Royal Holloway University of London

Discussion Paper Series

2005 – 02

Department of Economics
Royal Holloway College
University of London
Egham TW20 0EX

Clustering of Trading Activity in the DAX Index Options Market*

Alexander K. Koch[a], Zdravetz Lazarov[b] †

[a] Royal Holloway, University of London

[b] Bonn Graduate School of Economics, University of Bonn

February 27, 2005

Abstract

A common contention is that more liquid financial contracts draw trading volume from contracts for which they are close substitutes. This paper tests this hypothesis by analyzing clustering of trading activity in DAX index options. Contracts with identical maturities cluster around particular classes of strike prices. For example, options with strikes ending on 50 are less traded than options with strikes ending on 00. The degree of substitution between options with neighboring strikes depends on the strike price grid and options' characteristics. Our empirical analysis finds a positive relation between clustering and substitutiability between option contracts, providing support to the initial hypothesis.

JEL classification: C24; G10

Keywords: Clustering; Incidental Truncation; Index Options; Volume

---

*We are grateful for helpful comments from Albert Kyle, Bruce Lehmann, Georg Nöldeke, Jean-Charles Rochet, Philipp Schönbucher, Erik Theissen, and Oved Yosha. Both authors gratefully acknowledge financial support by the Bonn Graduate School of Economics. We thank the Deutsche Börse AG for providing the data. An earlier version of this paper received the "Young Economist Award" at the 17th Annual Congress of the European Economic Association, Venice in 2002.

†Correspondence address: Royal Holloway, University of London, Dept. of Economics, Egham TW20 0EX, United Kingdom. Tel.: +44 1784 44 33 09, E-mail: Alexander.Koch@rhul.ac.uk.

1 Introduction

A common contention is that more liquid financial contracts draw trading volume from contracts for which they are close substitutes. Working (1953)'s famous example of competing futures contracts illustrates this point. In 1950, the Chicago Board of Trade introduced a contract written on Pacific Northwest soft wheat, the risk of which could only be imperfectly hedged with an existing futures contract on hard wheat. Despite offering a better hedge for some market participants, the contract failed because it could not divert enough volume from the more liquid hard wheat futures market. Apparently, the two futures contracts were sufficiently close substitutes so that the liquidity advantage of the existing contract outweighed the hedging advantage that the new contract offered to some market participants.

Despite the intuitive appeal of the above argument we are not aware of any formal empirical analysis of the relation between trading volume and the degree of substitution between different financial assets. This paper fills this gap by analyzing variations in trading volume across option contracts with different strike prices. The advantage of using options instead of futures is that many important contract characteristics such as time to maturity as well as the underlying market can be held fixed.

On a typical day, trading activity in many index options markets follows a saw-toothed pattern (e.g. DAX, CAC40, and DJ EURO STOXX 50). For options with the same maturity date volume is high on one strike price, drops for the next highest strike price, and rebounds for the second-next highest strike price, etc. If option traders indeed trade off liquidity versus the precision of position we should be able to explain variations in the degree of this clustering around some strike prices by changes in factors that affect the degree of substitution between neighboring option contracts.

This paper focuses on DAX index options traded on Eurex since this market is highly liquid and has stable institutional features. The exchange segments the market for options with identical maturity dates into three *strike classes* according to their strike prices. In the following, *200-strike* options refer to the strike class containing all options traded on the 200 index point grid, i.e. with strikes ending on 000, 200, 400, 600, or 800;

100-strike options are traded on the 100 index point grid comprising strikes ending on 100, 300, 500, 700, or 900; *50-strike* options are traded on the 50 index point grid with strikes ending on 50. The additional hybrid *100/200-strike* class contains all options that are either 100- or 200-strike options.

Options with time to maturity exceeding one year all belong to the 200-strike class. The exchange introduces 100-strike options with maturity up to one year and 50-strike options with maturity up to six months. 200-strikes are more frequently traded than 100-strikes, and 100/200-strikes are more frequently traded than 50-strikes. This clustering of trading activity is partly due to differences in open interest. On average 200-strikes are older than 100-strikes and, therefore, have typically accumulated higher open interest than 100-strikes. Similarly, the open interest on 100/200-strikes is typically higher than on 50-strikes.

Our hypothesis is that clustering of trading activity depends on the degree of substitution between options with close strike prices. We maintain that the degree of substitution between options not only depends on open interest but also on the level of the DAX index, time to maturity, the volatility of DAX index returns, options' moneyness, and options' deltas. When options with close strike prices are good substitutes, traders have an interest in concentrating their trades on a subset of contracts to generate liquidity. We hypothesize that the sequential introduction of strike prices serves as a coordination device, making particular strike classes focal. 200-strikes are more attractive than 100-strikes and 100/200-strikes are more attractive than 50-strikes. We define a measure of clustering which compares the trading activity of neighboring options and estimate the relation between clustering and the options' characteristics to test our hypotheses.

Even though the DAX contract is highly liquid compared to other options, some contracts, in particular away from the money, often are not traded on a given day. For this reason any clustering measure will have censoring or truncation of the data. Our empirical analysis deals with this issue by applying a Heckman (1979)-style estimation procedure. Moreover, we ascertain the robustness of our results using an alternative measure based on aggregate trading volume across strike classes, which does not suffer from truncation.

Related Literature

The trade-off between liquidity and other contract characteristics is an important issue in derivatives markets, as illustrated above by Working (1953)'s famous example. Nevertheless, only a few papers explicitly address the relation between liquidity and substitution. Cuny (1993) analyzes financial innovation in a theoretical setup where futures exchanges have a monopoly in their contract and compete with other exchanges in the amount of liquidity their members are able to supply to hedgers. Economides and Siow (1988) consider financial innovation in a model of spatial competition where the division of markets is limited by a liquidity externality.

The paper also relates to the literature on price clustering. Even though price and trade clustering are different phenomena, the line of reasoning for the existence of price clustering applies partly to the context of trade clustering. In many financial markets transaction prices cluster around round price fractions (Grossman, Miller, Cone, and Fischel 1997, Gwilym, Clare, and Thomas 1998). Ball, Torous, and Tschoegl (1985) argue that the amount of information available in the market could determine the market participants' degree of price resolution. To simplify negotiations traders might restrict trading to a discrete price set that is coarser than the price set available (Harris 1991, Harris 1994). In options markets for the same underlying asset and the same maturity date there typically exist several option contracts which differ only in their strike prices. In the spirit of Harris (1991)'s negotiation hypothesis, traders might use discrete strike price sets which are coarser than the strike price set introduced by the exchange. This could facilitate negotiations which are along two dimensions for options, namely prices and strike prices. Moreover, as in Ball, Torous, and Tschoegl (1985)'s price resolution hypothesis, traders possibly choose their desired strike price gradation depending on how accurately they can forecast the value of the underlying asset on the maturity date of the option. If traders use coarser strike price gradations than the exchange's strike price set this results in clustering of trading activity for option series with the same time to maturity.

Finally, Judd and Leisen (2002)'s theoretical model addresses the impact of the strike price grid on the demand for options with different strike prices, reflected by their open interest. However, their analysis does not include liquidity considerations.

The remainder of this paper is organized as follows. Section 2 describes the institutional features of the DAX index options market. Section 3 explains what factors affect clustering. Section 4 provides summary statistics of clustering in trading activity. Section 5 contains the econometric analysis of clustering of trading activity. Section 6 gives robustness checks for our results. Section 7 summarizes the findings.

2 Market Description

DAX index options trade on the Eurex, which is an order driven electronic trading system that ranks orders and quotes by their price and time precedence. Market makers in DAX index options have to respond to at least 50 percent of quote requests during each trading day. These have to be filled within one minute with quotes not exceeding a maximum spread of 15 percent and with a minimum quoted depth of 20 contracts. In exchange they face lower transaction fees (Deutsche Börse 2001b).

Options on the DAX 30 stock index are European style and have a contract value of five Euros per index point. On every trading day the menu of available call and put options includes eight different maturity classes. All contracts expire on the third Friday of their respective maturity months or, if this is a holiday, on the last prior exchange day. For options in the first three maturity classes these expiry dates are in the three succeeding months, respectively. Contracts in maturity classes four, five, and six expire in the succeeding three quarterly expiration months (March, June, September, and December), respectively. Maturity classes seven and eight comprise the succeeding two half-year expiration months (June and December).

DAX index options' strike prices are restricted to lie on price grids with grid sizes of 50, 100, or 200 index points. The Eurex' rules for introducing new option series mandate that strike prices for option series with time to maturity of more than 12 months have a price gradation of 200 index points, options with a remaining term of six to 12 months have a price gradation of 100 index points, and options with less than six months to maturity have a price gradation of 50 index points (Deutsche Börse 2000a, Deutsche Börse 2001a).

The menu of option series ranges from a minimum of five strikes for maturities longer than six months to a minimum of nine strikes for shorter terms. New option series are introduced if the closing level of the DAX exceeded (dropped below) the average of the third- and second-highest (third- and second-lowest) existing exercise prices on the two preceding trading days. An option series is only cancelled if no market participant holds any open position (Deutsche Börse 2000a, Deutsche Börse 2001a).

In the DAX futures market contracts are valued at EUR 25 per index point. The futures' maturities do not always match those of the DAX index options, since contracts are available only for the succeeding three quarterly settlement dates.

3 Factors Affecting Clustering

Our hypothesis is that the degree of substitution between neighboring 50-strike and 100/200-strike options, and between neighboring 100-strike and 200-strike options determines the extent of clustering. When two options with neighboring strike prices are close substitutes, trading activity concentrates on the more attractive option. Additionally, we argue that the sequential nature of introducing strike classes leads to 200-strikes being more attractive than 100-strikes and to 100/200-strikes being more attractive than 50-strikes. We consider the following factors which affect the degree of substitution between options:

Level of the DAX index

When the level of the DAX index goes up the absolute difference in strike prices between neighboring options relative to the index level decreases and, therefore, becomes economically less meaningful.[1] The smaller the relative distance between options is (in terms of their strike prices) the higher the degree of substitution between them. Hence, clustering should increase when the level of the DAX index is high and decrease when the level of the index is low, ceteris paribus.

Time to maturity and volatility of index returns

Many investors in options markets are directional traders who pursue buy-and-hold strategies, i.e. they close their positions near maturity or exercise options. These traders are interested in the index level at or near maturity. The accuracy with which traders can predict the final index level decreases with increasing time to maturity and increasing volatility of the index returns. If investors' predictions become less precise, small differences in strike prices are less important to them.[2] Hence, in choosing between neighboring options which have small strike price distances, such as 50 or 100 index points, trading will concentrate on the more attractive strike classes. This means that clustering should increase with volatility and time to maturity.

Options' deltas

An option's delta gives the sensitivity of the option's price to changes in the index level. Market makers usually combine options with different deltas in order to minimize exposure to risk by keeping the delta of their total position close to zero. Other traders also require a particular delta for their hedging needs. For these types of traders two options with similar deltas are close substitutes. Therefore, one can expect clustering to decrease when the distance between options' deltas increases.

Options' moneyness

In options markets trading tends to concentrate around the at-the-money point and volume decreases for options that are farther away from the at-the-money point. We

[1]Harris (1991) uses a similar argument in the context of minimum price variation rules.
[2]This argument is similar to the price resolution hypothesis in Ball, Torous, and Tschoegl (1985).

expect clustering to increase for options that are farther away from the money since traders strive to coordinate trades to generate liquidity.

Options' open interest

Open interest is a sign of potential future turnover in an option because it affects the number of positions that will be closed out. If two neighboring options do not differ much in the previous factors then traders prefer the option with higher open interest. When options are close substitutes traders are interested in coordinating their trades in order to increase the liquidity in the option series they hold. One way coordination can be achieved is if some strike classes become focal. We argue that the sequential introduction of strike prices makes 200-strikes more attractive than 100-strikes and 100/200-strikes more attractive than 50-strikes.

4 Preliminary Characterization of Clustering in Trading Activity

Our data set comprises all transactions in DAX index futures and options contracts traded on Eurex during the period from 4 January 1999 until 31 July 2002 (908 trading days). We restrict our analysis to the first four maturity classes for which all strike price classes exist, and for which options are sufficiently liquid to allow estimating implied volatilities and option deltas.

On average there are 12.00 (13.55) transactions per 200-strike call (put) option on each trading day. The corresponding figures for 100-strike options are 13.24 (13.63) and 1.73 (1.51) for 50-strike options. Clearly, trading concentrates heavily on 100- and 200-strike options. At first glance, it seems that trading activity clusters more strongly on 100-strike options than on 200-strike options. However, these simple averages may be misleading because they do not account for the sequential nature of option introduction. For example, consider 200-strikes which are introduced earlier than 100-strikes. As time passes and the level of the index changes, some of the older 200-strike options go very deep in- or out-of-the-money while new 100-strike options are only introduced closer to the at-the-money point. Therefore, usually there exist more deep-in- or out-of-the-money 200-strikes than 100-strikes. Typically, far-away-from-the-money options witness low transaction volume or none at all. This distorts any measure of clustering based on the average number of transactions per option in favor of the 100-strike options. The same argument applies when comparing the average transaction volume per option for 100/200-strikes with that for 50-strikes.

To account for the above issue we recompute the average trading activity for the subset of options which have a "neighbor" belonging to the other strike class. For example, a 200- and a 100-strike option are neighbors if their strike prices are 100 index points apart. According to the new statistics, 200-strike call (put) options have the largest average trading volume, 18.09 (19.98), followed by the 100-strike options, 13.26 (13.69). Again the 50-strike options have much lower trading activity: 1.72 (1.50). Table 1 gives a more detailed breakdown of the average trading activity across options belonging to different maturity classes and different moneyness ranges. For both call and put options, trading concentrates around the at-the-money point and decreases with absolute moneyness. Similarly, trading activity decreases when moving to a higher maturity class.

[TABLE 1 about here]

A simple way to see how clustering of trading activity is related to time to maturity and moneyness is to divide the average trading volume of the higher strike class options by that of the lower strike class options. These statistics are reported in Table 2. The figures for both call and put options are similar. For the 200- versus 100-strike options there is a clear pattern that clustering increases with higher maturity classes and with larger absolute moneyness. The case of 100/200- versus 50-strike options shows less regular behavior. Clustering appears to increase with higher maturity classes, but this pattern is less pronounced than in the previous case. No clear-cut pattern for the impact of moneyness on clustering emerges.

[TABLE 2 about here]

To look at the effects of volatility and the index level on clustering we further refine our analysis. Table 3 reports separately the ratios of average trading volumes on trading days with high/low volatility and high/low index level. On days with volatility below the median (low volatility) clustering tends to be lower than on days with volatility above the median (high volatility). This is in line with the hypothesis in Section 3. The effect is more pronounced for 100/200- versus 50-strike options than for 200- versus 100-strike options. In contrast, for 100/200- versus 50-strike options the impact of the index level on clustering appears to be different from that for 200- versus 100-strike options. In the former case, clustering tends decrease with increasing index level, contrary to the hypothesis in Section 3. In the latter case, the impact of an increase in the index level is in line with our expectations.

[TABLE 3 about here]

The simple statistics above provide preliminary evidence for some of the hypotheses in Section 3. In the next part we move to a formal analysis of clustering that simultaneously gauges the effect of all factors and accounts for various econometric issues.

5 Analysis of Clustering in Trading Activity

This section introduces a measure of clustering relating daily trading activity of neighboring options belonging to different strike classes. Then this measure is regressed on variables that capture the impact of the factors described in Section 3.

5.1 Pairwise Measure of Clustering

A natural way to measure clustering is to define a function $f(T^+, T^-)$ which depends on the trading activity in two neighboring options belonging to different strike classes, where T^+ (T^-) is the trading volume in the more (less) attractive option. This function should increase if T^+ increases relative to T^-. A problem is how to deal with the cases when one or both of the options have zero trading volume. As can be seen in Table 4, roughly three quarters of all observations of pairs of 100/200- and 50-strike options, and roughly one half of all observations of pairs of 200- and 100-strike options fall into in this category. One way to deal with this issue is to restrict the clustering function to take values on a finite interval $[a, b]$. If one of the options has non-zero volume while the other option has zero volume then the function is set to one of the end points of the interval, i.e. $f(T^+, 0) = b$, $T^+ > 0$, and $f(0, T^-) = a$, $T^- > 0$. If both options have zero volume the function is set to some fixed value, e.g. $f(0, 0) = (a + b)/2$.

One needs to account for the mass points on a, b, and $(a + b)/2$ when regressing this measure on the factors that explain clustering. Standard models for limited dependent variables will give biased results in our setting because of the heteroscedasticity and autocorrelation in the residuals that are typical in financial data. In principle, it is possible to correct these biases, but this requires maximization of cumbersome likelihood functions. In the paper we pursue a simpler approach by only using observations for which both options have non-zero trading volume. This introduces a sample selection bias that can be corrected in a robust manner using a Heckman (1979)-style estimation procedure.

[TABLE 4 about here]

We consider separately clustering of 100/200- versus 50-strikes and of 200- versus 100-strike options. The measure of clustering between the i-th pair of 200- and 100-strike options, $PC_{i,t}^{200/100}$, is defined as the logarithm of the ratio of the number of transactions on the 200-strike over the number of transactions on the 100-strike on date t. Moreover, we define analogous measures of clustering for the open interest, $PO_{i,t}^{200/100}$, with the convention that these measures are set equal to zero when both the numerator and the denominator are zero. Formally,

$$\mathbf{PC^{200/100}_{i,t}} = \begin{cases} ln\left(\frac{T^{200}_{i,t}}{T^{100}_{i,t}}\right) & \text{if } T^{200}_{i,t} \cdot T^{100}_{i,t} > 0, \\ \text{not defined} & \text{otherwise,} \end{cases} \tag{1}$$

$$\mathbf{PO^{200/100}_{i,t}} = \begin{cases} ln\left(\frac{O^{200}_{i,t}}{O^{100}_{i,t}}\right) & \text{if } O^{200}_{i,t} \cdot O^{100}_{i,t} > 0, \\ 0 & \text{if } O^{200}_{i,t} + O^{100}_{i,t} = 0, \\ \text{not defined} & \text{otherwise,} \end{cases} \tag{2}$$

where $i \in \mathcal{N}^{200-100}_t$, $t = 1, \ldots, T$ and

$\mathcal{N}^{200-100}_t$: set of neighboring 100- and 200-strike options,
$T^{100}_{i,t}$: transaction volume on the 100-strike,
$T^{200}_{i,t}$: transaction volume on the 200-strike,
$O^{100}_{i,t}$: open interest in the 100-strike,
$O^{200}_{i,t}$: open interest in the 200-strike.

Table 4 reports summary statistics for the pairwise measure of clustering. The next section describes our econometric procedure and presents the estimation results for the pairwise measure of clustering.

[TABLE 4 about here]

5.2 Regression Results

Regressors

We consider the pairwise measures of clustering $PC^{100/50}_{i,t}$ and $PC^{200/100}_{i,t}$ separately for call and put options. As regressors we include the inverse of the DAX index level ($\frac{1}{DAX_t}$), time to maturity ($ttm_{i,t}$) as well as its square to account for possible non-linear maturity effects. The daily GARCH(1,1) DAX index volatility (vol_t) serves as a proxy for the impact of volatility on clustering. To gauge the effect of open interest on clustering we include the previous trading day's value of the appropriate measure of clustering for open interest, $PO_{i,t-1}$. The interaction between $ttm_{i,t}$ and $PO_{i,t-1}$ captures whether the importance of open interest diminishes with shorter remaining life time of the options. As explained in Section 3 clustering should decrease with larger difference between the deltas of neighboring options. The natural approach to measure this effect is to use the absolute value of the arithmetic difference between options' deltas as a regressor. However, this measure is sensitive to moneyness. To reduce potential moneyness effects we scale the absolute value of the arithmetic difference by the absolute value of the average deltas of the two options. Option deltas are computed using the Black and Scholes (1973) formula and the implied volatility, which is obtained using the procedure described in Appendix A. For near-to-maturity options implied volatilities become very unstable and computing them from the data makes little sense. Therefore, in all estimations we restrict the sample to options with maturity exceeding seven days.[3]

Clustering should increase with higher average moneyness and this impact could be asymmetric depending on whether the option is in or out of the money. An option pair is in (out of) the money if the average moneyness[4] of the two options is positive (negative). We use as regressors the average absolute moneyness ($avm_{i,t}$) of the two options in the pair interacted with the dummy $D_{i,t}$, that takes on value one if the pair of options is in the money on day t and zero otherwise, as well as its inverse $1 - D_{i,t}$. To account for possible nonlinear effects, we also include $avm_{i,t}^2$ and $avm_{i,t}$ multiplied by time to maturity.

To avoid problems due to serial correlation in the clustering measure and regressors, which could lead to spurious results, we include the lagged value of the clustering measure as an additional regressor.

Given the choice of dependent variable, our sample is restricted to option pairs for which the clustering measure is defined on the current and on the previous trading day. As can be seen in Table 4, a considerable number of option pairs are not usable as observations. Using a simple OLS estimator could potentially lead to biased results. To account for this problem we employ a two-stage estimation procedure based on Heckman (1979). As a first step, we estimate a probit equation for the inclusion of an option pair in the

[3]Using the odd days only yields virtually the same results.

[4]Moneyness is defined as $m = 1 - \frac{strike}{DAX}$ for call options and as $m = \frac{strike}{DAX} - 1$ for put options.

sample. Then we use this estimation to correct for sample selection bias in the linear regression.

Selection Specification

We assume there exist three latent variables, $u^{50}_{i,t}$, $u^{100}_{i,t}$, and $u^{200}_{i,t}$, which take on positive values whenever the option in the corresponding strike class is traded on the current and on the previous trading day, and which take on non-positive values otherwise, i.e. for $k \in \{50, 100, 200\}$, $t = 2, \ldots, T$, and $i \in \mathcal{K}^k_t$,

$$u^k_{i,t} \quad \begin{cases} > 0 & if\ T^k_{i,t} \cdot T^k_{i,t-1} > 0, \\ \leq 0 & \text{otherwise,} \end{cases} \tag{3}$$

where $\mathcal{K}^k_t$ is the set of k-strike options on day t.

To fix ideas, let us focus on the case of 200- versus 100-strike options. We assume that there exists a pair of latent variables $\left(y_{i,t}, y^*_{i,t}\right)$ defined as follows. It takes on the values of the clustering measure on the current and on the previous trading day whenever they are both defined; otherwise the pair is not observed. Formally, for $t = 2, \ldots, T$ and $i \in \mathcal{N}^{200-100}_t$,

$$\left(y_{i,t}, y^*_{i,t}\right) \quad \begin{cases} = \left(PC^{200/100}_{i,t}, PC^{200/100}_{i,t-1}\right) & if\ u^{100}_{q(i),t} \cdot u^{200}_{l(i),t} > 0, \\ \text{not observed} & \text{otherwise,} \end{cases} \tag{4}$$

where $\mathcal{N}^{200-100}_t$ is the set of neighboring 100- and 200-strike options on day t, and $q(i)$ and $l(i)$ are the corresponding indices of the 100- and 200-strike options in the $i - th$ pair, respectively.

The three latent variables depend on three sets of regressors, $z^{100}_{i,t}$, $z^{200}_{i,t}$, and $x_{i,t}$:

$$u^{100}_{q(i),t} = \left(z^{100}_{q(i),t}\right)'\gamma^{100} + \epsilon^{100}_{q(i),t}, \tag{5}$$

$$u^{200}_{l(i),t} = \left(z^{200}_{l(i),t}\right)'\gamma^{200} + \epsilon^{200}_{l(i),t}, \quad i \in \mathcal{N}^{200-100}_t, \quad t = 2, \ldots, T, \tag{6}$$

$$y_{i,t} = y^*_{i,t}\beta^* + x'_{i,t}\beta + \nu_{i,t}. \tag{7}$$

Assumption 1

(i) The residuals $\epsilon^{100}_{q(i),t}$, $\epsilon^{200}_{l(i),t}$, *and* $\nu_{i,t}$ *are trivariate normally distributed.*

(ii) $\epsilon^{100}_{q(i),t}$ *and* $\epsilon^{200}_{l(i),t}$ *are uncorrelated and their variances are normalized to one:* $var(\epsilon^{100}_{q(i),t}) \equiv var(\epsilon^{200}_{l(i),t}) \equiv 1$.

(iii) The conditional covariances between the residuals in Eqs. (5) and (7) and between the residuals in Eqs. (6) and (7) are linear functions of a subset of the regressors $x_{i,t} \cup z^{100}_{q(i),t}$ *and* $x_{i,t} \cup z^{200}_{l(i),t}$, *respectively:*

$$Cov[\epsilon^k_{j,t}, \nu_{i,t} | y^*_{i,t}, x_{i,t}, z^{100}_{q(i),t}, z^{200}_{l(i),t}, (y_{i,t}, y^*_{i,t}) \text{ observed}] = \left(b^k_{i,t}\right)' \xi^k, \tag{8}$$

$$b^k_{i,t} \subset x_{i,t} \cup z^k_{j,t}, \quad k \in \{100, 200\},\ i \in \mathcal{N}^{200-100}_t, \quad j \in \{q(i), l(i)\}, \quad t = 2, \ldots, T.$$

With the above assumption we obtain the following proposition.

Proposition 1

Under Assumption 1,

$$PC^{200/100}_{i,t} = PC^{200/100}_{i,t-1}\beta^* + x'_{i,t}\beta + M^{100}_{i,t}\left(b^{100}_{i,t}\right)'\xi^{100} + M^{200}_{i,t}\left(b^{200}_{i,t}\right)'\xi^{200} + \omega_{i,t}, \tag{9}$$

where

$$\omega_{i,t} = \nu_{i,t} - E\left[\nu_{i,t} | u^{100}_{q(i),t} > 0, u^{200}_{l(i),t} > 0\right], \quad i \in \mathcal{N}^{200-100}_t, \tag{10}$$

$$M^k_{i,t} = \frac{\phi\left(\left(z^k_{j,t}\right)'\gamma^k\right)}{\Phi\left(\left(z^k_{j,t}\right)'\gamma^k\right)}, \quad k \in \{100, 200\}, j \in \{q(i), l(i)\}, \quad t = 2, \ldots, T, \tag{11}$$

where $\phi()$ *and* $\Phi()$ *are the pdf and cdf of the standard normal distribution, respectively.*

The proof is a straightforward extension of that in Heckman (1979) and is omitted. For 100/200- versus 50-strikes similar equations are obtained. Note that in this case Assumption 1 has to be modified to apply separately to the set of pairs of 50-strikes neighboring a 100-strike and the set of pairs of 50-strikes neighboring a 200-strike.[5]

[5] In an alternative approach, which is not reported in the paper, we estimate probit equations for observing volume on individual options on a single trading day. However, among the four correction terms in the linear estimation, the two terms belonging to a single option in the pair are highly collinear. The specification used in the paper avoids this problem.

Probit Estimation

The dependent variable in the probit estimation is an indicator variable $I^k_{i,t,t-1}$ that takes on the value one whenever the ith option in strike class $k \in \{50, 100, 200\}$ witnesses trading activity on days t and $t-1$, and takes on the value zero otherwise.

The liquidity of options decreases with increasing time to maturity and moneyness. Therefore, the set of regressors $z_{i,t}$ comprises time to maturity and its square: $ttm_{i,t-1}$ and $ttm^2_{i,t-1}$. Since out-of-the-money options are more traded than in-the-money options, we include absolute moneyness of the option $|m_{i,t-1}|$ interacted with the dummy variables $D_{i,t-1}$ and $1-D_{i,t-1}$. The dummy $D_{i,t-1}$ takes on value one if the option is in the money on day $t-1$ and zero otherwise. To account for possible time to maturity interaction effects, we also include $|m_{i,t-1}| \cdot ttm_{i,t-1}$ and $|m_{i,t-1}| \cdot ttm^2_{i,t-1}$, both interacted with the dummy variables $D_{i,t-1}$ and $1-D_{i,t-1}$. Since options with higher open interest are more liquid, we add as a regressor the open interest, $O_{i,t-2}$. Moreover, we interact $O_{i,t-2}$ with $ttm_{i,t-1}$. Additionally, we include index volatility, $vol_{i,t-1}$, the effect of which is explained below. To control for possible autocorrelation in the residuals of the probit equation we also include the lagged dependent variable, $I^k_{i,t-2,t-3}$.

[TABLE 5 about here]

Table 5 reports summary statistics for the individual options in the different strike classes. For the probit estimation we include all options for which the dependent variable and its lagged value exist, i.e. all options that are listed on four consecutive trading days. Moreover, to avoid that the dependent variable covers overlapping trading days we restrict the sample to include only even days. After estimating the models including the full set of regressors we test for joint significance of variables using a Wald test and re-estimate the restricted specifications. For reasons of brevity we report results only for the restricted specifications, summarized in Table 6.

[TABLE 6 about here]

The coefficients on $ttm_{i,t-1}$ and $O_{i,t-2}$ are all significant and have the expected signs. The regressor $vol_{i,t-1}$ is significant with positive coefficient with the exception of 50-strike options. We can interpret this result as follows. When the DAX index becomes more volatile, demand for options tends to increase (*demand effect*) and at the same time clustering should increase (*substitution effect*). The two effects unambiguously should increase the probability of volume on 200-strikes, which is confirmed by the estimation results. For 100-strikes the impact of substitution is ambiguous: some transactions shift from 50- to 100-strike options, increasing the probability of observing positive volume on 100-strikes. However, some transactions shift from 100- to 200-strike options as well, decreasing the probability of observing positive volume on 100-strikes. Therefore, there is no prediction for the overall effect of volatility on 100-strikes. In contrast, for 50-strike call options the substitution effect is predicted to decrease the trading activity. The insignificant coefficient suggests that the demand and substitution effects cancel each other. Finally, the estimations confirm that increasing absolute moneyness negatively impacts the probability of observing trading volume.

Second-stage Estimation

To correct for potential selection bias we include the correction terms $M_{i,t}^k$ obtained from the probit estimation. The set of regressors $b_{i,t}^k$ in the covariance terms $\left(b_{i,t}^k\right)' \xi^k$ includes all of the regressors in $x_{i,t}$ and most of those in $z_{\cdot,t}^k$ from the corresponding probit estimations. To avoid collinearity problems we do not include lagged values of the regressors in $x_{i,t}$. For the same reason only the moneyness regressors from the linear specification $x_{i,t}$ are included in $b_{i,t}^k$ and not the corresponding variables from the probit equations.

Results for the second-stage regressions are reported in Tables 7 and 8. The coefficient on $ttm_{i,t}$ is always significant with the expected positive sign. With the exception of 100/200- versus 50-strike put options, the coefficient on $PO_{i,t-1}$ is significant with the correct sign. The coefficient on $\frac{1}{DAX_t}$ is always significant, with the exception of 100/200- versus 50-strike call options. For 100/200- versus 50-strike put options it is positive, which suggests that clustering decreases when the DAX index level increases, contrary to our predictions. In contrast, for 200- versus 100-strike options it has the correct sign. The coefficient on $vol_{i,t}$ is always significant with the expected sign. The results on the impact of moneyness on clustering are ambiguous. For 100/200- versus 50-strike options our estimates indicate that clustering decreases with higher average moneyness, contrary to our predictions. For the 200- versus 100-strike options the effect of moneyness on clustering in out-of-the-money options is positive, while for in-the-money options the effect is reversed. The delta regressor is always significant with a negative sign, in line with our hypothesis.

[TABLE 7 about here]

[TABLE 8 about here]

6 Robustness Checks

Clustering may depend on the past movements of the index. For example, consider the period when only 200-strike options exist. If the index value changes significantly during that period then trading spreads across different 200-strike options, each of them being near the money only for a short period of time. In this case, 200-strike options do not have a large built-up liquidity advantage over 100- and 50-strike options when these are introduced. Thus, all else equal, clustering should be low. In contrast, if the index does not move much then the older 200-strike options accumulate a larger liquidity advantage over 100- and 50-strike options. Thus, all else equal, clustering should be high. In our regressions the lagged value of the pairwise measure and open interest control for this effect to some degree. However, to ascertain that the above effect does not impact the validity of our results, we re-estimate the restricted specifications of the regressions in Tables 7 and 8 using only pairs of options which are simultaneously introduced. In this case, no option has an initial liquidity advantage. We report the results in Table 9. The regression coefficients do not change signs. The coefficients for the volatility and the delta difference become insignificant for the case of 100/200- versus 50-strike options.

Another way in which time to maturity may affect clustering is the following. Consider, for example a newly introduced 50-strike option and its older 100-strike neighbor which already has a built-up liquidity advantage. After the introduction of the 50-strike option, both options compete for attracting volume and the older option's liquidity advantage will tend to decrease over time. Clustering may decrease with decreasing time to maturity because of this or the effect in Section 3. In the regressions for pairs where both options have the same age (see Table 9) this liquidity advantage effect is not present. The coefficients on time to maturity are roughly the same as in the earlier regressions, which indicates that the effect of time to maturity is as described Section 3.

To explore the robustness of the two-stage estimation procedure we use an alternative measure of clustering. It is based on daily aggregated trading volumes and has the advantage that it is not truncated for the first maturity class. Although this measure does not allow us to isolate all the effects in Section 3 it can be used to obtain results free of selection bias for some factors. The measure is defined as follows. Consider 100- versus 200-strike options. First, we record the individual transaction volumes in every pair of neighboring options on each day. Then for each day we separately sum up over all option pairs the number of transactions on 200-strikes and 100-strikes, respectively. The aggregate measure of clustering of 200- versus 100-strikes, $AC_t^{200/100}$, is defined as the logarithm of the ratio of the transaction volume on 200-strikes over that on 100-strikes. To account for the impact of the open interest we define a similar measure $AO_t^{200/100}$. The measures for the case of 100/200- versus 50-strike options are computed analogously. Since we aggregate over different strike prices the effects of moneyness and differences in deltas can not be isolated. Moreover, clustering in the deep in- and out-of-the-money options impacts the value of this measure much less than clustering in the near-the-money options.

Estimation results for the first maturity class are reported in Table 10. The coefficients for time to maturity, open interest, and volatility have the correct signs and have almost identical patterns of significance as in the case of the pairwise measure. For the DAX index level the coefficients exhibit the same behavior for both the pairwise and the aggregate measures. This provides evidence for the robustness of the two-stage estimation procedure.

7 Conclusion

In the DAX index options market contracts with identical maturities cluster around particular classes of strike prices. For example, options with strikes ending on 50 are less traded than options with strikes ending on 00. Our main hypothesis is that this clustering of trading activity depends on the degree of substitution between options with neighboring strike prices. When two options with nearby strike prices are close substitutes, trading concentrates on the option belonging to the more attractive strike class. We maintain that 200-strikes are more attractive than 100-strikes, and that 100/200-strikes are more attractive than 50-strikes.

Our empirical analysis broadly finds a positive relation between clustering and factors that increase the degree of substitution between options, such as differences in open interest, time to maturity, the volatility of DAX index returns, options' moneyness, and differences in options' deltas. This lends support to our narrow hypotheses about the way in which these factors affect clustering and the attractiveness of different strike classes. More broadly, our results provide evidence in favor of the hypothesis that more liquid financial contracts draw trading volume from contracts for which they are close substitutes.

The trade-off between liquidity and other contract characteristics is an important issue in derivatives markets. In futures markets, the success of a newly introduced contract depends to a large extent on the degree of substitution with existing contracts (e.g., Working (1953)). In options markets, an exchange also needs to consider this trade-off when it decides on the grid size of strike prices for option contracts. On the one hand, if the grid is too coarse, overall trading volume might be low because some traders do not find a contract tailored to their needs. On the other hand, if the grid is too fine, overall volume on the exchange might decrease because individual contracts have too little trading volume, since demand is spread across too many contracts. Intuitively, if options with different strike prices are good substitutes the strike price grid can be coarser than if they are bad substitutes. Our paper is just a first step in understanding better empirically the trade-off between precision of position and liquidity. Further research into this issue promises to deliver important implications for market design.

A Estimation of the Implied Volatilities

On every trading day we match DAX option prices with the nearest to maturity DAX futures prices. Only transactions that are at most 5 minutes apart are considered. We obtain the implied spot level of the DAX for the corresponding five minute intervals by inverting a simple futures pricing formula. The fair price of a future is assumed to be the continuously compounded spot price of the underlying.[6] That is,

$$F_{t,v}(T_F) = S_{t,v} e^{r(T_F - t)}, \tag{12}$$

where

T_F : future's maturity,
$S_{t,v}$: (implied) underlying index in the vth five minute interval on day t,
$F_{t,v}$: nearest to maturity futures contract in the vth five minute interval on day t,
r : risk-free rate for future's term $(T_F - t)$.

We compute the implied spot price of the DAX index by inverting equation (12) and using the average futures price over the respective five minute interval. The appropriate risk-free interest rate is obtained by linearly interpolating EUR-Libor rates bracketing the option's maturity.[7] Based on this sample of matched option prices, spot prices, strike prices, and interest rates, we calculate the implied volatilities by inverting the Black and

[6]The DAX index is computed assuming reinvestment of dividends after corporate income tax on distributed gains (Deutsche Börse 2000b). German income tax law treats dividends as if they partially include corporate income tax. Thus, the above futures pricing formula is not exactly the fair price if the marginal investor's personal income tax rate differs from the corporate income tax rate. For a discussion of this issue see Hafner and Wallmeier (2001). In our data this problem appears to be a minor one.

[7]The interest rate convention for Libor rates is linear and, therefore, rates have to be converted to continuous compounding first.

Scholes (1973) formula. Following Hafner and Wallmeier (2001), we approximate the smile on every trading day by fitting via OLS a smooth differentiable spline function whose segments join at the at the money point. The general specification allows for quadratic function segments for the in- and out-of-the-money ranges, respectively. Let $M = strike/DAX$. We impose

$$\begin{aligned} \sigma_{IV} &= \alpha_0 + \alpha_1 M + \alpha_2 M^2 + D\left(\beta_0 + \beta_1 M + \beta_2 M^2\right) + \epsilon, \qquad (13) \\ D &= \begin{cases} 0 \quad if & \text{if strike below the at-the-money point,} \\ 1 \quad if & \text{if strike above the at-the-money point.} \end{cases} \end{aligned}$$

The restriction is imposed that the two segments join at the at-the-money-point ($M = 1$), i.e. $\beta_0 + \beta_1 + \beta_2 = 1$. Moreover, the smile is assumed to be a smooth, differentiable function with its minimum at $M = 1$, i.e.

$$\left.\frac{d\left(\beta_0 + \beta_1 M + \beta_2 M^2\right)}{dM}\right|_{M=1} = \beta_1 + 2\beta_2 = 0. \qquad (14)$$

Hence, the final specification to be fitted to the data is given by

$$\sigma_{IV} = \alpha_0 + \alpha_1 M + \alpha_2 M^2 + \alpha_3 D\left(1 - 2M + M^2\right) + \epsilon. \qquad (15)$$

Table 1:

Average trading volume for pairs of options

	Maturity class				Maturity class			
Moneyness[a]	1	2	3	4	1	2	3	4
Call options								
	50-strike options				100/200-strike options			
$m \leq -0.2$	0.31	0.78	0.46	0.43	4.24	7.42	6.46	14.58
$-0.2 < m \leq -0.15$	0.85	1.13	0.83	0.50	13.61	18.08	12.82	17.89
$-0.15 < m \leq -0.1$	2.15	1.74	0.51	0.39	25.50	30.73	17.10	14.76
$-0.1 < m \leq -0.05$	5.33	1.90	0.53	0.26	74.16	45.10	16.75	10.22
$-0.05 < m \leq 0$	15.60	1.71	0.44	0.28	173.83	45.10	10.33	7.28
$0 < m \leq 0.05$	4.64	0.57	0.20	0.14	67.51	15.22	4.39	3.13
$0.05 < m \leq 0.1$	0.30	0.11	0.05	0.07	9.86	3.25	1.91	1.47
$0.1 < m \leq 0.15$	0.11	0.09	0.07	0.04	4.32	1.60	1.56	0.82
$0.15 < m \leq 0.2$	0.07	0.07	0.09	0.02	2.74	1.08	1.08	0.61
$0.2 < m$	0.02	0.03	0.08	0.02	1.09	0.45	0.49	0.23
	100-strike options				200-strike options			
$m \leq -0.2$	1.16	1.74	1.24	1.48	1.84	3.01	2.88	4.78
$-0.2 < m \leq -0.15$	6.74	7.22	4.77	3.86	11.63	12.02	11.17	11.63
$-0.15 < m \leq -0.1$	17.61	18.86	9.28	5.34	21.96	27.39	17.53	12.98
$-0.1 < m \leq -0.05$	60.46	35.78	11.67	5.65	74.71	50.58	17.52	11.87
$-0.05 < m \leq 0$	155.31	37.37	7.76	4.91	181.07	52.15	12.72	9.62
$0 < m \leq 0.05$	61.30	12.49	3.00	1.79	76.18	18.37	5.57	4.33
$0.05 < m \leq 0.1$	6.55	2.01	0.65	0.54	10.63	3.54	2.03	1.76
$0.1 < m \leq 0.15$	2.04	0.73	0.37	0.27	4.07	1.51	1.26	0.86
$0.15 < m \leq 0.2$	0.90	0.26	0.28	0.18	2.12	0.88	0.69	0.61
$0.2 < m$	0.40	0.08	0.07	0.05	0.83	0.41	0.43	0.27
Put options								
	50-strike options				100/200-strike options			
$m \leq -0.2$	0.30	0.55	0.30	0.09	4.32	8.67	6.66	3.27
$-0.2 < m \leq -0.15$	0.93	0.73	0.44	0.14	14.15	19.96	11.30	5.64
$-0.15 < m \leq -0.1$	2.09	0.90	0.54	0.23	30.45	29.36	14.03	6.06
$-0.1 < m \leq -0.05$	4.31	1.20	0.34	0.27	67.53	39.00	14.18	7.60
$-0.05 < m \leq 0$	11.76	1.31	0.44	0.28	157.55	41.71	11.20	8.52
$0 < m \leq 0.05$	5.69	0.70	0.28	0.25	72.20	18.86	6.16	5.30
$0.05 < m \leq 0.1$	0.51	0.17	0.09	0.06	13.37	5.58	2.67	2.64
$0.1 < m \leq 0.15$	0.22	0.06	0.04	0.04	7.39	3.78	2.09	3.00
$0.15 < m \leq 0.2$	0.16	0.09	0.03	0.01	6.43	3.24	1.77	4.21
$0.2 < m$	0.09	0.06	0.05	0.03	5.47	2.47	2.28	7.24
	100-strike options				200-strike options			
$m \leq -0.2$	1.89	3.18	2.02	0.94	3.81	6.56	5.32	3.77
$-0.2 < m \leq -0.15$	8.73	11.29	5.02	2.49	13.86	20.52	10.59	7.12
$-0.15 < m \leq -0.1$	22.68	21.14	7.40	3.78	33.70	32.11	14.29	8.61
$-0.1 < m \leq -0.05$	56.33	29.97	8.60	4.85	80.03	46.89	16.67	10.93
$-0.05 < m \leq 0$	136.22	33.44	8.16	5.69	173.84	50.37	14.75	11.76
$0 < m \leq 0.05$	65.49	16.11	4.43	3.48	79.00	21.49	7.48	6.94
$0.05 < m \leq 0.1$	9.92	3.95	1.58	1.28	13.32	5.76	2.75	2.65
$0.1 < m \leq 0.15$	4.19	1.80	1.06	0.66	6.23	3.11	1.78	1.83
$0.15 < m \leq 0.2$	2.65	1.13	0.65	0.53	4.70	1.77	1.31	1.46
$0.2 < m$	1.46	0.69	0.51	0.42	2.69	1.18	1.01	1.25

[a] Moneyness is defined as $m = 1 - \frac{strike}{DAX}$ for call options and as $m = \frac{strike}{DAX} - 1$ for put options.

Table 2:

Clustering statistics for pairs of options

	100/200- vs 50-strike options				200- vs 100-strike options			
	Maturity class				Maturity class			
Moneyness[a]	1	2	3	4	1	2	3	4
			Call options					
			Ratio of average trading volumes[b]					
$m \leq -0.2$	13.49	9.54	13.97	34.14	1.59	1.73	2.32	3.23
$-0.2 < m \leq -0.15$	15.93	16.02	15.35	35.68	1.72	1.67	2.34	3.01
$-0.15 < m \leq -0.1$	11.86	17.63	33.46	37.73	1.25	1.45	1.89	2.43
$-0.1 < m \leq -0.05$	13.92	23.73	31.74	39.55	1.24	1.41	1.50	2.10
$-0.05 < m \leq 0$	11.15	26.32	23.25	25.83	1.17	1.40	1.64	1.96
$0 < m \leq 0.05$	14.56	26.55	21.60	22.19	1.24	1.47	1.86	2.42
$0.05 < m \leq 0.1$	33.14	30.76	35.15	20.61	1.62	1.76	3.10	3.26
$0.1 < m \leq 0.15$	39.99	17.88	23.38	20.13	2.00	2.07	3.39	3.17
$0.15 < m \leq 0.2$	38.03	15.50	12.19	28.70	2.36	3.35	2.50	3.33
$0.2 < m$	57.06	15.76	5.92	15.18	2.08	4.95	5.84	5.54
			Put options					
			Ratio of average trading volumes[b]					
$m \leq -0.2$	14.37	15.88	21.93	34.51	2.01	2.06	2.63	4.02
$-0.2 < m \leq -0.15$	15.22	27.32	25.81	40.68	1.59	1.82	2.11	2.85
$-0.15 < m \leq -0.1$	14.59	32.63	26.04	26.64	1.49	1.52	1.93	2.28
$-0.1 < m \leq -0.05$	15.67	32.47	41.59	27.92	1.42	1.56	1.94	2.25
$-0.05 < m \leq 0$	13.40	31.93	25.56	30.41	1.28	1.51	1.81	2.06
$0 < m \leq 0.05$	12.69	26.93	21.87	21.01	1.21	1.33	1.69	2.00
$0.05 < m \leq 0.1$	26.26	33.48	30.18	41.40	1.34	1.46	1.74	2.07
$0.1 < m \leq 0.15$	33.76	65.03	49.44	74.93	1.49	1.73	1.69	2.79
$0.15 < m \leq 0.2$	40.24	36.53	61.65	407.75	1.78	1.56	2.02	2.76
$0.2 < m$	61.52	40.06	49.80	226.57	1.84	1.73	1.98	2.97

[a] Moneyness is defined as $m = 1 - \frac{strike}{DAX}$ for call options and as $m = \frac{strike}{DAX} - 1$ for put options.

[b] The average trading volume of the higher strike class options is divided by that of the lower strike class options.

Table 3:

Clustering statistics for pairs of options

	100/200- vs 50-strike options				200- vs 100-strike options			
	Maturity class				Maturity class			
Moneyness[a]	1	2	3	4	1	2	3	4
Call options								
Ratio of average trading volumes for trading days with volatility above/below the median[b]								
$m \leq -0.2$	13.61/10.11	10.08/5.33	13.92/ -	34.51/28.36	1.59/1.58	1.64/2.76	2.26/2.79	3.05/3.89
$-0.2 < m \leq -0.15$	16.50/11.42	16.98/12.44	17.49/7.34	33.91/44.66	1.74/1.66	1.68/1.64	2.14/2.77	2.91/3.17
$-0.15 < m \leq -0.1$	14.41/7.10	19.66/14.66	33.45/33.48	37.18/40.81	1.23/1.29	1.43/1.49	1.82/1.97	2.46/2.40
$-0.1 < m \leq -0.05$	15.83/11.87	27.53/20.80	33.78/29.95	42.93/40.83	1.30/1.16	1.48/1.35	1.67/1.38	2.06/2.15
$-0.05 < m \leq 0$	13.65/9.55	30.03/24.00	23.52/23.00	27.13/40.18	1.23/1.11	1.49/1.33	1.70/1.59	2.15/1.78
$0 < m \leq 0.05$	20.69/11.28	30.71/23.62	23.91/19.34	27.96/23.34	1.27/1.22	1.61/1.36	1.92/1.79	2.70/2.11
$0.05 < m \leq 0.1$	33.74/32.39	38.17/24.98	46.95/22.30	47.26/13.10	1.68/1.55	1.89/1.61	2.95/3.42	2.91/3.85
$0.1 < m \leq 0.15$	61.25/24.24	27.02/10.54	31.91/10.46	34.83/23.37	1.95/2.11	2.02/2.15	3.52/2.96	3.18/3.15
$0.15 < m \leq 0.2$	32.93/66.06	19.76/8.77	15.48/5.62	41.36/127.00	2.10/3.02	3.46/3.07	2.34/3.17	4.39/1.86
$0.2 < m$	45.6/114.33	24.47/6.50	6.16/5.14	12.36/11.55	1.60/3.76	4.06/6.85	6.03/5.28	6.44/4.02
Ratio of average trading volumes for trading days with DAX index value above/below the median[b]								
$m \leq -0.2$	8.72/13.72	4.29/9.65	12.81/14.07	43.00/34.05	1.37/1.61	1.59/1.74	2.38/2.30	3.58/3.12
$-0.2 < m \leq -0.15$	10.48/17.50	10.95/17.00	20.03/14.05	48.21/34.71	1.65/1.74	1.61/1.68	2.50/2.25	2.69/3.27
$-0.15 < m \leq -0.1$	8.23/14.63	12.58/22.72	38.24/29.77	31.81/40.81	1.47/1.16	1.64/1.35	1.89/1.89	2.33/2.53
$-0.1 < m \leq -0.05$	11.19/17.47	21.26/27.06	37.86/26.27	34.06/44.83	1.35/1.15	1.53/1.29	1.60/1.38	2.53/1.76
$-0.05 < m \leq 0$	10.18/12.95	24.49/29.65	21.78/26.00	19.95/40.18	1.19/1.12	1.46/1.30	1.79/1.44	2.18/1.73
$0 < m \leq 0.05$	14.00/15.50	23.96/32.08	21.25/22.16	21.34/23.34	1.25/1.23	1.42/1.55	1.95/1.73	2.50/2.32
$0.05 < m \leq 0.1$	34.75/30.93	23.88/53.84	28.26/59.39	29.87/13.10	1.65/1.59	1.74/1.78	3.61/2.48	3.78/2.61
$0.1 < m \leq 0.15$	37.16/46.19	18.04/17.67	17.78/43.68	18.37/23.77	2.03/1.96	2.22/1.89	3.72/2.96	4.21/2.11
$0.15 < m \leq 0.2$	30.61/67.33	14.75/17.25	8.94/27.55	19.33/127.00	2.12/2.80	3.20/3.58	3.75/1.52	3.25/3.44
Put options								
Ratio of average trading volumes for trading days with volatility above/below the median[b]								
$m \leq -0.2$	16.53/9.83	20.98/10.13	20.64/32.25	30.15/178.50	2.09/1.77	2.15/1.88	2.90/1.98	3.75/4.59
$-0.2 < m \leq -0.15$	16.20/12.75	30.11/24.30	27.23/23.31	25.96/194.77	1.67/1.40	1.93/1.66	2.29/1.90	3.07/2.59
$-0.15 < m \leq -0.1$	15.26/13.25	33.64/31.51	19.93/42.01	19.69/37.78	1.53/1.40	1.68/1.37	2.03/1.84	2.31/2.25
$-0.1 < m \leq -0.05$	19.62/12.14	40.90/27.31	42.06/41.08	33.82/23.85	1.42/1.42	1.63/1.51	2.01/1.87	2.45/2.09
$-0.05 < m \leq 0$	17.42/11.02	35.52/29.51	30.00/22.22	33.24/27.94	1.30/1.26	1.59/1.44	1.89/1.73	2.30/1.85
$0 < m \leq 0.05$	16.03/10.58	31.32/23.91	21.43/22.40	29.14/15.58	1.27/1.15	1.42/1.26	1.90/1.48	2.11/1.85
$0.05 < m \leq 0.1$	31.49/19.95	36.09/29.73	32.99/25.91	52.57/27.62	1.35/1.33	1.54/1.34	1.90/1.49	1.78/2.94
$0.1 < m \leq 0.15$	33.12/36.37	63.55/70.43	55.26/29.08	107.88/31.00	1.44/1.65	1.77/1.64	1.50/2.44	2.42/4.26
$0.15 < m \leq 0.2$	38.95/52.71	40.17/22.47	84.33/7.20	371.50/ -	1.76/1.86	1.40/2.23	1.70/4.00	2.64/3.28
$0.2 < m$	63.15/31.50	41.53/22.00	51.05/-	261.00/20.00	1.77/2.84	1.56/3.06	1.92/3.70	2.92/3.47
Ratio of average trading volumes for trading days with DAX index value above/below the median[b]								
$m \leq -0.2$	12.85/17.75	20.00/13.94	21.83/22.03	- /30.24	1.72/2.28	1.99/2.11	2.51/2.75	3.15/5.37
$-0.2 < m \leq -0.15$	13.81/17.45	25.69/29.35	43.99/17.57	139.75/25.33	1.44/1.73	1.91/1.76	1.98/2.24	2.41/3.32
$-0.15 < m \leq -0.1$	11.79/19.86	29.56/36.71	41.31/18.27	43.24/17.76	1.64/1.35	1.79/1.31	2.30/1.61	2.56/2.01
$-0.1 < m \leq -0.05$	12.97/21.31	31.51/33.90	40.20/43.87	28.04/27.74	1.50/1.32	1.65/1.45	1.94/1.93	2.28/2.21
$-0.05 < m \leq 0$	12.07/16.16	29.49/36.60	22.90/31.36	22.70/51.17	1.25/1.32	1.48/1.55	1.73/1.95	2.04/2.09
$0 < m \leq 0.05$	11.14/15.25	23.97/31.40	17.83/33.92	14.75/38.30	1.25/1.16	1.35/1.31	1.78/1.57	2.01/1.98
$0.05 < m \leq 0.1$	21.44/31.84	36.86/31.22	27.35/34.91	34.74/45.86	1.51/1.23	1.44/1.48	1.87/1.57	2.59/1.81
$0.1 < m \leq 0.15$	32.81/34.30	74.26/61.26	109.92/32.17	50.13/88.70	1.66/1.40	1.64/1.78	1.40/2.03	2.33/3.08
$0.15 < m \leq 0.2$	33.91/43.15	20.28/41.85	- /45.18	- /363.00	2.04/1.67	1.55/1.57	1.92/2.09	2.44/2.89
$0.2 < m$	80.14/60.64	5.20/42.39	59.50/49.31	11.50/312.60	2.26/1.79	1.96/1.70	1.86/2.00	3.33/2.92

[a] Moneyness is defined as $m = 1 - \frac{strike}{DAX}$ for call options and as $m = \frac{strike}{DAX} - 1$ for put options.

[b] The average trading volume of the higher strike class options is divided by that of the lower strike class options.

Table 4:

Descriptive statistics for the pairwise measure of clustering

Maturity class	1	2	3	4	1	2	3	4
	100/200- vs 50-strike call options				100/200- vs 50-strike put options			
♯ pairs	29,046	25,790	20,273	21,395	29,046	25,790	20,273	21,395
Observations	12,153	6,914	2,506	1,639	13,285	6,684	2,338	1,650
Average $PC_{i,t}$	2.57	2.51	1.63	1.40	2.64	2.54	1.56	1.27
Maximum $PC_{i,t}$	6.43	5.75	4.75	4.97	6.45	5.46	5.28	5.18
Minimum $PC_{i,t}$	-2.61	-2.30	-3.18	-2.20	-2.64	-2.77	-2.94	-3.78
Standard deviation	1.28	1.21	1.23	1.24	1.24	1.16	1.26	1.21
♯ $T_{i,t}^{50} = 0$	16,747	18,709	17,518	19,517	15,736	19,003	17,728	19,513
♯ $T_{i,t}^{100/200} = 0$	5,175	5,776	6,190	7,686	3,404	4,032	4,755	5,411
♯ $T_{i,t}^{50} = T_{i,t}^{100/200} = 0$	5,029	5,609	5,941	7,447	3,279	3,929	4,548	5,179
	200- vs 100-strike call options				200- vs 100-strike put options			
♯ pairs	23,576	21,589	17,791	24,780	23,576	21,589	17,791	24,780
Observations	12,205	10,671	7,364	8,670	15,254	13,618	9,066	9,699
Average $PC_{i,t}$	0.30	0.41	0.51	0.69	0.33	0.47	0.53	0.64
Maximum $PC_{i,t}$	5.00	4.61	4.23	4.40	4.64	4.44	4.61	4.44
Minimum $PC_{i,t}$	-5.21	-3.83	-4.84	-4.20	-4.47	-4.06	-4.04	-3.85
Standard deviation	1.09	1.02	1.13	1.17	1.00	0.99	1.13	1.17
♯ $T_{i,t}^{100} = 0$	10,235	9,946	9,346	14,784	7,236	6,983	7,527	13,511
♯ $T_{i,t}^{200} = 0$	7,944	8,066	7,378	10,200	5,351	5,441	5,406	8,855
♯ $T_{i,t}^{100} = T_{i,t}^{200} = 0$	6,808	7,094	6,297	8,874	4,265	4,453	4,208	7,285

Table 5:

Descriptive statistics for individual options (maturity classes $1-4$)

	Call Options			Put Options		
	50-strike	100-strike	200-strike	50-strike	100-strike	200-strike
Observations	47,997	44,105	72,211	47,997	44,105	72,211
Observations with $\sharp$ trades>0	12,011	21,821	34,034	12,278	26,221	40,321
Average $\sharp$ trades per day	1.73	13.24	12	1.51	13.63	13.55
Maximum $\sharp$ trades per day	422	577	892	302	623	905
Minimum $\sharp$ trades per day	0	0	0	0	0	0
Standard deviation	8.2	40.06	38.49	6.84	36.13	37.9

Table 6:

Probit estimations for observing trading volume on two consecutive trading days (restricted specifications)

Dependent variable: $I_{t,t-1}$, takes on value one if the option has positive trading volume on days t and $t-1$ and zero otherwise.

	Call options			Put options		
Variable[a]	50-strike	100-strike	200-strike	50-strike	100-strike	200-strike
C	0.94^{***}	0.11	-0.16^{***}	0.86^{***}	-0.46^{***}	-0.35^{***}
	(0.09)	(0.07)	(0.06)	(0.10)	(0.06)	(0.06)
$I_{t-2,t-3}$[b]	0.76^{***}	1.10^{***}	1.33^{***}	0.71^{***}	1.08^{***}	1.25^{***}
	(0.04)	(0.03)	(0.02)	(0.04)	(0.03)	(0.02)
ttm_{t-1}	-0.05^{***}	-0.01^{***}	$-4.85E-3^{***}$	-0.05^{***}	-0.01^{***}	-0.01^{***}
	$(2.41E-3)$	$(5.53E-4)$	$(4.52E-4)$	$(2.79E-3)$	$(4.62E-4)$	$(5.41E-4)$
ttm_{t-1}^2	$1.90E-4^{***}$			$2.02E-4^{***}$		
	$(1.42E-5)$			$(1.57E-5)$		
O_{t-2}[c]	$1.50E-4^{***}$	$6.49E-5^{***}$	$2.87E-5^{***}$	$4.64E-5^{**}$	$1.05E-4^{***}$	$5.16E-5^{***}$
	$(1.34E-5)$	$(3.13E-6)$	$(1.27E-6)$	$(1.89E-5)$	$(6.93E-6)$	$(3.07E-6)$
$ttm_{t-1}\, O_{t-2}$				$1.35E-6^{***}$	$-3.45E-7^{***}$	$-1.04E-7^{***}$
				$(3.78E-7)$	$(7.38E-8)$	$(3.16E-8)$
vol_{t-1}		2.13^{***}	1.96^{***}		2.69^{***}	2.12^{***}
		(0.18)	(0.15)		(0.18)	(0.15)
$D_{t-1}\,\lvert m_{t-1}\rvert$[d]	-45.85^{***}	-10.86^{***}	-6.76^{***}	-28.26^{***}	-6.90^{***}	-6.27^{***}
	(4.29)	(0.81)	(0.46)	(3.74)	(0.30)	(0.34)
$(1-D_{t-1})\,\lvert m_{t-1}\rvert$	-10.70^{***}	-9.81^{***}	-7.99^{***}	-8.16^{***}	-6.00^{***}	-5.57^{***}
	(0.77)	(0.49)	(0.35)	(0.72)	(0.32)	(0.22)
$D_{t-1}\,\lvert m_{t-1}\rvert\, ttm_{t-1}$	0.76^{***}	-0.09^{***}	-0.05^{***}	0.36^{***}		0.02^{***}
	(0.12)	(0.03)	(0.01)	(0.11)		$(3.21E-3)$
$(1-D_{t-1})\,\lvert m_{t-1}\rvert\, ttm_{t-1}$	0.23^{***}	0.07^{***}	0.06^{***}	0.17^{***}	0.06^{***}	0.06^{***}
	(0.03)	(0.01)	(0.01)	(0.03)	(0.01)	$(4.39E-3)$
$D_{t-1}\,\lvert m_{t-1}\rvert\, ttm_{t-1}^2$	$-3.53E-3^{***}$	$5.05E-4^{***}$	$3.33E-4^{***}$	$-1.74E-3^{***}$		
	$(7.30E-4)$	$(1.78E-4)$	$(6.84E-5)$	$(6.67E-4)$		
$(1-D_{t-1})\,\lvert m_{t-1}\rvert\, ttm_{t-1}^2$	$-1.04E-3^{***}$	$-1.98E-4^{***}$	$-1.97E-4^{***}$	$-1.04E-3^{***}$	$-2.76E-4^{***}$	$-2.05E-4^{***}$
	$(2.01E-4)$	$(4.64E-5)$	$(2.90E-5)$	$(1.99E-4)$	$(4.61E-5)$	$(2.21E-5)$
Wald test[e]	1.25	2.25	2.62	1.71	1.32	4.11
	(0.53)	(0.32)	(0.27)	(0.19)	(0.72)	(0.13)
N	20,972	19,618	32,602	20,972	19,618	32,602
$I_{t,t-1}=1$	18,017	11,675	20,142	18,161	9,901	17,409

[a] Huber-White standard errors in parentheses; ***, **, and * indicate significance at 1 percent, 5 percent, and 10 percent level, respectively.

[b] Lagged dependent variable; takes on value one if the option has positive trading volume on days $t-2$ and $t-3$ and zero otherwise.

[c] Open interest of the corresponding option.

[d] D_{t-1}: Dummy which takes on value one if the option is in the money on day $t-1$ and zero otherwise.
$\lvert m_{t-1}\rvert$: (absolute) moneyness on day $t-1$, defined as $\lvert strike/DAX_{t-1}-1\rvert$

[e] p-value in parenthesis. The Wald test statistic is distributed $\chi^2(q)$ under the null hypothesis that the model with q restrictions is true.

Table 7:

Regressions for 100/200-strike options versus 50-strike options (restricted specifications)

Dependent Variable: PC_t, the pairwise measure of clustering.

Variable[a]	Call options			Put options		
		$M^{50\,b}$	$M^{100/200\,b}$		$M^{50\,b}$	$M^{100/200\,b}$
C	1.43***	0.50***	−3.23***	1.15***	0.38***	−0.63***
	(0.10)	(0.10)	(0.45)	(0.13)	(0.11)	(0.22)
PC_{t-1}	0.31***			0.34***		−0.13***
	(0.01)			(0.01)		(0.04)
ttm_t	0.03***	−0.03***	0.03***	0.03***	−0.02***	0.01***
	(4.12E-3)	(2.59E-3)	(0.01)	(4.50E-3)	(2.34E-3)	(1.60E-3)
ttm_t^2	$-2.31E-4^{***}$	$1.56E-4^{***}$	$-1.45E-4^{**}$	$-1.89E-4^{***}$	$1.13E-4^{***}$	
	(5.49E-5)	(2.47E-5)	(5.35E-5)	(4.16E-5)	(2.07E-5)	
PO_{t-1}	0.08***					
	(0.01)					
$ttm_t\, PO_{t-1}$				$8.46E-4^{***}$		
				(2.80E-4)		
$\frac{1}{DAX_t}$			5,624.03***	2,206.08***		
			(1,737.05)	(415.90)		
vol_t	1.74***	−0.76***		1.55***	−0.67**	−1.96***
	(0.34)	(0.25)		(0.38)	(0.30)	(0.75)
$D_t\, avm_t$[c]	−10.88***			−16.57***		
	(2.33)			(1.70)		
$(1-D_t)\ avm_t$	−5.35***	3.26***			2.44***	
	(0.79)	(0.73)			(0.43)	
$D_t\, avm_t^2$	152.11***	−17.61**		117.53***	−15.87***	33.41***
	(52.77)	(8.95)		(12.05)	(2.10)	(5.48)
$(1-D_t)\, avm_t^2$		−4.46***		−20.63***		
		(1.55)		(3.28)		
$D_t\, avm_t\, ttm_t$			−0.03*	0.12***		−0.12***
			(0.02)	(0.03)		(0.03)
$(1-D_t)\, avm_t\, ttm_t$						
$2\,\frac{\lvert\delta_t^{100/200}-\delta_t^{50}\rvert}{\lvert\delta_t^{100/200}+\delta_t^{50}\rvert}$[d]	−2.67***	−3.18***	8.06***	−1.46***	−2.29***	
	(0.78)	(0.94)	(1.64)	(0.54)	(0.64)	
$I_{t-2,t-3}$[e]	|	−0.13**	−0.58***	|	−0.27***	−0.34***
	|	(0.05)	(0.15)	|	(0.04)	(0.09)
O_{t-2}^{50}[f]	|		|	|	$-2.70E-5^{**}$	|
	|		|	|	(1.20E-5)	|
$O_{t-2}^{50}\, ttm_{t-1}$	|		|	|		|
$O_{t-2}^{100/200}$[f]	|	|	$2.96E-5^{**}$	|	|	$9.43E-5^{***}$
	|	|	(1.23E-5)	|	|	(2.14E-5)
$O_{t-2}^{100/200}\ ttm_{t-1}$	|	|		|	|	$-8.09E-7^{***}$
	|	|		|	|	(2.20E-7)
R^2		0.35			0.36	
Wald test[g]		32.70			30.85	
		(0.11)			(0.08)	
N		11,053			10,417	

[a] Variables not included in the full specification are marked by |. Newey-West standard errors in parentheses; ***, **, and * indicate significance at 1 percent, 5 percent level, and 10 percent level, respectively.

[b] This column reports the coefficients for the regressors obtained by interacting the variables with the inverse Mills ratio M^k computed from the corresponding selection equation $k = 50, 100/200$.

[c] D_t : Dummy which takes on value one if the pair of options is in the money on day t. avm_t: average (absolute) moneyness of the pair on day t, defined as $|(strike^{50} + strike^{100/200})/(2\,DAX_t) - 1|$.

[d] δ_t^k: Delta of the option in strike class $k = 50, 100/200$ on day t.

[e] Dummy which takes on value one if the option has positive trading volume on days $t-2$ and $t-3$ and zero otherwise.

[f] Open interest of the option in the corresponding strike class on day $t-2$. The variables O_{t-2}^k and $O_{t-2}^k\, ttm_{t-1}$ are interacted only with the inverse Mills ratio M^k computed from the corresponding selection equation, $k = 50, 100/200$.

[g] p-value in parenthesis. The Wald test statistic is distributed $\chi^2(q)$ under the null hypothesis that the model with q restrictions is true.

Table 8:

Regressions for 200-strike options versus 100-strike options (restricted specifications)

Dependent Variable: PC_t, the pairwise measure of clustering.

Variable[a]	Call options			Put options		
		$M^{100\,b}$	$M^{200\,b}$		$M^{100\,b}$	$M^{200\,b}$
C	0.30*** (0.05)	0.84*** (0.09)	−0.83*** (0.14)	0.15*** (0.05)	0.93*** (0.09)	−0.99*** (0.10)
PC_{t-1}	0.45*** (0.01)	−0.17*** (0.02)	−0.18*** (0.03)	0.34*** (0.01)	−0.02*** (0.01)	−0.17*** (0.02)
ttm_t	$1.38E-3$*** (4.35E-4)	$-2.97E-3$*** (8.93E-4)	$2.12E-3$* (1.20E-3)	$2.99E-3$*** (4.07E-4)	−0.01*** (9.94E-4)	$4.46E-3$*** (9.18E-4)
ttm_t^2					$1.76E-5$*** (3.57E-6)	
PO_{t-1}	0.13*** (0.01)			0.24*** (0.02)	−0.09*** (0.02)	
$ttm_t\, PO_{t-1}$	$-1.31E-3$*** (2.45E-4)	$4.56E-4$** (1.78E-4)		$-1.43E-3$*** (2.01E-4)	$7.18E-4$*** (1.21E-4)	
$\frac{1}{DAX_t}$	−1,787.84*** (249.71)			−1,249.25*** (206.20)		
vol_t	0.68*** (0.10)			0.71*** (0.09)	−0.16** (0.07)	
$D_t\, avm_t$[c]	−1.45*** (0.47)		−3.56*** (0.87)	−0.49** (0.21)		−2.80*** (0.88)
$(1-D_t)\, avm_t$	1.67*** (0.28)	−2.70*** (0.60)	3.14*** (0.71)	0.46*** (0.14)	−2.05*** (0.37)	5.68*** (1.00)
$D_t\, avm_t^2$		6.89*** (1.95)				8.55*** (2.52)
$(1-D_t)\, avm_t^2$					1.90*** (0.46)	−11.33*** (2.66)
$D_t\, avm_t\, ttm_t$					$-2.25E-3$* (1.17E-3)	0.01** (0.01)
$(1-D_t)\, avm_t\, ttm_t$	−4.52*** (1.01)	0.02*** (0.01)	−0.02*** (0.01)		0.01*** (2.00E-3)	−0.02*** (0.01)
$2\,\frac{\lvert\delta_t^{200}-\delta_t^{100}\rvert}{\lvert\delta_t^{100}+\delta_t^{200}\rvert}$[d]	−0.22*** (0.07)	0.92*** (0.18)	−1.48*** (0.29)	−0.20*** (0.08)	−0.31** (0.14)	
$I_{t-2,t-3}$[e]	|	0.70*** (0.06)	−1.22*** (0.09)	|	−0.02** (0.01)	−0.25*** (0.05)
O_{t-2}^{100}[f]	|	$3.97E-5$*** (8.15E-6)	|	|		|
$O_{t-2}^{100}\, ttm_{t-1}$	|	$-3.18E-7$*** (9.25E-8)	|	|		|
O_{t-2}^{200}[f]	|	|		|	|	
$O_{t-2}^{200}\, ttm_{t-1}$	|	|		|	|	
R^2		0.27			0.19	
Wald test[g]		30.27 (0.09)			28.55 (0.05)	
N		27,922			34,473	

[a] Variables not included in the full specification are marked by |. Newey-West standard errors in parentheses; ***, **, and * indicate significance at 1 percent, 5 percent level, and 10 percent level, respectively.

[b] This column reports the coefficients for the regressors obtained by interacting the variables with the inverse Mills ratio M^k computed from the corresponding selection equation $k = 100, 200$.

[c] D_t : Dummy which takes on value one if the pair of options is in the money on day t. avm_t: average (absolute) moneyness of the pair on day t, defined as $|(strike^{100} + strike^{200})/(2\,DAX_t) - 1|$.

[d] δ_t^k: Delta of the option in strike class $k = 100, 200$ on day t.

[e] Dummy which takes on value one if the option has positive trading volume on days $t-2$ and $t-3$ and zero otherwise.

[f] Open interest of the option in the corresponding strike class on day $t-2$. The variables O_{t-2}^k and $O_{t-2}^k\, ttm_{t-1}$ are interacted only with the inverse Mills ratio M^k computed from the corresponding selection equation, $k = 100, 200$.

[g] p-value in parenthesis. The Wald test statistic is distributed $\chi^2(q)$ under the null hypothesis that the model with q restrictions is true.

Table 9:

Regressions for options of the same age (re-estimated restricted specifications)

Dependent Variable: PC_t, the pairwise measure of clustering.

	100/200- vs 50-strike options		200- vs 100-strike options	
Variable[a]	Call options	Put options	Call options	Put options
C	1.55^{***}	1.08^{***}	0.37^{***}	0.21^{***}
	(0.16)	(0.22)	(0.06)	-0.06
PC_{t-1}	0.33^{***}	0.31^{***}	0.44^{***}	0.33^{***}
	(0.01)	(0.02)	(0.01)	-0.01
ttm_t	0.03^{***}	0.04^{***}	$1.17E-3^{**}$	$2.73E-3^{***}$
	(0.01)	(0.01)	(5.06E-4)	-4.94E-04
ttm_t^2	$-2.73E-4^{***}$	$-2.47E-4^{***}$		
	(7.47E-5)	(4.97E-5)		
PO_{t-1}	0.08^{***}		0.10^{***}	0.24^{***}
	(0.02)		(0.02)	-0.02
$ttm_t\,PO_{t-1}$		$4.84E-4$	$-9.82E-4^{***}$	$-1.44E-3^{***}$
		(3.76E-4)	(2.98E-4)	-2.70E-04
$\frac{1}{DAX_t}$		$3,526.80^{***}$	$-1,877.94^{***}$	$-1,087.98^{***}$
		(675.22)	(312.09)	-260.24
vol_t	0.66	0.30	0.64^{***}	0.52^{***}
	(0.62)	(0.75)	(0.13)	-0.12
$D_t\,avm_t$[b]	-11.72^{***}	-13.99^{***}	-2.77^{***}	-0.39
	(3.11)	(2.80)	(0.62)	-0.25
$(1-D_t)\,avm_t$	-6.65^{***}		1.73^{***}	0.15
	(1.05)		(0.33)	-0.18
$D_t\,avm_t^2$	117.20^{**}	89.13^{***}		
	(57.55)	(25.77)		
$(1-D_t)\,avm_t^2$		-16.86^{***}		
		(4.65)		
$D_t\,avm_t\,ttm_t$		0.14^{***}		
		(0.04)		
$(1-D_t)\,avm_t\,ttm_t$			-4.70^{***}	
			(1.09)	
$2\,\frac{\lvert\delta_t^k-\delta_t^l\rvert}{\lvert\delta_t^k+\delta_t^l\rvert}$[c]	-0.87	-0.75	-0.29^{***}	-0.27^{***}
	(1.20)	(0.77)	(0.09)	(0.09)
R^2	0.35	0.35	0.24	0.16
N	5,292	4,779	19,802	23085

[a] Coefficients for the regressors interacted with the inverse Mills ratios computed from the corresponding selection equations are not reported. Newey-West standard errors in parentheses; ***, **, and * indicate significance at 1 percent, 5 percent level, and 10 percent level, respectively.

[b] D_t : Dummy which takes on value one if the pair of options is in the money on day t. avm_t: average (absolute) moneyness of the pair on day t, defined as $|(strike^k + strike^l)/(2\,DAX_t) - 1|$, $k, l \in \{50, 100/200, 100, 200\}$.

[c] δ_t^k (δ_t^l): Delta of the option in the higher (lower) strike class $k \neq l$, $k, l \in \{50, 100/200, 100, 200\}$ on day t.

Table 10:

Regressions for aggregate measure of clustering (maturity class 1, restricted specifications)

Dependent Variable: AC_t, the aggregate measure of clustering.

	100/200- vs 50-strike options		200- vs 100-strike options	
Variable[a]	Call options	Put options	Call options	Put options
C	0.15	-0.71^{***}	0.16^{***}	0.12^{**}
	(0.13)	(0.23)	(0.06)	(0.05)
AC_{t-1}	0.37^{***}	0.46^{***}	0.38^{***}	0.38^{***}
	(0.04)	(0.04)	(0.06)	(0.05)
ttm_t	0.08^{***}	0.11^{***}		$2.70E-3^{**}$
	(0.01)	(0.01)		$(1.17E-3)$
ttm_t^2	$-1.81E-3^{***}$	$-1.37E-3^{***}$		
	$(2.44E-4)$	$(3.04E-4)$		
AO_{t-1}	0.17^{***}	0.36^{***}	0.12^{***}	
	(0.05)	(0.09)	(0.03)	
$ttm_t\, AO_{t-1}$		-0.02^{***}		
		$(4.73E-3)$		
$\frac{1}{DAX_t}$		$2,247.31^{***}$	$-1,273.76^{***}$	-570.66^{**}
		(773.30)	(389.46)	(281.22)
vol_t	2.44^{***}	2.27^{***}	1.08^{***}	0.52^{***}
	(0.54)	(0.67)	(0.35)	(0.18)
R^2	0.57	0.60	0.22	0.16
Wald Test[b]	1.10	-	3.55	6.82
	(0.58)	-	(0.32)	(0.08)
N	905	903	905	905

[a] Newey-West standard errors in parentheses; ***, ** and * indicate significance at 1 percent, 5 percent, and 10 percent level, respectively.

[b] p-value in parenthesis. The Wald test statistic is distributed $\chi^2(q)$ under the null hypothesis that the model with q restrictions is true.

References

Ball, Clifford A., Walter N. Torous, and Adrian E. Tschoegl, 1985, The degree of price resolution: The case of the gold market, *Journal of Futures Markets* 5, 29–43.

Black, Fisher, and Myron Scholes, 1973, The pricing of options and corporate liabilities, *Journal of Political Economy* 81, 637–59.

Cuny, Charles J., 1993, The role of liquidity in futures markets innovations, *Review of Financial Studies* 6, 57–78.

Deutsche Börse, 2000a, Eurex products, Frankfurt am Main.

——— , 2000b, German stock index dax, Frankfurt am Main.

——— , 2001a, Conditions for trading at eurex deutschland and eurex zürich, Frankfurt am Main.

——— , 2001b, Eurex - market making, Frankfurt am Main.

Economides, Nicholas, and Aloysius Siow, 1988, The division of markets is limited by the extent of liquidity (spatial competition with externalities), *American Economic Review* 78, 108–121.

Grossman, Sanford J., Merton H. Miller, Kenneth R. Cone, and Daniel R. Fischel, 1997, Clustering and competition in asset markets, *Journal of Law and Economics* 40, 23–60.

Gwilym, Owain Ap, Andrew Clare, and Stephen Thomas, 1998, Extreme price clustering in the London equity index futures and options markets, *Journal of Banking and Finance* 22, 1193–1206.

Hafner, Reinhold, and Martin Wallmeier, 2001, The dynamics of DAX implied volatilities, *International Quarterly Journal of Finance* 1, 1–27.

Harris, Lawrence, 1991, Stock price clustering and discreteness, *Review of Financial Studies* 4, 389–415.

Harris, Lawrence E., 1994, Minimum price variations, discrete bid-ask spreads, and quotation sizes, *Review of Financial Studies* 7, 149–78.

Heckman, James J., 1979, Sample selection bias as a specification error, *Econometrica* 47, 153–61.

Judd, Kenneth L., and Dietmar P.J. Leisen, 2002, Equilibrium open interest, mimeo, Stanford University.

Working, Holbrook, 1953, Futures trading and hedging, *American Economic Review* 43, 314–343.