

THE IDENTIFICATION OF PRICE JUMPS

Jan Hanousek
Evžen Kočenda
Jan Novotný

CERGE-EI

Charles University
Center for Economic Research and Graduate Education
Academy of Sciences of the Czech Republic
Economics Institute

Working Paper Series
(ISSN 1211-3298)

434

The Identification of Price Jumps

Jan Hanousek
Evžen Kočenda
Jan Novotný

CERGE-EI
Prague, March 2011



EVROPSKÝ SOCIÁLNÍ FOND
PRAHA & EU: INVESTUJEME DO VAŠÍ BUDOUCNOSTI

**ISBN 978-80-7343-235-5 (Univerzita Karlova. Centrum pro ekonomický výzkum
a doktorské studium)**

ISBN 978-80-7344-225-5 (Národohospodářský ústav AV ČR, v.v.i.)

The Identification of Price Jumps

Jan Hanousek^a, Evžen Kočenda^b, Jan Novotný^c

Abstract

We performed an extensive simulation study to compare the relative performance of many price-jump indicators with respect to false positive and false negative probabilities. We simulated twenty different time series specifications with different intraday noise volatility patterns and price-jump specifications. The double McNemar (1947) non-parametric test has been applied on constructed artificial time series to compare fourteen different price-jump indicators that are widely used in the literature. The results suggest large differences in terms of performance among the indicators, but we were able to identify the best-performing indicators. In the case of false positive probability, the best-performing price-jump indicator is based on thresholding with respect to centiles. In the case of false negative probability, the best indicator is based on bipower variation.

Abstrakt

Provedli jsme extenzivní simulační studii a srovnali výkonnost mnoha různých indikátorů cenových skoků na základě kritéria rozlišné pravděpodobnosti falešné pozitivní a falešné negativní identifikace. Simulovali jsme dvacet různých typů časových řad s různě specifikovaným chováním vnitrodenní volatility tvořené bílým šumem a ne-normálními cenovými skoky. Na simulovaných časových řadách jsme aplikovali dvojitý McNemarův (1947) neparametrický test a porovnali 14 různých indikátorů cenových skoků, které se nejvíce používají ve finanční ekonometrii. Výsledky ukazují velmi odlišné vlastnosti indikátorů, přičemž jsme byli schopni identifikovat ty nejpřesnější z nich. V případě kritéria porovnávajícího pravděpodobnosti falešné pozitivní indikace cenových skoků se jako nejlepší nástroj ukázal indikátor využívající filtrování pomocí centilů. V případě pravděpodobnosti falešné negativní indikace nejlépe fungoval indikátor založený na dvojlenném rozptylu (bipower variance).

Keywords: Price jumps; price-jump indicators; non-parametric testing; Monte Carlo simulations; financial econometrics

JEL classification: C14, C58, F37, G15, G17

The work was supported by the grant SVV–2011–263 801.

The authors thank Petr Zemčík and Miloslav Vošvrda for helpful comments. The authors gratefully acknowledge financial support from the GAČR grant (403/11/0020). The usual disclaimer applies.

^a CERGE-EI, Charles University and the Academy of Sciences, Prague, Czech Republic; Anglo-American University, Prague; The William Davidson Institute, Michigan; and CEPR, London.

^b CERGE-EI, Charles University and the Academy of Sciences, Prague, Czech Republic; Anglo-American University, Prague; CESifo, Munich; OEI, Regensburg; The William Davidson Institute, Michigan; CEPR, London; and the Euro Area Business Cycle Network.

^c CERGE-EI, Charles University and the Academy of Sciences, Prague, Czech Republic; NPI, Prague, Czech Republic; FNSPE, Czech Technical University in Prague, Czech Republic.

CERGE-EI is a joint workplace of the Center for Economic Research and Graduate Education, Charles University, and the Economics Institute of Academy of Sciences of the Czech Republic.

Address: CERGE-EI, P.O. Box 882, Politických vězňů 7, Prague 1, 111 21, Czech Republic

1 Motivation and relevant literature

Discontinuities in price evolution have been recognized as an essential part of the price time series generated on financial markets. Many studies, from the seminal work of Merton (1976) to Andersen et al. (2002), demonstrate that continuous-time models have to incorporate the discontinuous component known in the literature as price jumps. The presence of price jumps has serious consequences for financial risk management and pricing. Nyberg and Wilhelmsson (2009) discuss the importance of including event risk as recommended by the Basel II accord, which suggests employing a VAR model with a continuous component and price jumps representing event risks. Andersen et al. (2007) conclude that most of the standard approaches in the financial literature on pricing assets assume a continuous price path. Since this assumption is clearly violated in most cases the results tend to be heavily biased.¹ Before a price jump can be accounted for in an estimation stage, it first has to be identified. Surprisingly, the literature up to now does not offer a consensus on how to identify price jumps properly. Jumps are identified with various techniques that yield different results. The value-added of this paper is that we perform an extensive and detailed non-parametric study that employs a wide variety of price-jump indicators to identify the superior techniques. Specifically, we have employed the double McNemar (1947) test and compared the fourteen different price-jump indicators most frequently used in the literature employing simulated time series.

Researchers agree on the presence of price jumps, but they disagree about the source. One branch of the literature considers new information as a primary source of price jumps (see e.g. Merton, 1976; Lee and Mykland, 2008 and Lahaye et al., 2010). Joulin et al. (2008) and Bouchaud et al. (2004) conclude that price jumps are usually caused by a local lack of liquidity on the market. They also claim that news announcements have a negligible effect on the origin of price jumps. The behavioral finance literature provides other explanations for price jumps. Schiller (2005) claims that price jumps are caused by market participants who themselves create an environment that tends to cause extreme reactions and thus price jumps. Finally, price

¹ For illustration, Jarrow and Rosenfeld (1984), Nietert (2001) and Pan (2002) study pricing in the presence of jumps and all of them confirm the presence of the jump risk premium. Pricing with jumps using continuous-time diffusion equations was studied by Broadie and Jain (2008), where the authors consider the pricing of volatility and variance swaps. They conclude that the pricing of swaps significantly differs when jumps are taken into account, thus one cannot appropriately price the risk connected with jumps while ignoring the jumps. Carr and Wu (2010) use a jump-diffusion model to simultaneously price stock options and credit default swaps and find a significant presence of the interplay between credit and market risks. A similar confirmation of the change in the pricing mechanism was also shown by Duffie et al. (2000); Liu et al. (2003) and Johannes (2004).

jumps can be viewed as a manifestation of Black Swans, as discussed by Taleb (2007), where the jumps are rather caused by complex systemic interactions that cannot be easily tracked down. In this view, the best way to understand jumps is to be well aware of them and be ready to react to them properly, instead of trying to forecast them.

The key role price jumps play in financial engineering triggered interest in the financial econometrics literature, especially how to identify price jumps. Several different approaches evolved over the recent years. Generally, we can identify in the literature four groups of econometric price-jump indicators.

The first group is represented by the work of Ait-Sahalia (2004), Ait-Sahalia et al. (2009) and Ait-Sahalia and Jacod (2009a, b) using proper statistical methods to derive and analyze the jump statistics based on different analytic models. The indicators have well-defined analytic properties; however, they do not identify price jumps one by one but rather measure the jumpiness of the given period. These methods are more suitable to assess the jumpiness of ultra-high-frequency data, even though they were also employed in previous studies for high-frequency time series.

The second group of price-jump indicators comprises indicators based on bipower variation and is promoted in a series of papers: Barndorff-Nielsen and Shephard (2004, 2006); Barndorff-Nielsen et al. (2004); Barndorff-Nielsen et al. (2006). The method is based on two distinct measures of overall volatility, where the first one takes into account the entire price time movement while the second one ignores the contribution of the model-dependent price-jump component. The papers above also illustrate a broad range of application. This method was further improved by Lee and Mykland (2008), who develop a statistics for the exact identification of moments when particular price jumps occur and employ it for high-frequency stock returns. However, the main disadvantage of bipower variation-based methods lie in the sensitivity of the intraday volatility patterns, which leads to a high rate of jump mis-identification.

The third group is represented by a test developed by Jiang and Oomen (2008). This test relies on the difference between the swap variance and the realized variance. The authors claim that their test is better than the one based on bipower variation since it amplifies the discontinuities to a larger extent, as they show with a comparative analysis using Monte Carlo simulation. The amplification of discontinuities tends, according to the authors, to suppress the effects of intraday volatility patterns.

Finally, the fourth group of price jump-indicator techniques has its roots in Econophysics. Such techniques are based on the scaling properties of time series known in

physics, see e.g. Stanley and Mategna (2000). Gopikrishnan et al. (1999); Eryigit et al. (2009); Jiang et al. (2009) and Joulin et al. (2008) take normalized price time series—the normalization differs across these papers—and define the scaling properties of the tails of the distributions. Then, the scaling index enables them to define the jumpiness of the market purely based on how much of the weight lies in the tails and how this weight is distributed.

As mentioned above, there is still no clear consensus in the literature on how to identify price jumps properly. Bajgrowicz and Scaillet (2010) treat the problem of the spurious identification of price jumps by adaptive thresholds in the testing statistics. The problem with most of the price-jump indicators lies in what model they are built upon. This study illustrates the need for robustness when dealing with price jumps.

The question of the intraday patterns of overall volatility mentioned above is deeply studied in the literature. The work of Wood et al. (1985) documents a U-shaped pattern in the intraday equity volatility. Bollerslev et al. (2008) confirm this effect more robustly. In addition, Novotny (2010) employed many price-jump indicators and studied the difference in price-jump properties during the recent financial crisis using stocks from the NYSE.

Still, to our best knowledge, the literature lacks a deep non-parametric study based on a wide variety of price jump indicators. The literature suggests that identification techniques vary a lot, therefore direct comparison of different papers is not easy. We have focused on this gap in the research and perform a detailed Monte Carlo simulation study to compare price-jump indicators. We have compared price-jump indicators with respect to the false positive and false negative probabilities. We have simulated twenty different kinds of time series with various intraday noise volatilities and different price-jump specifications. Using these simulated time series, we have employed the double McNemar test and compared the fourteen different price-jump indicators most frequently used in the literature.

Our analysis revealed significant differences among the indicators. It was very often the case that one type of indicator clearly dominated the others with respect to the given criterion. Namely, the comparison with respect to the false positive probability was significantly dominated by the indicator based on thresholding with respect to centiles. On the other hand, the comparison with respect to the false negative probability yielded results in which the bipower variation-based indicator dominated. The differences in indicators is very often significant at the highest significance level, which further supports the initial suspicion that the results obtained using different price-jump indicators are not comparable.

2 Price-Jump Indicators

We employ a set of price-jump indicators divided into four groups as outlined in the introduction. These indicators are widely used in the empirical literature but the results of two or more indicators are rarely compared with respect to a single string of data. Hence, the results derived in different papers are hard to compare. In our study we perform a non-parametric comparison of the set of price-jump indicators whose details are outlined in section 4. In this section we first introduce the four groups of price-jump indicators. The technical details of the indicators are further elaborated in Appendix A.

2.1 Group 1: Ait-Sahalia

The first class of indicators goes in a similar direction as Ait-Sahalia (2004) and Ait-Sahalia and Jacod (2009a,b). The price process is assumed to be decomposed into the Gaussian component—corresponding to normal (white) noise—and the non-homogenous Poisson component—corresponding to price jumps. Under certain assumptions it holds that whenever a significant price jump appears, the price increment is dominated by the non-homogenous Poisson component. On the other hand, when the price movements are governed solely by Gaussian noise, the average and/or maximum magnitudes of such increments can be estimated (at a given confidence level). Therefore, one can invert such an argument and set a threshold value that will effectively distinguish the two components.

In practical cases, however, the proper threshold values require a knowledge of what should be estimated. Thus we employ an empirical approach and set the threshold by calculating certain threshold levels, or certain percentiles, of the distribution of returns observed over the entire sample. In addition, the financial time series often have intraday volatility patterns, i.e., the average absolute returns systematically differ over the trading day. To reflect this phenomenon, we further divide every trading day into several blocks and calculate percentiles over these blocks separately.

2.2 Group 2: Bipower Variation

The second group is based on bipower variation, as in Berndorff-Nielsen and Shephard (2004, 2006). Specifically, it is based on the difference between the two measures of variation: realized variation and bipower variation. Assuming the price generating process can be decomposed into two components—regular white noise and price jumps—the realized variation measures the variation in the prices coming from both the white noise and the price jumps, while the bipower

variation measures the variation coming from the white noise only. This measure can be applied in two different ways.

The first approach, proposed by Berndorff-Nielsen and Shephard (2004, 2006) and further elaborated by Huang and Tauchen (2005), involves the construction of a statistics whose purpose is to determine the presence of price jumps over a given time window, i.e., to test the hypothesis that a given time window contains price jumps at all. This statistics, known as the Max-Adjusted statistics, can be thus employed to identify the exact moment when a price jump occurs. Namely, we fix the length of the testing window and for every time step we test a given window ending at that time step for the presence of price jumps. Then, we say that a price jump occurs at that moment if the window ending at that moment contains a price jump while the window ending at the preceding time step does not.

A problem occurs for consecutive price jumps. If two price jumps are separated by an interval shorter than the given window, the second price jump cannot be identified. Hence, we modify the technique in such a way that after we identify a price jump, we replace it with an average calculated over a moving window of the same length. Since these observations by definition do not contain a price jump, their average also excludes price jumps.

The second approach, constructed by Lee and Mykland (2008), also employs bipower variation. However, it is by definition constructed as a statistics to identify price jumps and the moments when they occurred. The statistics compares the current price movement with the bipower variation calculated over a moving window with a given number of preceding observations, excluding the current one.

2.3 Group 3: Jiang-Oomen Statistics

Jiang and Oomen (2008) proposed another statistics to test for the presence of price jumps over a certain time window based on Swap Variance. It is claimed that this test amplifies the contribution of price jumps to a larger extent than bipower variation indicators and thus are less sensitive to intraday volatility patterns. Since the Jiang and Oomen statistics is constructed as a test statistics for a certain time window, the price-jump indicator is analogous to the one following the Berndorff-Nielsen and Shephard method: For every time step, we define a moving time window of a given length ending at the time step and test for the presence of price jumps over the window. Then, we identify a certain moment as the one when the price jump occurred if the window ending at the current time step contains a price jump and the one ending

at the previous time step does not. In addition, we define an analogous improved indicator, which involves replacing the identified price jumps with moving averages and thus allows for identification of consecutive price jumps.

2.4 Group 4: Statistical Finance

The last group of identification techniques come from the field of statistical finance, as it is called by Bouchaud (2002), although it is also known as Econophysics. This group of indicators relies on the scaling properties of price movements. We employ the price-jump index defined by Joulin et al. (2008).

The price index is defined as the absolute returns normalized with respect to the L_1 variance, i.e., the variance defined as an average of absolute returns over a certain moving window. The price-jump index has, as the literature confirms (Joulin et al., 2008), certain scaling properties of the tail part of its distribution. Thus, we define price jumps as those returns for which the price-jump index exceeds a certain empirically determined threshold.

3 Test to compare the performance of the different price-jump indicators

Here we introduce the procedure to compare the performance of the different price-jump indicators. The procedure itself is based on the double McNemar (1947) test, which is a non-parametric method used on nominal data. The intuition behind employing this method lies in the fact that, based on extensive simulations, we want to compare the price-jump indicators with each other rather than study their finite sample properties. Hence, the comparison will be based on the prediction accuracy of the indicators. This means that price-jump indicator A dominates indicator B if A has a significantly better accuracy of jump prediction. This strategy leads to a test procedure to compare the proportions of correctly and incorrectly predicted jumps. The main idea for this approach is natural: if the price-jump indicators are not different in terms of the accuracy of the prediction, it is hard to judge whether one indicator dominates the other in other ways (for binary models see Hanousek, 2000).

Since the jump process $\{Y_t\}$ could be understood as a binary process (0-1), 1 being associated with a jump, studying the prediction accuracy would lead to the following binary outcomes with the probabilities:

$p_{11} = Pr(1|1) = Pr(Y \text{ hat} = 1 | Y = 1)$, i.e., the probability of correct prediction when $Y = 1$;

$p_{22} = Pr(0|0) = Pr(Y \text{ hat} = 0 | Y = 0)$, i.e., the probability of correct prediction when $Y = 0$;

$p_{12} = Pr(1|0) = Pr(Y \text{ hat} = 1 | Y = 0)$, i.e., the probability of wrong prediction when $Y = 0$; (1)

$p_{21} = Pr(0|1) = Pr(Y \text{ hat} = 0 | Y = 1)$, i.e., the probability of wrong prediction when $Y = 1$.

In diagnostics terminology the above probabilities are usually called *sensitivity* (p_{11}), *selectivity* (p_{22}), *false positive* (p_{12}) and *false negative* (p_{21}) probabilities. It is clear that in different situations we might prefer different treatments of misclassification by giving to $Pr(1|0)$ and $Pr(0|1)$ different weights. For the sake of simplicity let us consider the case when both misclassifications have equal weights, i.e., we concentrate on the standard case in which the probability of correct prediction is maximized and where

$$Pr(\text{correct prediction}) = Pr(0|0) + Pr(1|1) = p_{11} + p_{22}. \quad (2)$$

Using a complementary approach, one can minimize the probability of incorrect prediction:

$$Pr(\text{incorrect prediction}) = Pr(0|1) + Pr(1|0) = p_{21} + p_{12}. \quad (3)$$

The above approach could be used to search for and study the “best” price-jump indicator in terms of the accuracy of prediction. However, the optimization procedures conducted on simulated data (potentially) might not capture the relationship, as well as the difference (in prediction) between the relationships between the studied indicators, and therefore it is better to use this approach for a pair-wise comparison of jump indicators.

In terms of any pair-wise comparison/test we can assume the following. We have the available outcomes of price-jump prediction given by two jump indicators denoted as I_1 and I_2 . The combination of their outcomes in terms of the accuracy of price-jump prediction can be summarized by Table 1.

Table 1. Pair-wise comparison of the prediction accuracy of two jump indicators.

		I_1 jump prediction		Σ (total)
		correct	incorrect	
I_2 jump prediction	Correct	n_{11}	n_{12}	$n_{1.}$
	incorrect	n_{21}	n_{22}	$n_{2.}$
Σ (total)		$n_{.1}$	$n_{.2}$	n

n is the total number of simulated returns, n_{11} denotes the number of cases when both indicators correctly identify a price jump. n_{12} is the number of cases when the I_2 correctly identifies a price jump and I_1 does not. n_{21} is the number of cases when the I_1 correctly identifies a price jump and I_2 does not. Finally, n_{22} denotes the number of cases when both indicators do not correctly identify a price jump. In other words, Table 1 is a contingency table, summarizing the outcomes of the two binary variables I_1 and I_2 using the accuracy of prediction as the additional classification dimension. Therefore we adopt standard notation for contingency tables and a dot used in a subscript indicates the corresponding marginal distribution, for example, $n_{1.}$ stands for the number of price jumps correctly identified by I_1 .

The statistical inference of whether jump indicator I_1 dominates I_2 in prediction accuracy can be assessed by testing the null hypothesis $H_0: n_{1.} = n_{2.}$, or equivalently $H_0: n_{12} = n_{21}$. This approach directly leads to the well-known McNemar (1947) test, whose underlying test statistics is

$\chi_1^2 = \frac{(n_{12} - n_{21})^2}{n_{12} + n_{21}}$ and is distributed asymptotically as χ_1^2 ($n_{12} + n_{21} \geq 8$). For smaller values of $n_{12} + n_{21}$ we can construct an exact test using probabilities in multinomial distribution.²

In Table 1 we can set various criteria for prediction accuracy, for example the classification used in the table. The example above used the approach where misclassifications $Pr(1|0)$ and $Pr(0|1)$ have the same weight in selecting price-jump indicators. We compare the correct identification of the price jump only with the incorrect identification in which we

² It is also recommended to conduct an exact test if 20% of $n \cdot n_i \cdot n_j$ is less than 5, or if any of $n \cdot n_i \cdot n_j$ is smaller than one (see for example Gibbons, 1997).

combine both types of misclassification. Since in reality those misclassifications have different impacts we test and compare the price-jump indicators using only false negative and false positive classifications. This means that we can treat misclassification only when an indicator predicts a jump but there was no jump (false positive) or when the indicator does not predict a jump, but we observe a jump (false negative). If we minimize the false positive criterion, the winning indicator would identify fewer returns as false price jumps and would potentially miss some true price jumps. A similar logic is valid for the false negative criterion.

The testing procedure in the simulation framework is applied as follows:

- 1) In the first step, we simulate 100 trading days and compare the indicators pair-wise.³ As the prediction criteria, we use:
 - a. the number of correctly identified price jumps and
 - b. the number falsely identified price jumps.

We conduct the McNemar-type test described above and we count it if indicator I_1 dominates I_2 (90%, 95% and 99% significance levels).

- 2) In the second step, we repeat each simulation 100 times. The results from the test procedure (the first step) are used as the input for the second step. The second test is again the McNemar-type test, where we compare the number of cases when one indicator dominates the other at a given confidence level. For both tests, we use three confidence levels—90%, 95% and 99%.

To summarize, first we use the test for a given (simulated) window of trading days to analyze if one jump indicator dominates the other in terms of the accuracy of the prediction of the price jump. The second step analyzes the results of repeated simulation using the same time window.

4 Data Generation of Artificial Time Series

The goal of this part is to compare the price-jump indicators to find the one that performs best. For that purpose, we perform an extensive simulation study with simulated data. We simulate the price of a virtual asset during a trading day: every trading day lasts seven trading hours or 420 trading minutes. The price time series is simulated at a 1-minute frequency as a discrete process generally defined using the Euler scheme:

³ We actually simulate 105 trading days and then cut off the first five trading days.

$$p_t - p_{t-1} = F_t,$$

where F_t is the time-dependent price generator. Generally, the drift is insignificant for high-frequency data.

4.1 Normal price movements

The most intuitive price generation process uses an iid normal distribution

$$p_t - p_{t-1} = \sigma Z_t,$$

where $Z_t = Z \propto N(0,1)$ and σ is a constant. This is the first intraday volatility pattern we employ.

4.2 Intraday Volatility Patterns

The flat intraday volatility pattern is, however, not close to real data. Therefore, we mimic the well-known U-shaped volatility pattern, which says that price time series show a significant increase in volatility at the beginning and end of the trading day. We implement three different intraday volatility patterns. The purpose is to test the behavior of the indicators under these intraday volatility patterns as well as to compare them over the broadest possible range of situations. The four different specifications for intraday volatility patterns further serve as a testing ground for a proper understanding of price-jump indicators.

4.2.1 Step function I

The second intraday volatility pattern is based on the assumption that volatility undergoes a two-regime switching process, where one regime is at the beginning and end of the trading day, while the other regime is at the middle of the trading day. Namely, we assume a price-generating process given as

$$p_t - p_{t-1} = \sigma_t Z_t,$$

where the volatility σ_t governs the two-regime process and is defined as

$$\sigma_t = \begin{cases} \sigma_{high} & t \in [0, \alpha \cdot Day) \\ \sigma_{low} & t \in [\alpha \cdot Day, \beta \cdot Day) \\ \sigma_{high} & t \in [\beta \cdot Day, Day] \end{cases},$$

where $\sigma_{low} < \sigma_{high}$ and α and β are parameters governing periods with different volatility regimes. Compared to the previous case, there is an artificial ‘‘jump’’ in volatility at the moment when the volatility changes from σ_{low} to σ_{high} .

4.2.2 Step function II

The third intraday volatility pattern is an extension of the previous one. We employ a four-level volatility regime to mimic the U-shaped volatility smile in a more subtle way. Such a definition also partially gets rid of the artificial jumps at the corners where volatility regimes change. Namely, we assume the price generating process is given as

$$p_t - p_{t-1} = \sigma_t Z_t,$$

with

$$\sigma_t = \begin{cases} 3\sigma_{high} & t \in [0, \alpha \cdot Day) \\ 2\sigma_{high} & t \in [\alpha \cdot Day, \beta \cdot Day) \\ 1\sigma_{high} & t \in [\beta \cdot Day, \gamma \cdot Day] \\ \sigma_{low} & t \in [\gamma \cdot Day, \delta \cdot Day] \\ 1\sigma_{high} & t \in [\delta \cdot Day, \varepsilon \cdot Day] \\ 2\sigma_{high} & t \in [\varepsilon \cdot Day, \phi \cdot Day) \\ 3\sigma_{high} & t \in [\phi \cdot Day, Day) \end{cases},$$

where $\sigma_{low} < \sigma_{high}$ and parameters $\alpha, \beta, \gamma, \delta, \varepsilon$ and ϕ define periods with different volatility regimes. In this case, the volatility pattern is smoother and mimics the empirical patterns better.

4.2.3 Linear-like smooth smile

The fourth volatility pattern mimics the U-shaped volatility smile more closely. We use three linear functions, which ensure a smooth transition in volatility between different parts of the trading day. Namely, we assume the price-generating process is given as

$$p_t - p_{t-1} = \sigma_t Z_t$$

with

$$\sigma_t = \begin{cases} 3\sigma_{high} - \frac{(3\sigma_{high} - \sigma_{low})}{\alpha \cdot Day}(t) & t \in [0, \alpha \cdot Day) \\ \sigma_{low} & t \in [\alpha \cdot Day, \beta \cdot Day) \\ 3\sigma_{low} - \frac{(3\sigma_{high} - \sigma_{low})}{(1-\beta) \cdot Day}(t) & t \in [\beta \cdot Day, Day] \end{cases}$$

and $\sigma_{low} < \sigma_{high}$. The parameters define periods with different volatility. $3\sigma_{high}$ was chosen to be able to compare this pattern with the previous one.

4.3 Volatility Specifications

We employ the four different intraday volatility patterns defined above with the parameters as follows.

Volatility Pattern A: The first type of intraday volatility pattern consists of a basic homogenous iid normal process, namely

$$p_t - p_{t-1} = \sigma Z_t,$$

where we use $\sigma = 0.0004$, which corresponds to the values observed in the real data (used in the literature and based on the annual realized volatility).

Volatility Pattern B: The second intraday volatility pattern is given as

$$p_t - p_{t-1} = \sigma_t Z_t$$

with

$$\sigma_t = \begin{cases} \sigma_{high} & t \in [0, Day / 4) \\ \sigma_{low} & t \in [Day / 4, 3Day / 4) \\ \sigma_{high} & t \in [3Day / 4, Day] \end{cases},$$

using the values $\sigma_{low} = 0.0001$ and $\sigma_{high} = 0.0004$.

Volatility Pattern C: The third intraday volatility pattern is defined as

$$p_t - p_{t-1} = \sigma_t Z_t$$

with volatility defined as

$$\sigma_t = \begin{cases} 3\sigma_{high} & t \in [0, 45 \text{ min}) \\ 2\sigma_{high} & t \in [45 \text{ min}, 90 \text{ min}) \\ 1\sigma_{high} & t \in [90 \text{ min}, 135 \text{ min}) \\ \sigma_{low} & t \in [135 \text{ min}, 285 \text{ min}) \\ 1\sigma_{high} & t \in [285 \text{ min}, 330 \text{ min}) \\ 2\sigma_{high} & t \in [330 \text{ min}, 375 \text{ min}) \\ 3\sigma_{high} & t \in [375 \text{ min}, 420 \text{ min}] \end{cases}.$$

The 45-minute step corresponds approximately to $Day/9$, thus the trading day has three periods of approximately the same duration: the first at the beginning of the day with decreasing volatility, the second at the middle of the day with increasing volatility and the third at the end of the day with increasing volatility. We use $\sigma_{low} = 0.0001$ and $\sigma_{high} = 0.0002$.

Volatility Pattern D: This pattern prevents a possible critique that could emerge in the previous cases: whenever we change the volatility regime, we introduce an artificial jump of average size σ_{high} . This can have a negative effect on the performance of some indicators; therefore we make the transition smoother. Thus, volatility is defined as

$$\sigma_t = \begin{cases} 3\sigma_{high} - \frac{(3\sigma_{high} - \sigma_{low})}{135\text{min}}(t) & t \in [0, 135 \text{ min}) \\ \sigma_{low} & t \in [135 \text{ min}, 285 \text{ min}), \\ \sigma_{low} - \frac{(3\sigma_{high} - \sigma_{low})}{135\text{min}}(t - 285) & t \in [285 \text{ min}, 420 \text{ min}] \end{cases}$$

where $\sigma_{low} < \sigma_{high}$. We use $\sigma_{low} = 0.0001$ and $\sigma_{high} = 0.0002$.

4.4 Price-Jump Specification

This study focuses on price jumps, so we extend the price movements defined above with non-normal price jumps. The Euler scheme for price evolution with price jumps is defined as

$$p_t - p_{t-1} = F_t = \sigma_t Z_t + J \cdot j_t,$$

where $\sigma_t Z_t$ is the term defined above and $J \cdot j_t$ is the term generating price jumps. We conveniently define j_t as a Poisson process with a rate of price-jump arrival λ_j :

$$j_t = \begin{cases} 0 & p_j \\ 1 & 1 - p_j \end{cases},$$

where $p_j = e^{-\lambda_j}$ and parameter J govern the size of the jumps. For single-size price jumps, $J = \pm J_{param}$, where both signs have the same probability of occurring. In the most sophisticated cases, parameter J can have a value from any statistical distribution.

Due to the independence of increments, the probability to observe n jumps at a time step is given as

$$P(N \text{ of Jumps} = n) = \frac{e^{-\lambda} (\lambda)^n}{n!}.$$

By definition, we assume that only one price jump per time step can occur and thus we define first the probability that no price jump will occur as $P(N \text{ of Jumps} = 0) = e^{-\lambda}$ and the probability that one price jump will occur as a complement value $P(N \text{ of Jumps} = 1) = 1 - e^{-\lambda}$.

We employ five different specifications of price jumps. These five specifications are combined with the four different groups of indicators. Thus we will have twenty different price time series (excluding four different time series without price jumps).⁴

Price Jumps 1–3: The first three price-jump specifications have the same rate of arrival and a constant size of jump $J = \pm const$. Both signs occur with the same probability.

Price Jump 1: We employ combinations of $J = 5\sigma_{jump}$ and $\lambda = 5/N_{Day}$.

Price Jump 2: We employ combinations of $J = 7\sigma_{jump}$ and $\lambda = 5/N_{Day}$.

Price Jump 3: We employ combinations of $J = 9\sigma_{jump}$ and $\lambda = 5/N_{Day}$.

The parameter $\sigma_{jump} = 0.0004$ and N_{day} means the number of minutes per trading day.

Price Jumps 4–5: The next two price-jump specifications use a uniform distribution to select the size of price jumps. We select price jumps from a given distribution $J \propto \pm U(a, b)$, with $0 < a < b$, and both signs occur with the same probability. We use the following specifications:

Price Jump 4: $J \propto \pm U(5\sigma_{jump}, 9\sigma_{jump})$ and $\lambda_j = 5/N_{Day}$.

Price Jump 5: $J \propto \pm U(5\sigma_{jump}, 9\sigma_{jump})$ and $\lambda_j = 15/N_{Day}$.

The parameter for volatility is chosen as $\sigma_{jump} = 0.0004$ and N_{day} was defined above.

5 Comparison Strategy

The goal of the simulation procedure is to compare price-jump indicators with each other, understand their properties and select the most appropriate indicator for real data.

⁴ An alternative approach to estimate price jumps is to assume that the error term follows a given stochastic distribution and combine it with a non-homogenous Poisson term describing a jump. Then, Score Method of Moments or Simulated Method of Moments could be employed to estimate the parameters of the model. See for example Jiang and Oomen (2007), who estimate an affine jump diffusion model for a series of returns from the S&P 500 index. However, this approach relies on the proper specification of the underlying model, including the intraday volatility pattern as well as the distribution. Thus, this approach is not appropriate for our analysis.

5.1 Price Jump Indicators

We employ the following extensive list of price-jump indicators that are defined in the Appendix A1:

1. Centiles as defined in A1.1: The price jump is identified as those returns below the 0.5th centile or above the 99.5th centile. Centiles are calculated for the entire sample.
2. Block-centiles as defined in A1.2: The price jump is identified as those returns below the 0.5th centile or above the 99.5th centile. Every trading day is divided into 15-minute blocks and centiles are calculated for every block separately for the entire sample.
3. $Z_{RJ,TP}$ as defined in A2.1 with a 99% confidence interval (CI) and length of moving window $n = 60$.
4. $Z_{RJ,TP}$ as defined in A2.1 with a 99% CI and $n = 120$.
5. Improved $Z_{RJ,TP}$ as defined in A2.2 with a 99% CI and $n = 60$.
6. Improved $Z_{RJ,TP}$ as defined in A2.2 with a 99% CI and $n = 120$.
7. ξ -statistics as defined in A2.3 with a 99% CI and $n = 60$.
8. ξ -statistics as defined in A2.3 with a 99% CI and $n = 120$.
9. JO_{Ratio} as defined in A3.1 with a 99% CI and $n = 60$.
10. JO_{Ratio} as defined in A3.1 with a 99% CI and $n = 120$.
11. Improved JO_{Ratio} as defined in A3.2 with a 99% CI and $n = 60$.
12. Improved JO_{Ratio} as defined in A3.2 with a 99% CI and $n = 120$.
13. Price-jump index as defined in A4.1: The price jump is identified as those returns with $pji > 4$ and $n = 120$.
14. Price-jump index as defined in A4.1: The price jump is identified as those returns with $pji > 4$. $n = 420$.

5.2 Artificial Time Series

We employ a Monte Carlo simulation technique to simulate an artificial time series with price jumps. We simulate all the combinations of four different intraday volatility patterns (specified in section 4.3) and five different price-jump specifications (specified in section 4.4), thus there are 20 different time series in total. Every trading day is sampled at a one-minute frequency, starting at 9:01 and ending at 16:00: seven hours in total, which gives 420 trading minutes per trading day. We further match the end of the trading day with the beginning of the next trading day and thus produce continuous time series.

We simulate 105 trading days and define price-jump indicators. Then, we cut off the first five days, which serve to settle down the simulation as well as produce the necessary observations for the moving windows. In addition, the Jiang-Oomen statistics-based indicators require absolute levels. For that purpose, we set an initial value $p_0 = 100$ and produce price levels instead of returns.

5.3 Relative Comparison of Price Jump Indicators

In the last step we perform an extensive comparison of the performance of the different price-jump indicators. We follow the methodology outlined in section 3 based on the McNemar (1947) test.

6 Results

We compared 14 different price-jump indicators with respect to false positive and false negative probabilities. We present all the details of the comparison in Appendix B. In Table 2 we present a summary of our results: the number of cases when a given price-jump indicator dominates the other indicators with respect to both false positive and false negative probabilities. Several times there were two indicators that were dominating the other indicators. In such a case, we counted both indicators as dominating the given simulated specification.

Table 2: Summary of the analysis based on false positive and false negative probabilities

	Indicator No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14
False positive	No. of dominances	13	3	0	0	0	0	1	2	0	0	0	0	0	0
False negative	No. of dominances	2	0	0	0	0	0	10	15	0	0	0	0	0	0

In the case of the false positive probability—the false identification of non-jump cases—the best indicator seems to be the indicator No. 1 based on centiles. This indicator dominated others the most often. In addition, there are other indicators, which performs in some specifications well, namely No. 2—the one based on block-centiles—and Nos. 7 and 8 based on the ξ -statistics with 99% CI and $n=60$ or $n = 120$, respectively.

The other case, false negative probability—jumps that occur are not identified—shows that the best performing statistics is the indicator No. 8, the ξ -statistics with 99% CI and $n = 120$. In addition, the analysis shows that even the version with time window $n=60$ performs well since these two statistics are in many cases statistically indistinguishable.

The analysis further reveals that the performance of price-jump indicators is not homogeneously distributed among all the indicators but rather their performance is dominated by a few best indicators. This can be further seen in the results, where most of the time when one indicator dominates another, it dominates it at the highest significance level.

7 Conclusion

We performed an extensive simulation study to compare the relative performance of a broad class of price-jump indicators with respect to false positive and false negative probabilities. We simulated twenty different time-series specifications with different intraday noise volatility patterns and price-jump specifications and using these artificial time series, we employed the double McNemar test and compared fourteen different price-jump indicators that are widely used in the literature. We compared them with respect to false positive and false negative probabilities. The results suggest large differences among the indicators in terms of their performance. In the case of false positive probability, the best-performing price-jump indicator is the one based on thresholding with respect to centiles. In the case of false negative probability, the best indicator is the one based on bipower variation. Significant differences among the indicators further confirms the fact that any meta-analysis based on different price-jump indicators is not possible since the indicators tend to perform in very different ways.

References

Ait-Sahalia, Y., 2004. Disentangling Diffusion from Jumps. *Journal of Financial Economics*, Vol. 74, pp. 487-528.

Ait-Sahalia, Y., Cacho-Diaz, J., and Hurd, T., 2009. Portfolio Choice with Jumps: A Closed Form Solution. *Annals of Applied Probability*, Vol. 19, pp. 556-584.

Ait-Sahalia, Y. and Jacod, J., 2009a. Testing for Jumps in a Discretely Observed Process. *Annals of Statistics*, Vol. 37, pp. 184-222.

Ait-Sahalia, Y. and Jacod, J., 2009b. Estimating the Degree of Activity of Jumps in High Frequency Data. *Annals of Statistics*, Vol. 37, pp. 2202-2244.

Andersen, T., Benzoni, L., and Lund, J., 2002. An Empirical Investigation of Continuous-Time Equity Return Models. *Journal of Finance*, Vol. 57, Issue 3, pp. 1239-1284.

Andersen, T., Bollerslev, T., and Diebold, F., 2007. Roughing It Up: Including Jump Components in the Measurement, Modeling, and Forecasting of Return Volatility. *Review of Economics and Statistics*, Vol. 89, Issue 4, pp. 701-720.

Bajgrowicz, P. and Scaillet, O., 2010. Detecting Spurious Jumps in High-Frequency Data, available at SSRN: <http://ssrn.com/abstract=1343900>.

Barndorff-Nielsen, O.E., Graversen, S.E., Jacod, J., Podolskij, M., and Shephard, N., 2004. A Central Limit Theorem for Realised Power and Bipower Variations of Continuous Semimartingales, No 2004-W29, Economics Papers, Economics Group, Nuffield College, University of Oxford.

Barndorff-Nielsen, O., Hansen, P., Lunde, A., and Shephard, N., 2006. Designing Realised Kernels to Measure the Ex-post Variation of Equity Prizes in the Presence of Noise. Working Paper, Nuffield College, Oxford University.

Barndorff-Nielsen, O. and Shephard, N., 2004. Power and Bipower Variation with Stochastic Volatility and Jumps. *Journal of Financial Econometrics*, Vol. 2, pp. 1-37.

Barndorff-Nielsen, O. and Shephard, N., 2006. Econometrics of Testing for Jumps in Financial Economics Using Bipower Variation. *Journal of Financial Econometrics*, Vol. 4, pp. 1-30.

Bollerslev, T., Law, T.H., and Tauchen, G., 2008. Risk, Jumps, and Diversification. *Journal of Econometrics*, Vol. 144, Issue 1, pp. 234-256.

Bouchaud, J-P., 2002. An Introduction to Statistical Finance. *Physica A*, Vol. 313, Nos. 1-2, pp. 238-251.

Bouchaud, J-P., Kockelkoren, J., and Potters, M., 2004. Random Walks, Liquidity Molasses and Critical Response in Financial Markets. Finance (CFM) Working Paper Archive 500063, Finance, Capital Fund Management.

Broadie, M. and Jain, A., 2008. The Effect of Jumps and Discrete Sampling on Volatility and Variance Swaps. *International Journal of Theoretical and Applied Finance*, Vol. 11, No. 8, pp. 761-797.

Carr, P. and Wu, L., 2010. Stock Options and Credit Default Swaps: A Joint Framework for Valuation and Estimation. *Journal of Financial Econometrics*, Vol. 8, No. 4, pp. 409-449.

Duffie, D., Pan, J., and Singleton, K., 2000. Transform Analysis and Asset Pricing for Affine Jump Diffusions. *Econometrica*, Vol. 68, pp. 1343-1376.

Eryigit, M., Cukur, S., and Eryigit, R., 2009. Tail Distribution of Index Fluctuations in World Market. *Physica A*, Vol. 388, pp. 1879-1886.

Gibbons, J.D., 1997. *Nonparametric Methods for Quantitative Analysis (3rd ed.)*, American Sciences Press, Columbus, Ohio.

Gopikrishnan, P., Plerou, V., Nunes Amaral, L. A., Meyer, M., and Stanley, H. E., 1999. Scaling of the Distribution of Fluctuations of Financial Market Indexes. *Physical Review E*, Vol. 60, No. 5, pp. 5305-5316.

Hanousek, J., 2000. Testing of (Non-nested) Choice Models via Accuracy of Prediction: A Non-Parametric Approach, *Central European Journal of Operations Research*, Vol. 8, Issue 3, pp. 259-264.

Huang, X. and Tauchen, G., 2005. The Relative Contribution of Jumps to Total Price Variance. *Journal of Financial Econometrics*, Vol. 3, Issue 4, pp. 456-499.

Jarrow, R.A. and Rosenfeld, E.R., 1984. Jump Risks and the Intertemporal Capital Asset Pricing Model. *Journal of Business*, Vol. 57, No. 3, pp. 337-351.

Jiang, J., Li, W., Cai, X., and Wang, Q. A., 2009. Empirical Study of Recent Chinese Stock Market. *Physica A*, Vol. 388, pp. 1893-1907.

Jiang, G. and Oomen, R., 2007. Estimating Latent Variables and Jump Diffusion Models Using High-Frequency Data. *Journal of Financial Econometrics*, Vol. 5, No. 1, pp. 1-30.

Jiang, G. and Oomen, R., 2008. Testing for Jumps When Asset Prices Are Observed with Noise: A Swap Variance Approach. *Journal of Econometrics*, Vol. 144, No. 2, pp. 352-370.

Johannes, M., 2004. The Statistical and Economic Role of Jumps in Interest Rates. *Journal of Finance*, Vol. 59, pp. 227-260.

Joulin, A., Lefevre, A., Grunberg, D., and Bouchaud, J.-P., 2008. Stock Price Jumps: News and Volume Play a Minor Role. *Quantitative Finance Papers by arXiv.org*, 0803.1769, <http://arxiv.org/abs/0803.1769>.

Lahaye, J., Laurent, S., and Neely, C.J., 2010. Jumps, Cojumps and Macro Announcements. *Journal of Applied Econometrics*, n/a. doi: 10.1002/jae.1149.

Lee, S.S. and Mykland, P.A., 2008. Jumps in Financial Markets: A New Nonparametric Test and Jump Dynamics. *The Review of Financial Studies*, Vol. 21, No. 6, pp. 2535-2563.

Liu, J., Longstaff, F., and Pan, J., 2003. Dynamic Asset Allocation with Event Risk. *Journal of Finance*, Vol. 58, pp. 231-259.

McNemar, Q., 1947. Note on the Sampling Error of the Difference between Correlated Proportions or Percentages. *Psychometrika*, Vol. 12, No. 2, pp. 153-157.

Merton, R. C., 1976. Option Pricing when Underlying Stock Returns are Discontinuous. *Journal of Financial Economics*, Vol. 3, Nos. 1-2, pp. 125-144.

Nietert, B., 2001. Pricing Relevance of Stock Price Jumps: CAPM and Variance Decomposition. EFMA 2001 Lugano Meetings, available at SSRN: <http://ssrn.com/abstract=268546>.

Novotny, J., 2010. Were Stocks during the Financial Crisis More Jumpy: A Comparative Study, CERGE-EI Working Paper No. 416.

Nyberg, P. and Wilhelmsson, A., 2009. Measuring Event Risk. *Journal of Financial Econometrics*, Vol. 7, No. 3, pp. 265-287.

Pan, J., 2002. The Jump-risk Premia Implicit in Options: Evidence from an Integrated Time-series Study. *Journal of Financial Economics*, Vol. 63, pp. 3-50.

Shiller, R.J., 2005. *Irrational Exuberance*. Princeton University Press, Princeton.

Stanley, H. E. and Mantegna, R. N., 2000. *An Introduction to Econophysics*. Cambridge University Press, Cambridge.

Taleb, N., 2007. *The Black Swan: The Impact of the Highly Improbable*. Random House, New York.

Wood, R., McNish, T., and Ord, J., 1985. An Investigation of Transaction Data for NYSE Stocks. *Journal of Finance*, Vol. 25, 723-739.

Appendix A: Price-Jump Indicators – The Details

In this section, we provide technical details for all of the price-jump indicators we tested.

A1 Group 1: Ait-Sahalia

This class of indicator assumes that the underlying price increment process is given as $\Delta S = \sigma \Delta X + \Delta J$, where the price increment means $\Delta S = S_t - S_{t-\Delta t}$, where we assume that we observe the realization of prices in equidistant time steps Δt , i.e., ΔS denotes a price change over the time interval Δt . In this definition, X corresponds to the Brownian motion and J is a β -stable process. The increments of the two components can be expressed as $\Delta X = (\Delta t)^{1/2} X_1$ and $\Delta J = (\Delta t)^{1/\beta} J_1$, where the equalities are equalities in distribution. In this specification, X corresponds to the Brownian motion and J is a β -stable process.

The different magnitudes in the two components can be used to discriminate between the noise components and the big price jumps coming solely from the J -process.⁵ The big price jumps cause $\Delta S = \Delta J$ (in distribution), while in the presence of no big price jumps, which is most of the time, $\Delta S = \sigma(\Delta t)^{1/2} X_1$. Therefore, we can, for a given Δt , choose a threshold value equal to $\alpha(\Delta t)^\gamma$, with $\alpha > 0$ and $\gamma \in (0, 1/2)$, such that if $\Delta S > \alpha(\Delta t)^\gamma$ then ΔS is at a given moment dominated by J with a certain probability.

This argument can be inverted. Assuming that we know the rate of the arrival of big jumps, we can easily construct a threshold based on the centile value. Therefore we will use centiles as a threshold to discriminate price jumps from the noise. Using centiles, however, can produce biased results due to the intraday volatility patterns. The intraday volatility pattern means that $\sigma = \sigma(t)$. In addition, the J -process can also be different either across different phases of the trading day or across different trading days. To account for the former, we divide every trading day into several trading blocks and assume that inside every trading block the price process is constant no matter the trading day. In this case, we can apply the same logic block by block. Namely, we calculate the centiles for the same block over different trading days and threshold price jumps for every trading block separately.

⁵ The J -process contributes to a large amount of small price jumps; however, we want to focus on big price jumps only. The goal is not to completely determine the properties of the J -process but rather to determine how to discriminate extreme price movements.

A1.1 Global Centiles

We define price jumps as those returns that are higher/lower than a given upper/lower centile. Centiles are calculated based on the observation of the entire sample, where we use the 99.5th centile as the upper threshold and the 0.5th centile as the lower threshold.

A1.2 Centiles over Block-Windows

To compensate for intraday volatility, we divide every trading day into several 15 minute-long blocks. Then, we apply the procedure defined above for every trading block separately, i.e., we calculate the upper/lower threshold for every trading block independently and then define the price jumps as those price movements that are higher/lower than the corresponding threshold values.

A2 Group 2: Bipower Variation

The two different measures for variation, as defined by Barndorff-Nielsen and Shephard (2004, 2006) are Realized Variation defined as $RV_t = \sum_{i=2}^n r_i^2$ and Bipower Variation defined as

$$BV_t = \mu_1^{-2} \left(\frac{n-1}{n-2} \right) \sum_{i=3}^n |r_i| |r_{i-1}|, \quad \text{with} \quad \mu_\alpha = E(|Z|^\alpha) \quad \text{for} \quad Z \sim N(0,1), \quad \text{or generally}$$
$$\mu_\alpha = 2^{\alpha/2} \Gamma\left(\frac{\alpha+1}{2}\right) / \sqrt{\pi}.$$

A2.1 The Max-adjusted Statistics

The difference between the two variations is the key ingredients; however, one also needs to estimate the conditional standard deviation $\int \sigma^4$. There are at least two possible ways to estimate this: Anderson, Bollerslev and Diebold (2007) introduced tripower quarticity,

$$TP_j = n \mu_{4/3}^{-3} \frac{n-1}{n-3} \sum_{i=j-n+4}^j |r_i|^{4/3} |r_{i-1}|^{4/3} |r_{i-2}|^{4/3},$$
 to measure the conditional standard deviation, while

Barndorff-Nielsen and Shephard (2004, 2006) used Quadpower Quarticity,

$$QP_j = n \mu_1^{-4} \frac{n-1}{n-4} \sum_{i=j-n+5}^j |r_i| |r_{i-1}| |r_{i-2}| |r_{i-3}|.$$

Barndorff-Nielsen and Shephard (2004, 2006) then proposed several different asymptotically equal statistics to estimate the presence of price jumps.

According to Huang and Tauchen (2005), the best statistics is $Z_{RJ,TP}$ defined as:

$$Z_{RJ,TP} = \frac{RJ}{\sqrt{\left(\left(\frac{\pi}{2}\right)^2 + \pi - 5\right)\left(\frac{1}{n}\right) \max\left(1, \frac{TP}{BV^2}\right)}}$$

with $RJ_j = (RJ_j - BJ_j) / RJ_j$. The null hypothesis states that there is no jump in a given period. If the statistics exceeds the critical value $\Phi^{-1}(\alpha)$, then we reject the null hypothesis of no price jump at confidence level α .

Realized Variation and Bipower Variation are forward-looking, however, we need a backward-looking specification re-defined as:

$$RV_j = \sum_{i=j-n+2}^j r_i^2,$$

$$BV_j = \mu_1^{-2} \left(\frac{n-1}{n-2}\right) \sum_{i=j-n+3}^j |r_i| |r_{i-1}|.$$

The statistics thus refer to a window of length n ending at time step j . Thus, observing a significant jump at time step j means that somewhere in the window of length n ending at time step j was at least one significant jump. Thus, the change between periods with no price jump and periods with a price jump can serve as an indicator for the moments when jumps happened the first time. This also assumes that the average time between two jumps will be much larger than the window used in this statistics. On the other hand, a very short time window skews the results with a small-sample bias. Since we assume more than one price jump per day, we employ $n = 60$ and $n = 120$.

The indicator for a price jump is defined as follows: price jumps are those prices for which $Z_{t-1} \leq \Phi^{-1}(\alpha)$ and $Z_t > \Phi^{-1}(\alpha)$. By definition, the indicator cannot distinguish two consecutive steps, otherwise we would have to work with the absolute levels of the statistics.

A2.2 Max-adjusted Statistics: Improved Identification Method

The improvement works as described in the main section, namely returns identified as price jumps are replaced by the average value calculated over the same length as was used for identification. The replaced value is clearly not a price jump, otherwise the price jump would

not be identified as a price jump. Therefore, we define a pair of improved indicators based on the above-defined Max-adjusted statistics with $n = 60$ and $n = 120$.

A2.3 Lee-Mykland

The statistics of Lee and Mykland (2008) is based on bipower variation and is given as

$$L(i) = \frac{r_i}{\hat{\sigma}(i)}, \text{ with } \hat{\sigma}^2(i) = \frac{1}{n-2} \sum_{j=i-n+2}^{i-1} |r_j| |r_{i-1}|. \text{ Then}$$

$$\frac{\max_i |L(i)| - C_n}{S_n} \rightarrow \xi,$$

where ξ has a cumulative distribution function $P(\xi \leq x) = \exp(-e^{-x})$, and the two constants are given as $C_n = \frac{(2 \log n)^{1/2}}{c} - \frac{\log z + \log(\log n)}{2c(2 \log n)^{1/2}}$, $S_n = \frac{1}{c(2 \log n)^{1/2}}$ and $c = \sqrt{2} / \sqrt{\pi}$. Whenever the ξ -statistics exceeds the critical value ξ_{CV} , we reject the null hypothesis of no price jump at time t_i .

Lee and Mykland recommends $n_{15-min} = 156$ and $n_{5-min} = 270$. In our analysis, we use $n = 60$ and $n = 120$.

A3 Group 3: Jiang-Oomen Statistics

The Jiang and Oomen (2008) statistics is based on Swap Variance defined as

$$SwV = 2 \sum_{i=2}^n (R_i - r_i),$$

where $R_i = \frac{P_i - P_{i-1}}{P_i}$, $P_i = \exp(p_i)$ and $r_i = p_i - p_{i-1}$. The authors claim that employing swap variance further amplifies the contribution coming from price jumps and thus makes the estimator less sensitive to intraday variation.

A3.1 Jiang-Oomen Statistics-based Price-Jump Indicator

The Jiang-Oomen statistics is defined as

$$JO_{Ratio} = \frac{nBV}{\sqrt{\Omega_{SwV}}} \left(1 - \frac{RV}{SwV} \right),$$

where the Realized Variation RV and the Bipower Variation BV are defined as above. The statistics is asymptotically equal to $z \sim N(0, 1)$ and tests the null hypothesis that a given window does not contain any price jump. The indicator for a price jump is defined as those price movements for which $JO_{t-1} \leq \Phi^{-1}(\alpha)$ and $JO_t > \Phi^{-1}(\alpha)$. The same comments as for the Max-adjusted statistics apply. We use two price-jump indicators with $n = 60$ and $n = 120$.

A3.2 Jiang-Oomen Statistics: Improved Identification Method

We use the same improvement technique as in section A2.1 and define two improved indicators based on the Jiang-Oomen statistics with $n = 60$ and $n = 120$.

A4 Group 4: Statistical Finance

The scaling properties of returns can be studied using different techniques (see Stanley and Mantegna, 2000, and references therein), where we employ the price-jump index, as defined by Joulin et al. (2008), for this study.

A4.1 Price-Jump Index

The price-jump index is defined as

$$Pji_i = \frac{|r_i|}{\frac{1}{n} \sum_{j=i-n+1}^i |r_j|},$$

where n governs the length of the moving window over which we normalize the absolute returns at a given time moment. The empirical observations suggest (Joulin et al., 2008) that the scaling properties behave as $P(pji > s) \sim s^{-\alpha}$, therefore we define a price jump as those price returns where the price-jump index exceeds a given threshold \hat{s} . In our analysis, we choose $\hat{s} = 4$ and $n = 120$ and $n = 420$.

Appendix B: Simulation Results

This appendix summarizes all the results from the simulations. First, we compare the price-jump indicators with respect to the false positive probability and then with respect to the false negative probability. We simulate 20 different combinations of intraday volatility patterns and price jumps, as they are defined in the preceding sections. Table B1 contains the notation for the combinations used in the tables below.

Table B1: Notation for combinations of different intraday volatility patterns and price jumps

	Volatility Pattern A	Volatility Pattern B	Volatility Pattern C	Volatility Pattern D
Price Jump 1	A1	B1	C1	D1
Price Jump 2	A2	B2	C2	D2
Price Jump 3	A3	B3	C3	D3
Price Jump 4	A4	B4	C4	D4
Price Jump 5	A5	B5	C5	D5

In the tables below, the different price-jump indicators are denoted by the numbers introduced in section 5.1.

Table B1 presents the results of every type of false probability and every combination of intraday volatility pattern and price-jump specification. To eliminate the necessity of having the same note describe the contents of each table, we describe how the tables should be interpreted here. Each table presents a pair-wise comparison of price jumps as denoted above. The price-jump indicator corresponding to a row is denoted by A, the price-jump indicator corresponding to a column is denoted as B. Therefore, whenever the entry in the table contains A, the row indicator dominates in performance the one in the column, and vice versa for B. We use the conventional * for 90%, ** for 95% and *** for 99% confidence levels. The equals symbol (=) means that we cannot reject the null hypothesis that both indicators are equal with respect to a given false probability.

B1 False positive probability: Indicator predicts a jump that does not exist

Table B2: Combination A1

A\B	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1														
2	B**													
3	B**	B**												
4	B**	B**	A**											
5	B**	B**	B**	B**										
6	B**	B**	=	B**	=									
7	B**	A**	A**	A**	A**	A**								
8	A**	A**	A**	A**	A**	A**	A**							
9	B**	B**	A**	A**	A**	A**	B**	B**						
10	B**	B**	A**	A**	A**	A**	B**	B**	A**					
11	B**	B**	A**	A**	A**	A**	B**	B**	B**	B**				
12	B**	B**	A**	A**	A**	A**	B**	B**	B**	B**	A**			
13	B**	A**	A**	A**	A**	A**	B**	B**	A**	A**	A**	A**		
14	B**	A**	A**	A**	A**	A**	B**	B**	A**	A**	A**	A**	A**	

Table B3: Combination B1

A\B	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1														
2	B***													
3	B***	B***												
4	B***	B***	A***											
5	B***	B***	B***	B***										
6	B***	B***	B***	B***	=									
7	B***	B***	A***	A***	A***	A***								
8	B***	B***	A***	=	A***	A***	B***							
9	B***	B***	A***	A***	A***	A***	=	A***						
10	B***	B***	A***	A***	A***	A***	A***	A***	A***					
11	B***	B***	A***	B***	A***	A***	B***	B**	B***	B***				
12	B***	B***	A***	A***	A***	A***	B***	A***	B***	B***	A***			
13	B***	B***	B***	B***	A***	A***	B***	B***	B***	B***	B***	B***		
14	B***	B***	A***	=	A***	A***	B***	=	B***	B***	A**	B***	A***	

Table B4: Combination C1

A\B	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1														
2	B***													
3	B***	B***												
4	B***	B***	=											
5	B***	B***	B***	B***										
6	B***	B***	B***	B***	=									
7	B***	B***	A***	A***	A***	A***								
8	B***	B***	A***	A***	A***	A***	B***							
9	B***	B***	A***	A***	A***	A***	B***	A***						
10	B***	B***	A***	A***	A***	A***	B***	A***	A***					
11	B***	B***	A***	A***	A***	A***	B***	B***	B***	B***				
12	B***	B***	A***	A***	A***	A***	B***	B**	B***	B***	A***			
13	B***	B***	B***	B***	A***	A***	B***	B***	B***	B***	B***	B***		
14	B***	B***	B***	B***	A***	A***	B***	B***	B***	B***	B***	B***	B***	

Table B5: Combination D1

A\B	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1														
2	B***													
3	B***	B***												
4	B***	B***	=											
5	B***	B***	B***	B***										
6	B***	B***	B***	B***	B**									
7	B***	B***	A***	A***	A***	A***								
8	B***	B***	A***	A***	A***	A***	B***							
9	B***	B***	A***	A***	A***	A***	B***	B***						
10	B***	B***	A***	A***	A***	A***	B***	A**	A***					
11	B***	B***	A***	A*	A***	A***	B***	B***	B***	B***				
12	B***	B***	A***	A***	A***	A***	B***	B***	B***	B***	A***			
13	B***	B***	B***	B***	A***	A***	B***	B***	B***	B***	B***	B***		
14	B***	B***	B***	B***	A***	A***	B***	B***	B***	B***	B***	B***	B***	

Table B6: Combination A2

A/B	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1														
2	B***													
3	B***	B***												
4	B***	B***	A***											
5	B***	B***	B***	B***										
6	B***	B***	B***	B***	=									
7	B***	A***	A***	A***	A***	A***								
8	=	A***	A***	A***	A***	A***	A***							
9	B***	B***	A***	A***	A***	A***	B***	B***						
10	B***	B***	A***	A***	A***	A***	B***	B***	A***					
11	B***	B***	A***	A***	A***	A***	B***	B***	B***	B***				
12	B***	B***	A***	A***	A***	A***	B***	B***	B***	B***	A***			
13	B***	A***	A***	A***	A***	A***	B***	B***	A***	A***	A***	A***		
14	B***	A***	A***	A***	A***	A***	B***	B***	A***	A***	A***	A***	A***	A**

Table B7: Combination B2

A/B	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1														
2	B***													
3	B***	B***												
4	B***	B***	A***											
5	B***	B***	B***	B***										
6	B***	B***	B***	B***	=									
7	B***	B***	A***	B**	A***	A***								
8	B***	B***	A**	B***	A***	A***	B***							
9	B***	B***	A***	A*	A***	A***	A***	A***						
10	B***	B***	A***	A***	A***	A***	A***	A***	A***					
11	B***	B***	A***	B***	A***	A***	B***	=	B***	B***				
12	B***	B***	A***	B*	A***	A***	=	A***	B**	B***	A***			
13	B***	B***	B***	B***	A***	A***	B***	B***	B***	B***	B***	B***		
14	B***	B***	A***	B***	A***	A***	B***	A***	B***	B***	A***	B***	A***	

Table B8: Combination C2

A/B	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1														
2	B***													
3	B***	B***												
4	B***	B***	A***											
5	B***	B***	B***	B***										
6	B***	B***	B***	B***	=									
7	B***	B***	A***	A***	A***	A***								
8	B***	B***	A***	A***	A***	A***	B***							
9	B***	B***	A***	A***	A***	A***	B***	A***						
10	B***	B***	A***	A***	A***	A***	B***	A***	A***					
11	B***	B***	A***	B***	A***	A***	B***	B***	B***	B***				
12	B***	B***	A***	A***	A***	A***	B***	B***	B***	B***	A***			
13	B***	B***	B***	B***	A***	A***	B***	B***	B***	B***	B***	B***		
14	B***	B***	B***	B***	A***	A***	B***	B***	B***	B***	B***	B***	B***	

Table B9: Combination D2

A/B	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1														
2	B***													
3	B***	B***												
4	B***	B***	A***											
5	B***	B***	B***	B***										
6	B***	B***	B***	B***	B***									
7	B***	B***	A***	A***	A***	A***								
8	B***	B***	A***	A***	A***	A***	B***							
9	B***	B***	A***	A***	A***	A***	B***	B***						
10	B***	B***	A***	A***	A***	A***	B***	A***	A***					
11	B***	B***	B**	B***	A***	A***	B***	B***	B***	B***				
12	B***	B***	A***	=	A***	A***	B***	B***	B***	B***	A***			
13	B***	B***	B***	B***	A***	A***	B***	B***	B***	B***	B***	B***		
14	B***	B***	B***	B***	A***	A***	B***	B***	B***	B***	B***	B***	B***	

Table B10: Combination A3

A/B	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1														
2	B***													
3	B***	B***												
4	B***	B***	A***											
5	B***	B***	B***	B***										
6	B***	B***	B***	B***	=									
7	B***	A***	A***	A***	A***	A***								
8	=	A***	A***	A***	A***	A***	A**							
9	B***	B***	A***	B***	A***	A***	B***	B***						
10	B***	B***	A***	A**	A***	A***	B***	B***	A***					
11	B***	B***	B***	B***	A***	A***	B***	B***	B***	B***				
12	B***	B***	A***	B***	A***	A***	B***	B***	=	B**	A***			
13	B***	A***	A***	A***	A***	A***	B***	B***	A***	A***	A***	A***		
14	B***	A***	A***	A***	A***	A***	B**	B***	A***	A***	A***	A***	A**	

Table B11: Combination B3

A/B	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1														
2	B***													
3	B***	B***												
4	B***	B***	A***											
5	B***	B***	B***	B***										
6	B***	B***	B***	B***	B***									
7	B***	B***	A***	B***	A***	A***								
8	B***	B***	B***	B***	A***	A***	B***							
9	B***	B***	A***	B***	A***	A***	A**	A***						
10	B***	B***	A***	=	A***	A***	A***	A***	A***					
11	B***	B***	B***	B***	A***	A***	B***	B***	B***	B***				
12	B***	B***	A***	B***	A***	A***	=	A***	B***	B***	A***			
13	B***	B***	B***	B***	A***	A***	B***	B***	B***	B***	B***	B***		
14	B***	B***	A**	B***	A***	A***	=	A***	B***	B***	A***	B*	A***	

Table B12: Combination C3

A/B	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1														
2	B***													
3	B***	B***												
4	B***	B***	A***											
5	B***	B***	B***	B***										
6	B***	B***	B***	B***	B***									
7	B***	B***	A***	A***	A***	A***								
8	B***	B***	A***	B***	A***	A***	B***							
9	B***	B***	A***	B*	A***	A***	B***	A***						
10	B***	B***	A***	A***	A***	A***	B***	A***	A***					
11	B***	B***	B***	B***	A***	A***	B***	B***	B***	B***				
12	B***	B***	A***	B***	A***	A***	B***	B**	B***	B***	A***			
13	B***	B***	B***	B***	A***	A***	B***	B***	B***	B***	B***	B***		
14	B***	B***	B***	B***	A***	A***	B***	B***	B***	B***	B***	B***	B***	B***

Table B13: Combination D3

A/B	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1														
2	B***													
3	B***	B***												
4	B***	B***	A***											
5	B***	B***	B***	B***										
6	B***	B***	B***	B***	B***									
7	B***	B***	A***	A***	A***	A***								
8	B***	B***	A***	B*	A***	A***	B***							
9	B***	B***	A***	B***	A***	A***	B***	B***						
10	B***	B***	A***	A**	A***	A***	B***	A**	A***					
11	B***	B***	B***	B***	A***	A***	B***	B***	B***	B***				
12	B***	B***	=	B***	A***	A***	B***	B***	B***	B***	A***			
13	B***	B***	B***	B***	A***	A***	B***	B***	B***	B***	B***	B***		
14	B***	B***	B***	B***	A***	A***	B***	B***	B***	B***	B***	B***	B***	B***

Table B14: Combination A4

A/B	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1														
2	B***													
3	B***	B***												
4	B***	B***	A***											
5	B***	B***	B***	B***										
6	B***	B***	B***	B***	=									
7	B***	A***	A***	A***	A***	A***								
8	=	A***	A***	A***	A***	A***	A***							
9	B***	B***	A***	A***	A***	A***	B***	B***						
10	B***	B***	A***	A***	A***	A***	B***	B***	A***					
11	B***	B***	A***	A***	A***	A***	B***	B***	B***	B***				
12	B***	B***	A***	A***	A***	A***	B***	B***	B***	B***	A***			
13	B***	A***	A***	A***	A***	A***	B***	B***	A***	A***	A***	A***		
14	B***	A***	A***	A***	A***	A***	B***	B***	A***	A***	A***	A***	A*	

Table B15: Combination B4

A/B	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1														
2	B***													
3	B***	B***												
4	B***	B***	A***											
5	B***	B***	B***	B***										
6	B***	B***	B***	B***	B*									
7	B***	B***	A***	B***	A***	A***								
8	B***	B***	A***	B***	A***	A***	B***							
9	B***	B***	A***	A***	A***	A***	A***	A***						
10	B***	B***	A***	A***	A***	A***	A***	A***	A***					
11	B***	B***	A***	B***	A***	A***	B***	=	B***	B***				
12	B***	B***	A***	=	A***	A***	A**	A***	B***	B***	A***			
13	B***	B***	B***	B***	A***	A***	B***	B***	B***	B***	B***	B***		
14	B***	B***	A***	B***	A***	A***	B***	A***	B***	B***	A***	B***	A***	

Table B16: Combination C4

A/B	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1														
2	B***													
3	B***	B***												
4	B***	B***	A***											
5	B***	B***	B***	B***										
6	B***	B***	B***	B***	B*									
7	B***	B***	A***	A***	A***	A***								
8	B***	B***	A***	A***	A***	A***	B***							
9	B***	B***	A***	A***	A***	A***	B***	A***						
10	B***	B***	A***	A***	A***	A***	B***	A***	A***					
11	B***	B***	A***	B***	A***	A***	B***	B***	B***	B***				
12	B***	B***	A***	A***	A***	A***	B***	=	B***	B***	A***			
13	B***	B***	B***	B***	A***	A***	B***	B***	B***	B***	B***	B***		
14	B***	B***	B***	B***	A***	A***	B***	B***	B***	B***	B***	B***	B***	

Table B17: Combination D4

A/B	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1														
2	B***													
3	B***	B***												
4	B***	B***	A***											
5	B***	B***	B***	B***										
6	B***	B***	B***	B***	B***									
7	B***	B***	A***	A***	A***	A***								
8	B***	B***	A***	A***	A***	A***	B***							
9	B***	B***	A***	A***	A***	A***	B***	B**						
10	B***	B***	A***	A***	A***	A***	B***	A***	A***					
11	B***	B***	=	B***	A***	A***	B***	B***	B***	B***				
12	B***	B***	A***	=	A***	A***	B***	B***	B***	B***	A***			
13	B***	B***	B***	B***	A***	A***	B***	B***	B***	B***	B***	B***		
14	B***	B***	B***	B***	A***	A***	B***	B***	B***	B***	B***	B***	B***	

Table B18: Combination A5

A/B	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1														
2	=													
3	B***	B***												
4	B***	B***	A***											
5	B***	B***	B***	B***										
6	B***	B***	B***	B***	A*									
7	=	=	A***	A***	A***	A***								
8	=	=	A***	A***	A***	A***	=							
9	B***	B***	A***	B***	A***	A***	B***	B***						
10	B***	B***	A***	B***	A***	A***	B***	B***	A***					
11	B***	B***	B***	B***	A***	A***	B***	B***	B***	B***				
12	B***	B***	A***	B***	A***	A***	B***	B***	B***	B***	A***			
13	B**	B**	A***	A***	A***	A***	=	B*	A***	A***	A***	A***		
14	=	=	A***	A***	A***	A***	=	=	A***	A***	A***	A***	=	

Table B19: Combination B5

A/B	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1														
2	=													
3	B***	B***												
4	B***	B***	A***											
5	B***	B***	B***	B***										
6	B***	B***	B***	B***	A***									
7	B***	B***	A***	B***	A***	A***								
8	B***	B***	A***	B***	A***	A***	B***							
9	B***	B***	B***	B***	A***	A***	B***	B***						
10	B***	B***	=	B***	A***	A***	B***	B***	A***					
11	B***	B***	B***	B***	A***	B***	B***	B***	B***	B***				
12	B***	B***	B***	B***	A***	=	B***	B***	B***	B***	A***			
13	B***	B***	A***	B***	A***	A***	B***	B***	A***	A***	A***	A***		
14	B***	B***	A***	B***	A***	A***	A***	A***	A***	A***	A***	A***	A***	A***

Table B20: Combination C5

A/B	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1														
2	=													
3	B***	B***												
4	B***	B***	A***											
5	B***	B***	B***	B***										
6	B***	B***	B***	B***	A***									
7	B***	B***	A***	A***	A***	A***								
8	B***	B***	A***	A***	A***	A***	B***							
9	B***	B***	A*	B***	A***	A***	B***	B***						
10	B***	B***	A***	B***	A***	A***	B***	B***	A***					
11	B***	B***	B***	B***	A***	A***	B***	B***	B***	B***				
12	B***	B***	B***	B***	A***	A***	B***	B***	B***	B***	A***			
13	B***	B***	A***	B***	A***	A***	B***	B***	A***	A***	A***	A***		
14	B***	B***	A***	B***	A***	A***	B***	B***	A***	A***	A***	A***	B***	

Table B21: Combination D5

A/B	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1														
2	=													
3	B***	B***												
4	B***	B***	A***											
5	B***	B***	B***	B***										
6	B***	B***	B***	B***	A***									
7	B***	B***	A***	A***	A***	A***								
8	B***	B***	A***	A***	A***	A***	B***							
9	B***	B***	B***	B***	A***	A***	B***	B***						
10	B***	B***	A***	B***	A***	A***	B***	B***	A***					
11	B***	B***	B***	B***	A***	B***	B***	B***	B***	B***				
12	B***	B***	B***	B***	A***	A***	B***	B***	B***	B***	A***			
13	B***	B***	A***	B***	A***	A***	B***	B***	A***	A***	A***	A***		
14	B***	B***	A***	B***	A***	A***	B***	B***	A***	A***	A***	A***	B***	

B2 False negative probability: Indicator does not predict an existing jump

Table B22: Combination A1

A\B	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1														
2	B***													
3	B***	B***												
4	B***	B***	B***											
5	B***	B***	A***	A***										
6	B***	B***	A***	A***	B***									
7	B***	A***	A***	A***	A***	A***								
8	B***	A***	A***	A***	A***	A***	B***							
9	B***	B***	B***	B***	B***	B***	B***	B***						
10	B***	B***	B***	B***	B***	B***	B***	B***	B***					
11	B***	B***	B***	B***	B***	B***	B***	B***	A***	A***				
12	B***	B***	B***	B***	B***	B***	B***	B***	B***	A***	B***			
13	B***	B***	A***	A***	B***	B***	B***	B***	A***	A***	A***	A***		
14	B***	B***	A***	A***	B***	B***	B***	B***	A***	A***	A***	A***	A***	

Table B23: Combination B1

A\B	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1														
2	B***													
3	B***	B***												
4	B***	B***	B***											
5	B***	A***	A***	A***										
6	B***	B***	A***	A***	B***									
7	A***	A***	A***	A***	A***	A***								
8	A***	A***	A***	A***	A***	A***	A***							
9	B***	B***	B***	B***	B***	B***	B***	B***						
10	B***	B***	B***	B***	B***	B***	B***	B***	B***					
11	B***	B***	B***	A***	B***	B***	B***	B***	A***	A***				
12	B***	B***	B***	B***	B***	B***	B***	B***	B***	A***	B***			
13	B***	B***	A***	A***	B***	B***	B***	B***	A***	A***	A***	A***		
14	B***	B***	A***	A***	B***	B***	B***	B***	A***	A***	A***	A***	A***	

Table B24: Combination C1

A\B	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1														
2	B***													
3	B***	B***												
4	B***	B***	B***											
5	B***	B*	A***	A***										
6	B***	B***	A***	A***	B***									
7	B***	A***	A***	A***	A***	A***								
8	B***	A***	A***	A***	A***	A***	B***							
9	B***	B***	B***	B***	B***	B***	B***	B***						
10	B***	B***	B***	B***	B***	B***	B***	B***	B***					
11	B***	B***	B***	A***	B***	B***	B***	B***	A***	A***				
12	B***	B***	B***	B***	B***	B***	B***	B***	B***	A***	B***			
13	B***	B***	A***	A***	B***	B***	B***	B***	A***	A***	A***	A***		
14	B***	B***	A***	A***	B***	B***	B***	B***	A***	A***	A***	A***	A***	

Table B25: Combination D1

A\B	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1														
2	B***													
3	B***	B***												
4	B***	B***	B***											
5	B***	A***	A***	A***										
6	B***	B***	A***	A***	B***									
7	A***	A***	A***	A***	A***	A***								
8	A***	A***	A***	A***	A***	A***	B***							
9	B***	B***	B***	B***	B***	B***	B***	B***						
10	B***	B***	B***	B***	B***	B***	B***	B***	B***					
11	B***	B***	B***	A***	B***	B***	B***	B***	A***	A***				
12	B***	B***	B***	B***	B***	B***	B***	B***	B***	A***	B***			
13	B***	B***	A***	A***	B***	B***	B***	B***	A***	A***	A***	A***		
14	B***	B***	A***	A***	B***	B***	B***	B***	A***	A***	A***	A***	A***	

Table B26: Combination A2

A/B	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1														
2	B***													
3	B***	B***												
4	B***	B***	B***											
5	A*	A***	A***	A***										
6	B***	A***	A***	A***	B***									
7	A***	A***	A***	A***	A***	A***								
8	A***	A***	A***	A***	A***	A***	=							
9	B***	B***	B***	B***	B***	B***	B***	B***						
10	B***	B***	B***	B***	B***	B***	B***	B***	B***					
11	B***	B***	B***	A***	B***	B***	B***	B***	A***	A***				
12	B***	B***	B***	B***	B***	B***	B***	B***	B**	A***	B***			
13	B***	B***	B***	A***	B***	B***	B***	B***	A***	A***	A***	A***		
14	B***	B***	B***	A***	B***	B***	B***	B***	A***	A***	A***	A***	=	

Table B27: Combination B2

A/B	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1														
2	B***													
3	B***	B***												
4	B***	B***	B***											
5	A***	A***	A***	A***										
6	B***	A***	A***	A***	B***									
7	A***	A***	A***	A***	A***	A***								
8	A***	A***	A***	A***	A***	A***	=							
9	B***	B***	B***	=	B***	B***	B***	B***						
10	B***	B***	B***	B***	B***	B***	B***	B***	B***					
11	B***	B***	B***	A***	B***	B***	B***	B***	A***	A***				
12	B***	B***	B***	=	B***	B***	B***	B***	=	A***	B***			
13	B***	B***	B***	A***	B***	B***	B***	B***	A***	A***	A***	A***		
14	B***	B***	B***	A***	B***	B***	B***	B***	A***	A***	A***	A***	=	

Table B28: Combination C2

A/B	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1														
2	B***													
3	B***	B***												
4	B***	B***	B***											
5	B***	A***	A***	A***										
6	B***	=	A***	A***	B***									
7	A***	A***	A***	A***	A***	A***								
8	A***	A***	A***	A***	A***	A***	A***							
9	B***	B***	B***	B***	B***	B***	B***	B***						
10	B***	B***	B***	B***	B***	B***	B***	B***	B***					
11	B***	B***	B***	A***	B***	B***	B***	B***	A***	A***				
12	B***	B***	B***	B***	B***	B***	B***	B***	B***	A***	B***			
13	B***	B***	A***	A***	B***	B***	B***	B***	A***	A***	A***	A***		
14	B***	B***	A***	A***	B***	B***	B***	B***	A***	A***	A***	A***	=	

Table B29: Combination D2

A/B	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1														
2	B***													
3	B***	B***												
4	B***	B***	B***											
5	A**	A***	A***	A***										
6	B***	A***	A***	A***	B***									
7	A***	A***	A***	A***	A***	A***								
8	A***	A***	A***	A***	A***	A***	A***							
9	B***	B***	B***	B**	B***	B***	B***	B***						
10	B***	B***	B***	B***	B***	B***	B***	B***	B***					
11	B***	B***	B***	A***	B***	B***	B***	B***	A***	A***				
12	B***	B***	B***	B***	B***	B***	B***	B***	=	A***	B***			
13	B***	B***	=	A***	B***	B***	B***	B***	A***	A***	A***	A***		
14	B***	B***	=	A***	B***	B***	B***	B***	A***	A***	A***	A***	=	

Table B30: Combination A3

A/B	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1														
2	B***													
3	B***	B***												
4	B***	B***	B***											
5	A***	A***	A***	A***										
6	A***	A***	A***	A***	B***									
7	A***	A***	A***	A***	A***	A***								
8	A***	A***	A***	A***	A***	A***	=							
9	B***	B***	B***	A*	B***	B***	B***	B***						
10	B***	B***	B***	B***	B***	B***	B***	B***	B***					
11	B***	B***	B***	A***	B***	B***	B***	B***	A***	A***				
12	B***	B***	B***	A***	B***	B***	B***	B***	A***	A***	B***			
13	B***	B***	=	A***	B***	B***	B***	B***	A***	A***	A***	A***		
14	B***	B***	=	A***	B***	B***	B***	B***	A***	A***	A***	A***	=	

Table B31: Combination B3

A/B	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1														
2	B***													
3	B***	B***												
4	B***	B***	B***											
5	A***	A***	A***	A***										
6	A***	A***	A***	A***	B***									
7	A***	A***	A***	A***	A***	A***								
8	A***	A***	A***	A***	A***	A***	=							
9	B***	B***	B***	A***	B***	B***	B***	B***						
10	B***	B***	B***	B***	B***	B***	B***	B***	B***					
11	B***	B***	B***	A***	B***	B***	B***	B***	A***	A***				
12	B***	B***	B***	A***	B***	B***	B***	B***	A***	A***	B***			
13	B***	B***	=	A***	B***	B***	B***	B***	A***	A***	A***	A***		
14	B***	B***	=	A***	B***	B***	B***	B***	A***	A***	A***	A***	=	

Table B32: Combination C3

A/B	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1														
2	B***													
3	B***	B***												
4	B***	B***	B***											
5	A***	A***	A***	A***										
6	B***	A***	A***	A***	B***									
7	A***	A***	A***	A***	A***	A***								
8	A***	A***	A***	A***	A***	A***	A***							
9	B***	B***	B***	=	B***	B***	B***	B***						
10	B***	B***	B***	B***	B***	B***	B***	B***	B***					
11	B***	B***	B***	A***	B***	B***	B***	B***	A***	A***				
12	B***	B***	B***	A**	B***	B***	B***	B***	=	A***	B***			
13	B***	B***	=	A***	B***	B***	B***	B***	A***	A***	A***	A***		
14	B***	B***	=	A***	B***	B***	B***	B***	A***	A***	A***	A***	=	

Table B33: Combination D3

A/B	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1														
2	B***													
3	B***	B***												
4	B***	B***	B***											
5	A***	A***	A***	A***										
6	A***	A***	A***	A***	B***									
7	A***	A***	A***	A***	A***	A***								
8	A***	A***	A***	A***	A***	A***	=							
9	B***	B***	B***	A***	B***	B***	B***	B***						
10	B***	B***	B***	B***	B***	B***	B***	B***	B***					
11	B***	B***	B***	A***	B***	B***	B***	B***	A***	A***				
12	B***	B***	B***	A***	B***	B***	B***	B***	A***	A***	B***			
13	B***	B***	=	A***	B***	B***	B***	B***	A***	A***	A***	A***		
14	B***	B***	=	A***	B***	B***	B***	B***	A***	A***	A***	A***	=	

Table B34: Combination A4

A/B	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1														
2	B***													
3	B***	B***												
4	B***	B***	B***											
5	B***	A***	A***	A***										
6	B***	A***	A***	A***	B***									
7	A***	A***	A***	A***	A***	A***								
8	A***	A***	A***	A***	A***	A***	B**							
9	B***	B***	B***	B***	B***	B***	B***	B***						
10	B***	B***	B***	B***	B***	B***	B***	B***	B***					
11	B***	B***	B***	A***	B***	B***	B***	B***	A***	A***				
12	B***	B***	B***	B***	B***	B***	B***	B***	B***	A***	B***			
13	B***	B***	B***	A***	B***	B***	B***	B***	A***	A***	A***	A***		
14	B***	B***	B**	A***	B***	B***	B***	B***	A***	A***	A***	A***	=	

Table B35: Combination B4

A/B	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1														
2	B***													
3	B***	B***												
4	B***	B***	B***											
5	A***	A***	A***	A***										
6	B***	A***	A***	A***	B***									
7	A***	A***	A***	A***	A***	A***								
8	A***	A***	A***	A***	A***	A***	A**							
9	B***	B***	B***	A***	B***	B***	B***	B***						
10	B***	B***	B***	B***	B***	B***	B***	B***	B***					
11	B***	B***	B***	A***	B***	B***	B***	B***	A***	A***				
12	B***	B***	B***	=	B***	B***	B***	B***	=	A***	B***			
13	B***	B***	=	A***	B***	B***	B***	B***	A***	A***	A***	A***		
14	B***	B***	=	A***	B***	B***	B***	B***	A***	A***	A***	A***	=	

Table B36: Combination C4

A/B	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1														
2	B***													
3	B***	B***												
4	B***	B***	B***											
5	B***	A***	A***	A***										
6	B***	B**	A***	A***	B***									
7	A***	A***	A***	A***	A***	A***								
8	A***	A***	A***	A***	A***	A***	=							
9	B***	B***	B***	B**	B***	B***	B***	B***						
10	B***	B***	B***	B***	B***	B***	B***	B***	B***					
11	B***	B***	B***	A***	B***	B***	B***	B***	A***	A***				
12	B***	B***	B***	B***	B***	B***	B***	B***	B***	A***	B***			
13	B***	B***	A***	A***	B***	B***	B***	B***	A***	A***	A***	A***		
14	B***	B***	A***	A***	B***	B***	B***	B***	A***	A***	A***	A***	A***	

Table B37: Combination D4

A/B	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1														
2	B***													
3	B***	B***												
4	B***	B***	B***											
5	A***	A***	A***	A***										
6	B***	A***	A***	A***	B***									
7	A***	A***	A***	A***	A***	A***								
8	A***	A***	A***	A***	A***	A***	A***							
9	B***	B***	B***	=	B***	B***	B***	B***						
10	B***	B***	B***	B***	B***	B***	B***	B***	B***					
11	B***	B***	B***	A***	B***	B***	B***	B***	A***	A***				
12	B***	B***	B***	=	B***	B***	B***	B***	B***	A***	B***			
13	B***	B***	=	A***	B***	B***	B***	B***	A***	A***	A***	A***		
14	B***	B***	A**	A***	B***	B***	B***	B***	A***	A***	A***	A***	A**	

Table B38: Combination A5

A\B	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1														
2	B***													
3	B***	B***												
4	B***	B***	B***											
5	A***	A***	A***	A***										
6	A***	A***	A***	A***	B***									
7	A***	A***	A***	A***	A***	A***								
8	A***	A***	A***	A***	A***	A***	B***							
9	B***	B***	B***	A***	B***	B***	B***	B***						
10	B***	B***	B***	A***	B***	B***	B***	B***	B***					
11	B***	B***	B*	A***	B***	B***	B***	B***	A***	A***				
12	B***	B***	B***	A***	B***	B***	B***	B***	=	A***	B***			
13	A***	A***	A***	A***	B***	A***	B***	B***	A***	A***	A***	A***		
14	A***	A***	A***	A***	B***	A***	B***	B***	A***	A***	A***	A***	A***	A***

Table B39: Combination B5

A\B	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1														
2	B***													
3	B***	B***												
4	B***	B***	B***											
5	A***	A***	A***	A***										
6	B***	B***	A***	A***	B***									
7	A***	A***	A***	A***	A***	A***								
8	A***	A***	A***	A***	A***	A***	A***							
9	B***	B***	B***	A***	B***	B***	B***	B***						
10	B***	B***	B***	A***	B***	B***	B***	B***	B***					
11	B***	B***	A***	A***	B***	B*	B***	B***	A***	A***				
12	B***	B***	B***	A***	B***	B***	B***	B***	=	A***	B***			
13	A***	A***	A***	A***	B***	A***	B***	B***	A***	A***	A***	A***		
14	A***	A***	A***	A***	B***	A***	B***	B***	A***	A***	A***	A***	A***	A***

Table B40: Combination C5

A\B	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1														
2	B***													
3	B***	B***												
4	B***	B***	B***											
5	A***	A***	A***	A***										
6	B***	A**	A***	A***	B***									
7	A***	A***	A***	A***	A***	A***								
8	A***	A***	A***	A***	A***	A***	=							
9	B***	B***	B***	A***	B***	B***	B***	B***						
10	B***	B***	B***	A***	B***	B***	B***	B***	B***					
11	B***	B***	A***	A***	B***	B***	B***	B***	A***	A***				
12	B***	B***	B***	A***	B***	B***	B***	B***	B**	A***	B***			
13	A***	A***	A***	A***	B***	A***	B***	B***	A***	A***	A***	A***		
14	A***	A***	A***	A***	B***	A***	B***	B***	A***	A***	A***	A***	A***	

Table B41: Combination D5

A\B	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1														
2	B***													
3	B***	B***												
4	B***	B***	B***											
5	A***	A***	A***	A***										
6	B***	=	A***	A***	B***									
7	A***	A***	A***	A***	A***	A***								
8	A***	A***	A***	A***	A***	A***	A***							
9	B***	B***	B***	A***	B***	B***	B***	B***						
10	B***	B***	B***	A***	B***	B***	B***	B***	B***					
11	B***	B***	A***	A***	B***	B***	B***	B***	A***	A***				
12	B***	B***	B***	A***	B***	B***	B***	B***	=	A***	B***			
13	A***	A***	A***	A***	B***	A***	B***	B***	A***	A***	A***	A***		
14	A***	A***	A***	A***	B***	A***	B***	B***	A***	A***	A***	A***	A***	

Working Paper Series
ISSN 1211-3298
Registration No. (Ministry of Culture): E 19443

Individual researchers, as well as the on-line and printed versions of the CERGE-EI Working Papers (including their dissemination) were supported from the European Structural Fund (within the Operational Programme Prague Adaptability), the budget of the City of Prague, the Czech Republic's state budget and the following institutional grants:

- Center of Advanced Political Economy Research [Centrum pro pokročilá politicko-ekonomická studia], No. LC542, (2005-2011);
- Economic Aspects of EU and EMU Entry [Ekonomické aspekty vstupu do Evropské unie a Evropské měnové unie], No. AVOZ70850503, (2005-2011);
- Economic Impact of European Integration on the Czech Republic [Ekonomické dopady evropské integrace na ČR], No. MSM0021620846, (2005-2011);

Specific research support and/or other grants the researchers/publications benefited from are acknowledged at the beginning of the Paper.

(c) Jan Hanousek, Evžen Kočenda, Jan Novotný, 2011

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical or photocopying, recording, or otherwise without the prior permission of the publisher.

Published by
Charles University in Prague, Center for Economic Research and Graduate Education (CERGE)
and
Economics Institute ASCR, v. v. i. (EI)
CERGE-EI, Politických vězňů 7, 111 21 Prague 1, tel.: +420 224 005 153, Czech Republic.
Printed by CERGE-EI, Prague
Subscription: CERGE-EI homepage: <http://www.cerge-ei.cz>

Phone: + 420 224 005 153
Email: office@cerge-ei.cz
Web: <http://www.cerge-ei.cz>

Editor: Michal Kejak
Editorial board: Jan Kmenta, Randall Filer, Petr Zemčík

The paper is available online at http://www.cerge-ei.cz/publications/working_papers/.

ISBN 978-80-7343-235-5 (Univerzita Karlova. Centrum pro ekonomický výzkum a doktorské studium)
ISBN 978-80-7344-225-5 (Národohospodářský ústav AV ČR, v. v. i.)



CERGE-EI
P.O.BOX 882
Politických vězňů 7
111 21 Praha 1
Czech Republic
<http://www.cerge-ei.cz>