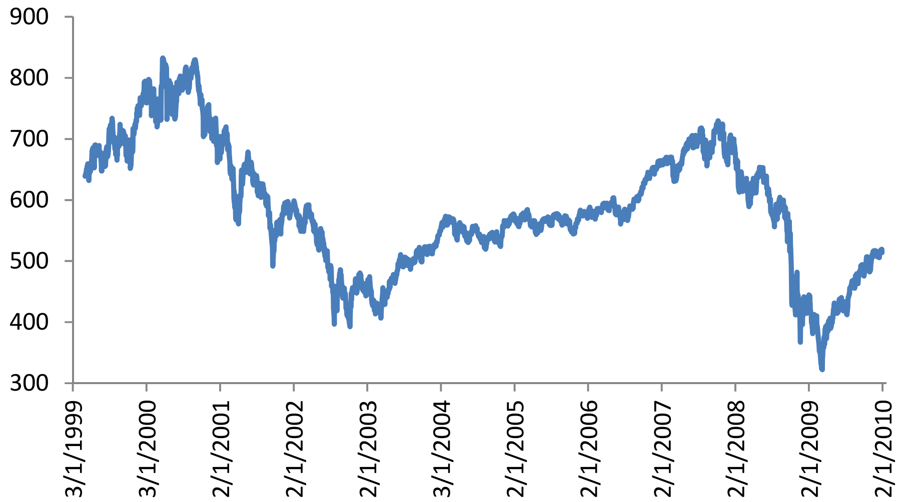

Figure 4: $WPC_N - IS_N$ Scatter Graph