Paper: Clustering of Trading Activity in the DAX Index Options Market

Authors: Alexander K. Koch (presenting and attending), Zdravetz Lazarov

Affiliation: Bonn Graduate School of Economics, University of Bonn

Clustering of Trading Activity in the DAX Index Options Market

April 29, 2002

Abstract

Trades in DAX index options 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. Clustering is higher when options with close strike prices are good substitutes. The degree of substitution between options with neighboring strikes depends on the strike price grid and options' characteristics. Using regression analysis we analyze the relation between clustering, grid size, and the options' characteristics. To our knowledge this paper is the first to explore how the grid size of strike prices affects options' trading volume.

JEL classification: C24; G10

Keywords: Clustering; Incidental Truncation; Index Options; Volume

1 Introduction

On a typical day, trading activity in DAX index options on the Eurex follows a sawtooth pattern. For options with the same maturity date turnover is high on one strike price, drops for the next highest strike price, and rebounds for the second-next highest strike price, etc.[1] Why does trading activity cluster around some strike prices? What factors determine the cross-sectional distribution of trading volume for option series with identical maturity dates? Answers to these questions have important implications for market design. When an exchange decides on the grid size of strike prices for option contracts it has to consider the effect this has on the overall trading volume. On the one hand, if the grid is too coarse, 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 because demand is divided among 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.

The phenomenon of clustering has been extensively analyzed in a different context. In many financial markets transaction prices cluster around round price fractions (Grossman, Miller, Cone, and Fischel 1997, Gwilym, Clare, and Thomas 1998). Ball, Torus, 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 option markets the exchange has to choose not only the minimum tick size but also the strike price grid and the maturity structure of option contracts. 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,

[1]The same phenomenon can be found in other options on European stock indices, such as the French CAC40, the Swiss SMI, and the DJ EURO STOXX 50. It is interesting to note that clustering in strike prices is generally not observed in U.S. index options. We are grateful to Bruce Lehmann for pointing this out to us.

traders may 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, Torus, 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.

This paper examines the relation between trading activity in different contracts in the DAX index options market. The Eurex' institutional features segment the markets for options with identical maturity dates according to their strike prices into three *strike classes*. In the following, *200-strike* options refer to the strike class containing all options traded on the 200 index point grid, i.e. with strike prices 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 starts introducing 100-strike options one year before maturity and 50-strike options six months before maturity. 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 the 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.

If options with close strike prices are good substitutes traders would like to concentrate their trades on one contract 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 introduce two measures of clustering of trading activity and estimate the relation between clustering and the options' characteristics to test our hypotheses. A regression analysis supports our predictions about the impact of the options' characteristics on clustering and about the relative attractiveness of the different strike classes.

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 contains the empirical analysis of clustering of trading activity. Section 5 summarizes results and gives conclusions.

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 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 expiry 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 remaining time to maturity of more than 12 months have a price gradation of 200 index points, those with a remaining term of six to 12 months have a price gradation of 100 index points, and those 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 €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, i.e. the maximum term is nine months.

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 introduction of strike classes leads to 200-strikes being more attractive than 100-strikes and 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.[2] The smaller the relative distance is between options (in terms of their strike prices) the higher is 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 option 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.[3] 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 increase with the absolute value of differences in options' deltas.

[2] Harris (1991) uses a similar argument in the context of minimum price variation rules.

[3] This argument is similar to the price resolution hypothesis in Ball, Torus, and Tschoegl (1985).

Options' moneyness

In options markets trading tends to concentrate around the at-the-money point and volume decreases for options that are further away from the at-the-money point. We expect clustering to increase for options that are farther away from the money since traders strive to coordinate trades to generate volume.

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. The effect of open interest on clustering should be strongest for longer term options because there are fewer opportunities to close out positions when time to maturity decreases.

When options are close substitutes traders are interested in coordinating their trades in order to increase the volume on 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 Analysis of Clustering in Trading Activity

We define two measures of clustering in trading activity. The first measure captures clustering in an aggregate sense by comparing total transaction volumes across different strike classes. The second measure gauges clustering between pairs of neighboring options belonging to different strike classes. We regress the two measures on the factors described in Section 3 to explain trade clustering between 50- and 100/200-strike options as well as between 100- and 200-strike options.

Our data set comprises all transactions in DAX index futures and options contracts traded on the Eurex during the period from January 4, 1999 until September 29, 2000 (445 trading days). We restrict our analysis to the first four maturity classes for which all strike price classes coexist.[4]

4.1 Aggregate Clustering in Transaction Volume

4.1.1 Aggregate Measures of Clustering

Our measure of aggregate clustering is roughly the ratio of the average numbers of transactions per option for two different strike classes. Using simple averages would give a distorted measure of clustering. To see this, consider 200-strikes which are introduced earlier than 100-strikes. As time passes and the level of the index changes, some of the older option series go very deep in- or out-of-the-money. Therefore, at a given point of time one can expect to find more deep-in- or deep-out-of-the-money 200-strikes than 100-strikes. Typically, far-away-from-the-money options witness less trading or none at all. This would bias average transaction volume in favor of the 100-strikes. To overcome this problem, our aggregate measure of clustering compares only the transaction volumes of options which are "neighbors". A 100-strike (50-strike) and a 200-strike (100/200-strike) option are neighbors if their strike prices differ by 100 (50) index points.

The aggregate measure of clustering in transaction volume for 100- versus 200-strike options is computed for every trading day according to the following procedure. First, we record the individual transaction volume for the two options in every pair of neighboring 200-strike and 100-strike options on that day. Then we separately sum up over all option pairs the number of transactions on 200-strikes and on 100-strikes, respectively. The measure of clustering for 100- versus 200-strikes, $AC_t^{200/100}$, is defined as the logarithm of the ratio of the total transaction volume on 200-strike options over that on 100-strike options. To account for the impact of open interest on clustering we define a the measure $AO_t^{200/100}$ which is computed following the same steps as above with one slight modification. If both 100- and 200-strikes have zero open interest, this is treated as if the open interests where equal and positive, i.e. we then define $AO_t^{200/100}$ to be

[4] According to the Eurex' rules for introducing new options only the first four maturity classes should contain 50-strike options (Deutsche Börse 2000a, Deutsche Börse 2001a). In our data set, 50-strike options are available on 200 (20) days for maturity class five (six). In these maturity classes the remaining life time of the options exceeds six months. Our focus on the first four maturity classes avoids having to account for discretionary exceptions to exchange rules.

zero. Corresponding measures for 50- versus 100/200-strikes, $AC_t^{100/50}$ and $AO_t^{100/50}$, are defined in a similar fashion. Formally,

$$\mathbf{AC_t^{200/100}} = ln\left(\frac{\sum_{i\in\mathcal{K}_t^{100/200}} T_{it}^{200}}{\sum_{i\in\mathcal{K}_t^{100/200}} T_{it}^{100}}\right), \tag{1}$$

$$\mathbf{AO_t^{200/100}} = \begin{cases} ln\left(\frac{\sum_{i\in\mathcal{K}_t^{100/200}} O_{it}^{200}}{\sum_{i\in\mathcal{K}_t^{100/200}} O_{it}^{100}}\right) & \sum_{i\in\mathcal{K}_t^{100/200}} (O_{it}^{200} + O_{it}^{100}) > 0 \\ 0 & \sum_i \left(O_{it}^{200} + \sum_{i\in\mathcal{K}_t^{100/200}} O_{it}^{100}\right) = 0 \end{cases}, \tag{2}$$

$$\mathbf{AC_t^{100/50}} = ln\left(\frac{\sum_{i\in\mathcal{K}_t^{50/100}} T_{it}^{100/200}}{\sum_{i\in\mathcal{K}_t^{50/100}} T_{it}^{50}}\right), \tag{3}$$

$$\mathbf{AO_t^{100/50}} = \begin{cases} ln\left(\frac{\sum_{i\in\mathcal{K}_t^{50/100}} O_{it}^{100/200}}{\sum_{i\in\mathcal{K}_t^{100/200}} O_{it}^{50}}\right) & \sum_{i\in\mathcal{K}_t^{50/100}} \left(O_{it}^{100/200} + O_{it}^{50}\right) > 0 \\ 0 & \sum_{i\in\mathcal{K}_t^{50/100}} \left(O_{it}^{100/200} + O_{it}^{50}\right) = 0 \end{cases}, \tag{4}$$

for dates $t = 1, \ldots, T$, where

$\mathcal{K}_t^{100/200}$: *Set of neighboring 100- and 200-strike options,*
T_{it}^{100} : *Transaction volume on the i-th 100-strike,*
T_{it}^{200} : *Transaction volume on the i-th 200-strike,*
O_{it}^{100} : *Open interest in the i-th 100-strike,*
O_{it}^{200} : *Open interest in the i-th 200-strike,*
$\mathcal{K}_t^{50/100}$: *Set of neighboring 50- and 100/200-strike options,*
T_{it}^{50} : *Transaction volume on the i-th 50-strike,*
$T_{it}^{100/200}$: *Transaction volume on the i-th 100/200-strike,*
O_{it}^{50} : *Open interest in the i-th 50-strike,*
$O_{it}^{100/200}$: *Open interest in the i-th 100/200-strike.*

Summary statistics for the four measures of clustering are reported in Table 1. For all measures means and medians are positive, indicating that there is clustering in both trading activity and open interest. The pattern of clustering in trading activity matches that in open interest. Clustering is much more severe for 50- versus 100/200-strikes than for 100- versus 200-strikes. Note that beyond the first maturity class the aggregate measure of clustering is not always defined because on some trading days there are no transactions in 50-, 100-, 200-, or 100/200-strikes. Truncation is most severe in the case of 50- versus 100/200-strikes. Among the call (put) options 316 (325) out of 1,335 observations in maturity classes two to four are truncated.[5] For 100- versus 200-strikes there are only 20 (22) truncated observations among call (put) options. Similarly, the aggregate measure for open interest is truncated, albeit to a

[5]Among these truncated observations for call (put) options only on two (one) trading days the total transaction volume on 100/200-strikes is zero. In all cases of truncation there is no volume on 50-strikes.

lesser degree than the aggregate measure of clustering.

[TABLE 1 about here]

4.1.2 Regression Results

This section describes the results of regressions using as independent variables the aggregate measures of clustering $AC_t^{100/50}$ and $AC_t^{200/100}$, respectively. Calls and puts are considered separately. As regressors we include the inverse of the DAX index level ($\frac{1}{DAX_t}$), time to maturity (ttm_t) as well as its square to account for possible non-linear maturity effects. The daily GARCH(1,1) volatility of the index logreturns (vol_t) serves as a proxy for the impact of volatility on clustering. Section 3 outlined the impact of differences in option deltas on clustering.

Neighboring options' deltas differ not only because, ceteris paribus, they have different strike prices, but also because changing the strike price implies moving along the implied volatility smile. Visual inspection of the option delta as a function of the strike price (accounting for changing implied volatility) reveals that it can be well approximated by a linear function.[6] For every trading day and maturity class we compute the derivative of the at-the-money delta with respect to the strike price. The absolute value of this derivative serves as a regressor which captures the effect of differences in deltas ($\left|\frac{\partial \delta_t}{\partial strike}\right|_{m=1}$). For near-to-maturity options, computing the implied volatility from the data makes little sense since implied volatilities are very unstable. Therefore, we include the delta-regressor only for options with time to maturity exceeding seven days. To capture the impact of open interest on clustering we include the previous trading day's value of the appropriate aggregate measure of clustering for open interest, AO_{t-1}.

Clustering of 50-strike versus 100/200-strike options

As noted in Section 4.1.1, for all but the first maturity class on a considerable number of days there are no transactions on 50-strike options. Running regressions only for those

[6] The procedure for estimating the implied volatility smile is described in Appendix A.

days on which the measure is defined potentially biases results. We use a two-stage estimation procedure to account for this problem.

First, we estimate regressions restricting the sample to the first maturity class for which the aggregate measure of clustering is always defined. Table 2 summarizes the results. The first set of regressions contains only short-term options with less than eight days to maturity and does not include the delta-regressor. The second set of regressions uses the remaining sample and includes the full set of regressors.

[TABLE 2 about here]

After estimating the models including the full set of regressors (*full models*) we test against the full models for joint significance of variables using a Wald test and re-estimate the restricted specifications.

All coefficients in the restricted specifications are significant except for some intercepts. For short-term call options, ttm_t and vol_t explain 58 percent of the variation in the clustering measure. With the same regressors and ttm_t^2, 67 percent of the variation in the clustering measure for short-term put options is explained. For call options with time to maturity exceeding seven days, an R^2 of 0.37 is achieved with ttm_t, AO_{t-1}, and vol_t. For put options, $\frac{1}{DAX_t}$ and the delta-regressor lead to an R^2 of 0.34. Contrary to the prediction in Section 3 we obtain a positive coefficient on $\frac{1}{DAX_t}$.

Next, we consider maturity classes two to four. For the pooled data we apply a two-step estimation procedure similar to that in Heckman (1979) to account for the potential selection bias introduced by the truncation of the sample when there are no transactions on 50-strikes.[7]

We assume that there exists a latent variable u_t that takes on positive values whenever the aggregate transaction volume on 50-strikes is nonzero and that takes on non-positive

[7]The incidences when 100/200-strike options have zero total transaction volume are ignored because there are too few truncated observations to estimate a second selection equation (cf. Note 5).

values otherwise, i.e.

$$u_t \begin{cases} > 0 & if \ \sum_{i \in \mathcal{K}_t^{50/100}} T_{it}^{50} > 0 \\ \leq 0 & otherwise \end{cases} . \tag{5}$$

We also assume that there exists a latent variable y_t that takes on the values of the aggregate measure of clustering $AC_t^{100/50}$ whenever it is defined, i.e. whenever $u_t > 0$, and that is unobserved otherwise, i.e

$$y_t \begin{cases} = AC_t^{100/50} & if \ u_t > 0 \\ unobserved & otherwise \end{cases} . \tag{6}$$

The two latent variables are assumed to depend on two sets of regressors, z_t and x_t,

$$u_t = z_t' \gamma + \epsilon_t \tag{7}$$

$$y_t = x_t' \beta + \nu_t, \quad t = 1, \ldots, T. \tag{8}$$

Assumption 1

The residuals ϵ_t and ν_t are bivariate normally distributed.

The variance of the residuals ϵ_t is normalized to one: $Var(\epsilon_t) \equiv 1$.

The covariance between the residuals in the two equations is a linear function of regressors x_t and z_t: $Cov(\epsilon_t, \nu_t | b_t) = b_t' \xi, \quad b_t = x_t \cup z_t, \quad t = 1, \ldots, T$.

The covariance specification captures heteroscedasticity in the latent variable equation (8) and a possibly non-constant correlation between the residuals in Eqs. (7) and (8). With these assumptions, we obtain the following proposition.

Proposition 1

Under Assumption 1,

$$AC_t^{100/50} = x_t'\beta + M_t b_t'\xi + \omega_t, \tag{9}$$

where

$$\omega_t = \nu_t - E\left[\nu_t | u_t > 0\right], \tag{10}$$

$$M_t = \frac{\phi\left(z_t'\gamma\right)}{\Phi\left(z_t'\gamma\right)}, \quad t = 1, \ldots, T, \tag{11}$$

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

The proof is a straightforward extension of that given in Heckman (1979) and is omitted. Based on Eq. (7) we estimate by maximum likelihood the probability of observing positive transaction volume on 50-strikes. The regressors z_t in the probit equation include $\frac{1}{DAX_t}$, ttm_t, ttm_t^2, vol_t , and the delta-regressor. They should impact the probability of observing transactions on 50-strikes in the same way they impact clustering. In order to avoid truncation problems in the probit estimation due to open interest, we use the simple ratio of the previous trading day's open interest in all 50-strike options over that in all 100/200-strike options as a regressor ($\frac{open50_{t-1}}{open100/200_{t-1}}$) instead of the aggregate measure of open interest. Since this is roughly the inverse of the aggregate measure for open interest, we should expect a positive coefficient.

Additionally, we include the total open interest in 50-strike options at the end of the previous trading day ($open50_t$), the number of option pairs ($pairs_t$), and the inverse of the current trading day's number of transactions ($\frac{1}{transactions_t}$). A higher level of open interest should increase the probability of observing volume on 50-strikes, since it becomes more likely that some of the option holders want to close a position. The more option pairs are included the greater should be the probability that transactions will occur on some 50-strike option. $\frac{1}{transactions_t}$ is a measure of transaction frequency. After appropriate scaling it can be interpreted as the average number of minutes that elapse between transactions on that particular trading day.[8] Larger time intervals between

[8]George and Longstaff (1993) use this measure in their study of trading activity in the S&P100 index

trades, i.e. a lower trading frequency in the option market, increase the incentive of traders to coordinate transactions on the more attractive 100/200-strikes. This effect decreases the probability of observing volume on 50-strikes. Table 3 summarizes the results. All coefficients, except those on $\frac{1}{DAX_t}$, have the expected signs. As in the first maturity class, we test against the full model using a Wald test to obtain the restricted specifications reported. Contrary to our prediction, the sign for the coefficient on $\frac{1}{DAX_t}$ is always negative and significant.

[TABLE 3 about here]

From these results we obtain an estimate of the correction term M_t for the second-step regression based on Eq. (11). Because the second-step regression relies on an estimated quantity and due to Ass. 1, the residuals in the second-step regression are heteroscedastic. To consistently estimate standard errors, we use the Newey and West (1987) estimator. For the covariance specification, based on Ass. 1, we specify the set of regressors b_t to include all the significant regressors in the restricted probit model and the full set of regressors x_t in the second-step regression. Again, a Wald test is applied to obtain the restricted specifications, which are reported together with the other results in Table 4.

[TABLE 4 about here]

Focusing on the restricted models, all regressors are significant and have the expected signs, with the exception of the coefficient on $\frac{1}{DAX_t}$, which again has the wrong sign and is significant. Fig. 1 plots the estimated covariance between the residuals in Eqs. (7) and (8).

For puts and calls the covariance is negative with few exceptions. Assuming that the correlation of the error terms is constant, we can interpret the coefficients in terms of their impact on the variance in the second-step regression. The variance of the residuals

options market. A similar measure - the inverse of the square root of the number of transactions - is used by Harris (1991) and Gwilym, Clare, and Thomas (1998) in their studies of price clustering.

for both call and put options increases with higher volatility of the DAX index returns and with the length of time intervals between transactions. Moreover, for call options, the variance increases with time to maturity and with larger values of the delta-regressor, whereas it is lower for larger values of the ratio of open interests. Intuitively, this means that the variance of the unexplained part of clustering is lower when neighboring options are better substitutes.

[FIGURE 1 about here]

It is instructive to compare these results to simple regressions, not corrected for selectivity bias, which are also reported in Table 4. For call options the coefficient on ttm_t has the wrong sign and is significant while the coefficient on the delta-regressor is not significant. Moreover, the R^2 is considerably lower than in the regression corrected for selectivity bias. Simply adding the inverse Mills ratio M_t, i.e. assuming a homoscedastic selection model, increases R^2 from 0.17 to 0.29. Similar results hold for put options. Including only the inverse Mills ratio M_t increases R^2 from 0.13 to 0.29. These results underscore the importance of accounting for incidental truncation in the regression framework.

Clustering of 100-strike versus 200-strike options

Truncation is a minor issue for the aggregate measure of clustering for 100-strike versus 200-strike options since there are only 20 (22) truncated observations among call (put) options (cf. Table 1). These are not sufficiently many to estimate a selection model. To make results comparable to the case of 50-strike versus 100/200-strike options, we run separate regressions for the first maturity class. Then we pool the data for maturity classes one to four, using only options with time to maturity exceeding seven days. Table 5 summarizes results for the first maturity class.

For call options with less than eight days to maturity only AO_{t-1} is significant and it has the right sign. The R^2 is only 0.08 compared to an R^2 of 0.61 for 50-strike versus 100/200-strike options (cf. Table 2). Only the intercept is significant for put options.

For call options with time to maturity exceeding seven days only AO_{t-1} and the delta-regressor are significant and they have the expected signs. The R^2 of 0.40 is comparable to that in the case of 50-strike versus 100/200-strike call options. For put options only the coefficient on the delta-regressor is significant and the R^2 of 0.17 is much lower than that for 50- versus 100/200-strike puts.

Table 6 summarizes results for the second set of regressions. For both call and put options all coefficients have their expected signs. The regressors ttm_t, AO_{t-1}, $\frac{1}{DAX_t}$, and the delta-regressor explain roughly 30 percent of the variation in the aggregate measure of clustering for call options. In the case of 100- versus 200-strike put options ttm_t, vol_t, and the delta-regressor explain roughly 17 percent of the variation in clustering.

[TABLE 5 about here]

[TABLE 6 about here]

4.2 Pairwise Clustering in Transaction Volume

4.2.1 Pairwise Measures of Clustering

We define pairwise measures of clustering to gauge more closely how the factors identified in Section 3 affect the degree of substitution between individual options. That is, for every option pair a separate measure of clustering is computed for all dates. For example, the pairwise measure of clustering between the i-th pair of 200- and 100-strike options, $PC_{it}^{200/100}$, is defined as the logarithm of the ratio of the number of transactions on the i-th 200-strike over the number of transactions on the i-th 100-strike on date t. Accordingly, we compute the pairwise measure of trade clustering between the i-th 50- and 100/200-strike options on date t, $PC_{it}^{100/50}$. Moreover, we define analogous measures of clustering for the open interest, $PO_{it}^{200/100}$ and $PO_{it}^{100/50}$ with the convention that these are set equal zero when both numerator and denominator are zero. Formally,

$$
\begin{aligned}
\mathbf{PC_{it}^{200/100}} &= ln\left(\frac{T_{it}^{200}}{T_{it}^{100}}\right), && (12)\\
\mathbf{PO_{it}^{200/100}} &= \begin{cases} ln\left(\frac{O_{it}^{200}}{O_{it}^{100}}\right) & O_{it}^{200}+O_{it}^{100}>0 \\ 0 & O_{it}^{200}+O_{it}^{100}=0 \end{cases}, && (13)\\
& i \in \mathcal{K}_t^{100/200}, \quad t=1,\ldots,T, \\
\mathbf{PC_{it}^{100/50}} &= ln\left(\frac{T_{it}^{100/200}}{T_{it}^{50}}\right), && (14)\\
\mathbf{PO_{it}^{100/50}} &= \begin{cases} ln\left(\frac{O_{it}^{100/200}}{O_{it}^{50}}\right) & O_{it}^{100/200}+O_{it}^{50}>0 \\ 0 & O_{it}^{100/200}+O_{it}^{50}=0 \end{cases}, && (15)\\
& i \in \mathcal{K}_t^{50/100}, \quad t=1,\ldots,T,
\end{aligned}
$$

where

$\mathcal{K}_t^{100/200}$: Set of neighboring 100- and 200-strike options,
T_{it}^{100} : Transaction volume on the i-th 100-strike,
T_{it}^{200} : Transaction volume on the i-th 200-strike,
O_{it}^{100} : Open interest in the i-th 100-strike,
O_{it}^{200} : Open interest in the i-th 200-strike,
$\mathcal{K}_t^{50/100}$: Set of neighboring 50- and 100/200-strike options,
T_{it}^{50} : Transaction volume on the i-th 50-strike,
$T_{it}^{100/200}$: Transaction volume on the i-th 100/200-strike,
O_{it}^{50} : Open interest in the i-th 50-strike,
$O_{it}^{100/200}$: Open interest in the i-th 100/200-strike.

Table 7 reports summary statistics for the measures of clustering. As for the aggregate measure, means and medians are always positive and clustering is greater for 50- versus 100/200-strike options than for 100- versus 200-strike options. In all cases the samples of the measures for trade clustering are severely truncated because often there are no trades on individual options. Roughly three quarters of the sample for 50- versus 100/200-strike options are truncated for both puts and calls. In the case of 100- versus 200-strike options roughly one half of the sample for call option pairs and about one third of the sample for put options are truncated.

[TABLE 7 about here]

4.2.2 Regression Results

Selection bias is a potential problem for regressions using the pairwise measure, and we carry out a two-step estimation procedure similar to the one employed for the aggregate measure of clustering in Section 4.2.2. Because individual options for all strike classes often are not traded we have to estimate three selection equations.

We assume there exist three latent variables, u_{it}^{50}, u_{it}^{100}, and u_{it}^{200}, which take on positive values whenever the transaction volume on the option in the corresponding strike class is nonzero, and which take on non-positive values otherwise, i.e.

$$u_{it}^{k} \quad \begin{cases} > 0 & if\ T_{it}^{k} > 0 \\ \leq 0 & otherwise \end{cases}, \quad i \in \mathcal{K}_t^k, \quad t = 1, \ldots, T, k \in \{50, 100, 200\}, \tag{16}$$

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

To fix ideas, we focus on the case of 100- vs 200-strike options in the following. We assume there exists a latent variable y_{it} that takes on the values of the pairwise measure of clustering $PC^{200/100}$ whenever it is defined, and that is unobserved otherwise. The pairwise measure $PC_{it}^{200/100}$ is only defined if both options in the pair have positive transaction volume. That is,

$$y_{it} \quad \begin{cases} = PC_{it}^{200/100} & if\ u_{it}^{100} > 0 \text{ and } u_{it}^{200} > 0 \\ unobserved & otherwise \end{cases}, \quad i \in \mathcal{K}_t^k,\ t = 1, \ldots, T, \tag{17}$$

where $\mathcal{K}_t^{100/200}$ is the set of neighboring 100- and 200-strike options on day t.

The three latent variables depend on three sets of regressors, z_{it}^{100}, z_{it}^{200}, and $x_{it}^{100/200}$, and their residuals are assumed to be uncorrelated.

$$u_{it}^{100} = \left(z_{it}^{100}\right)'\gamma^{100} + \epsilon_{it}^{100}, \tag{18}$$

$$u_{it}^{200} = \left(z_{it}^{200}\right)'\gamma^{200} + \epsilon_{it}^{200}, \quad i \in \mathcal{K}_t^k, \quad t = 1, \ldots, T, \tag{19}$$

$$y_{it} = x_{it}'\beta + \nu_{it}. \tag{20}$$

Assumption 2

The residuals ϵ_{it}^{100}, ϵ_{it}^{200}, and ν_{it} are trivariate normally distributed.

ϵ_{it}^{100} and ϵ_{it}^{200} are uncorrelated and their variances are normalized to one:

$var(\epsilon_{it}^{100}) \equiv var(\epsilon_{it}^{200}) \equiv 1.$

The covariances between the residuals in Eqs. (18) and (20) and between the residuals in Eqs. (19) and (20) are linear functions of regressors $x_{it} \cup z_{it}^{100}$ and $x_{it} \cup z_{it}^{200}$, respectively:

$Cov[\epsilon_{it}^{k}, \nu_{it} | b_{it}^{k}] = \left(b_{it}^{k}\right)' \xi_t^k, \quad b_{it}^{k} = x_{it} \cup z_{it}^{k}, \quad k \in \{100, 200\}, \quad i \in \mathcal{K}_t^k, \quad t = 1, \ldots, T.$

With the above assumption we obtain the following proposition.

Proposition 2

Under Assumption 2,

$$PC_{it}^{200/100} = x_{it}'\beta + M_{it}^{100}\left(b_{it}^{100}\right)'\xi_t^{100} + M_{it}^{200}\left(b_{it}^{200}\right)'\xi_t^{200} + \omega_{it}, \tag{21}$$

where

$$\omega_{it} = \nu_{it} - E\left[\nu_{it} | u_{it}^{100} > 0, u_{it}^{200} > 0\right], \tag{22}$$

$$M_{it}^{k} = \frac{\phi\left(\left(z_{it}^{k}\right)'\gamma^k\right)}{\Phi\left(\left(z_{it}^{k}\right)'\gamma^k\right)}, \quad k \in \{100, 200\}, \quad i \in \mathcal{K}_t^k, \quad t = 1, \ldots, T, \tag{23}$$

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

The proof is again a straightforward extension of that in Heckman (1979). For 50- versus 100/200-strikes Ass. 2 applies 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.

Probit estimation

In the selection equations the set of regressors z_{it} comprises $\frac{1}{DAX_{it}}$, ttm_{it}, ttm_{it}^2, and vol_{it}, as in the selection equation for the aggregate measure of clustering in Section 4.1.2. The absolute value of the option's delta captures the impact of the option's risk on the probability of observing transactions. Additional regressors are the absolute moneyness of the option, defined as $\left|\frac{strike}{DAX_{it}} - 1\right|$, and interactions with ttm_{it} and ttm_{it}^2. Moreover, we include the option's open interest at the end of the previous trading day,

$O_{i,t-1}$. We include the interaction $ttm_{it}O_{i,t-1}$ to test the prediction in Section 3 that open interest should matter more for the choice of strike prices when time to maturity is long.

For reasons of brevity we report results only for the pooled data with time to maturity exceeding seven days, summarized in Table 8. The coefficients on ttm_{it}, ttm_{it}^2, $O_{i,t-1}$, and the delta-regressors are all significant and have the expected signs. Coefficients on the interactions between moneyness and ttm_{it} also have the expected signs when they are significant. The interaction $ttm_{it}O_{i,t-1}$ is significant only in the probit equations for 200-strike options. Coefficients take on negative values for both call and put options, contrary to our predictions.

[TABLE 8 about here]

For call options the coefficient on $\frac{1}{DAX_{it}}$ always has a negative sign and is significant except for 200-strike options. For put options coefficients are significant only for 100- and 200-strike options, and they are positive. Some care is required in interpreting these results. For small values of $\frac{1}{DAX_{it}}$ the economic importance of strike price differences decreases, and options become better substitutes. In the case of 100-strike options, some transactions shift from 50-strike options to 100-strike options, increasing the probability of observing positive volume on 100-strikes. However, some transactions shift from 100-strike options to 200-strike options as well, decreasing the probability of observing positive volume on 100-strikes. While the impact on 100-strikes is ambiguous, for 50-strikes this should reduce the probability of observing volume and increase it for 200-strikes. Hence, the signs of the coefficients on $\frac{1}{DAX_{it}}$ for 200-strike call options are inconsistent with our hypothesis.

The regressor vol_{it} always has a significant coefficient that is positive for 100- and 200-strike options and that is negative for 50-strikes. 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, following a similar reasoning to that for $\frac{1}{DAX_{it}}$, and the data reveal that the increased demand effect prevails over the substitution effect. This is not surprising, since clustering of 100- versus 200-strikes is typically less pronounced than clustering of 50-strikes versus 100/200-strikes. In contrast, for 50-strikes the substitution effect outweighs the increased demand effect.

Second-stage estimation

In the second-step regressions the set of regressors x_{it} contains $\frac{1}{DAX_{it}}$, ttm_{it}, ttm_{it}^2, and vol_{it}, as well as the pairwise measure for open interest $PO_{i,t-1}$, which have the same interpretation as in the regressions for the aggregate measure of clustering in Section 4.1.2. The interaction between ttm_{it} and $PO_{i,t-1}$ captures whether the importance of open interest diminishes with shorter remaining life time of the options.

To capture the impact of differences in option deltas we include additional regressors. If the attractive option has a higher delta than its counterpart it becomes less attractive relative to the neighbor because it is riskier. To illustrate how this affects clustering, consider two neighboring 100- and 200-strike call options. For small differences between the two options' deltas options are good substitutes and trades should concentrate on the more attractive option. However, if the 200-strike has a higher delta than its 100-strike counterpart, this has a countervailing effect on clustering since some traders prefer the less risky 100-strike option over the 200-strike. In the other case, if the 100-strike has a higher delta than the 200-strike, this exacerbates clustering.

To measure differences in deltas we use the absolute value of the log ratio of the two deltas, i.e. $|ln\,(\delta_{it}^1) - ln\,(\delta_{it}^2)|$. Because the impact of differences in options' deltas is predicted to be asymmetric we include two regressors. The first regressor takes on the absolute value of the log ratio of the two deltas whenever the option from the lower strike class (i.e. 50- or 100-strike) has a larger absolute delta than its counterpart (i.e. 100/200- or 200-strike) and takes on zero otherwise. The second regressor takes on the absolute value of the log ratio of the two deltas whenever the option from the higher strike class has a larger absolute delta than its counterpart and takes on zero otherwise. Moreover, when comparing transaction volumes for two neighboring options, one would expect the option that is farther away from the money to be less actively traded, ceteris paribus. For example, for 100- versus 200-strike call options, clustering should increase for option pairs that are farther away from the at-the-money point. If the 100-strike's absolute distance from the at-the-money point is greater than that for the 200-strike, this exacerbates clustering. However, if the 200-strike lies farther away from the money than the 100-strike, clustering will be less severe since some trades are drawn away from the 200-strike. To account for this we include the average absolute moneyness of the two options in the pair ($\frac{\left|m_{it}^1+m_{it}^2\right|}{2}$) and a dummy ($I_{it}^m$) which takes on value one if the option from the lower strike class (i.e. 50- or 100-strike) has a larger absolute moneyness than its counterpart (i.e. 100/200- or 200-strike) and takes on zero otherwise.

To correct for potential selection biases we include the correction terms M_{it}^k obtained from the probit estimation. The set of regressors b_{it}^k in the covariance terms $\left(b_{it}^k\xi^k\right)'$ includes all the regressors in x_{it} as well as those from the corresponding probit estimations, z_{it}^k. To avoid collinearity problems the average value of absolute moneyness is used instead of the corresponding variables in the probit equations and only the delta-regressors that capture differences in deltas are included.

Results for the second-stage regressions are reported in Table 9. The coefficients on ttm_{it} and $PO_{i,t-1}$ are always significant and have the right signs. The interaction between $PO_{i,t-1}$ and ttm_{it} is significant only for 100-strike versus 200-strike options. It always has a negative coefficient, contrary to our predictions. Clustering for 50- versus 100/200-strike options decreases when the DAX index level increases, as in the case of aggregate clustering. In contrast, for 100-strike versus 200-strike options the coefficient on $\frac{1}{DAX_{it}}$ has the expected sign. The coefficient on vol_{it} always has the right sign, when it is significant. For 100-strike versus 200-strike options, clustering increases with moneyness, as expected. Moneyness has the opposite effect on clustering for 50-strike versus 100/200-strike call options. We do not have an explanation for this result. The dummy I_{it}^m is always significant with the predicted positive coefficient. There is no clear pattern for the interaction between I_{it}^m and moneyness. The absolute value of the first delta-regressor is smaller than the absolute value of the second delta-regressor in all cases,

which confirms our conjecture about the impact of options' riskiness on clustering.
Selection is important for 50-strike versus 100/200-strike options. As for the aggregate measure of clustering ttm is significant and has the wrong sign in the regressions not corrected for selectivity bias. Adding the two inverse Mills ratios M_{it}^{k} to these regressions increases R^2 from 0.25 (0.21) to 0.41 (0.41) for call (put) options. For 100-strike versus 200-strike options selection is less important. Adding the two inverse Mills ratios to the simple regressions not corrected for selectivity bias only leads to a minor increase in R^2, from 0.14 (0.13) percent to 0.16 (0.14) for call (put) options.

[TABLE 9 about here]

5 Conclusion

This paper analyzes the impact of options' characteristics on the cross-sectional distribution of trading activity in contracts with different strike prices in the DAX index options market. In this market trading clusters around particular strike prices.
The 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.
The empirical analysis is based on two measures of clustering of trading activity which we regress on options' characteristics. The first measure of clustering is roughly the log of the ratio of aggregate transaction volumes in two different strike classes. In the case of 50-strike versus 100/200-strike options, this aggregate measure of clustering is not always defined since on some trading days there is no volume on 50-strike options. We find that this sample truncation leads to selectivity bias in a simple regression framework. To overcome this problem we use a two-step estimation procedure similar to the classic Heckman (1979) approach.

The second measure of clustering is the log ratio of transaction volume on two options with neighboring strike prices. It allows for a more detailed analysis of how individual options' characteristics, such as moneyness, affect transaction clustering. For this measure the sample is severely truncated because individual options in all strike classes frequently witness no turnover, and the measure is not defined. To correct for selectivity bias we again use a Heckman-style two-step estimation procedure. This correction appears to be particularly important for the 50-strike versus 100/200-strike case.

Our analysis finds that differences in open interest and other factors, such as time to maturity, the volatility of DAX index returns, options' moneyness, and the options' deltas, impact clustering of trading activity. The signs of the coefficients in the regressions generally confirm our hypotheses about the way in which these factors affect clustering and the attractiveness of different strike classes.

To our knowledge this paper is the first to analyze the impact of the strike price grid on the cross-sectional distribution of trading volume in options markets. In future research we want to establish the relationship between strike price gradations, the cross-sectional distribution of trading volume and overall trading volume. This would have 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 5-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.[9] That is,

$$F_{t,m}(T_F) = S_{t,m} e^{r(T_F - t)}, \tag{24}$$

where

T_F : future's maturity,

$S_{t,m}$: (implied) underlying index in the mth 5-minute interval on day t,

$F_{t,m}$: nearest-to-maturity futures contract in the mth 5-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 Eq. (24) and using the average futures price over the respective 5-minute interval. The appropriate risk-free interest rate is obtained by linearly interpolating EUR-Libor rates bracketing the option's maturity.[10] 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 Scholes (1973) formula. Following Hafner and Wallmeier (2000), we approximate the smile on every trading day by fitting a smooth differentiable spline function whose segments join at the at the money point via ordinary least squares. The general specification allows for quadratic function segments for the in- and out-of-the-money ranges, respectively. That is,

$$\begin{aligned} \sigma_{IV} &= \alpha_0 + \alpha_1 M + \alpha_2 M^2 + D\left(\beta_0 + \beta_1 M + \beta_2 M^2\right) + \epsilon \qquad (25) \\ D &= \begin{cases} 0 \quad if \quad M \leq 1, & \text{(strike below the at-the-money point)} \\ 1 \quad if \quad M > 1 & \text{(strike above the at-the-money point)} \end{cases} . \end{aligned}$$

[9] The DAX index is computed assuming reinvestment of dividends after corporate income tax on distributed gains (Deutsche Börse 2000b). German income tax law in effect in 1999 and 2000 treats dividends as if they included 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 (2000). In our data this problem appears to be a minor one.

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

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 minimum at $M^* = 1$, i.e.

$$\left.\frac{d\left(\beta_0+\beta_1 M+\beta_2 M^2\right)}{dM}\right|_{M^*=1} = \beta_1 + 2\beta_2 = 0. \tag{26}$$

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

$$\sigma_{IV} = \alpha_0 + \alpha_1 M + \alpha_2 M^2 + \alpha_3 D\left(1 - 2M + M^2\right) + \epsilon. \tag{27}$$

References

Ball, Clifford A., Walter N. Torus, 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.

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.

George, Thomas J., and Francis A. Longstaff, 1993, Bid-ask spreads and trading activity in the S&P 100 index options market, *Journal of Financial and Quantitative Analysis* 28, 381–97.

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, 2000, The dynamics of DAX implied volatilities, Proceedings of the 13th Annual CBOT Research Symposium.

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

——— , 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.

Newey, Whitney, and Kenneth West, 1987, A simple positive semi-definite, heteroscedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703–8.

Maturity Class	**1**		**2**		**3**		**4**	
	AC	**AO**	**AC**	**AO**	**AC**	**AO**	**AC**	**AO**
Aggregate Measure of Clustering for 50- vs 100/200-strike Call Options								
Mean	2.72	2.28	3.32	2.45	2.71	2.22	2.76	3.70
Median	2.76	2.25	3.25	2.41	2.77	2.20	2.83	3.41
Maximum	6.92	3.27	6.14	3.78	5.63	5.10	5.31	9.54
Minimum	0.45	1.26	0.74	0.72	-2.22	-2.06	-0.41	2.04
Std. Dev.	0.81	0.35	0.90	0.46	1.24	0.95	0.94	1.48
Observations	445	444	420	444	321	429	278	418
Total number of days	445	445	445	445	445	445	445	445
Truncated observations	0	1	25	1	124	16	167	27
Aggregate Measure of Clustering for 50- vs 100/200-strike Put Options								
Mean	2.80	2.45	3.47	2.61	2.80	2.30	2.92	3.43
Median	2.80	2.46	3.43	2.56	2.85	2.47	2.89	3.16
Maximum	6.54	3.47	6.85	4.03	5.75	6.14	5.51	10.44
Minimum	-0.29	1.69	1.18	1.74	-2.71	-4.59	-0.13	1.83
Std. Dev.	0.90	0.36	0.86	0.40	1.24	1.12	0.99	1.22
Observations	445	444	413	444	316	424	281	415
Total number of days	445	445	445	445	445	445	445	445
Truncated observations	0	1	32	1	129	21	164	30
Aggregate Measure of Clustering for 100- vs 200-strike Call Options								
Mean	0.20	0.38	0.43	0.57	0.54	0.60	1.01	1.09
Median	0.19	0.36	0.38	0.56	0.55	0.65	1.04	1.16
Maximum	1.52	1.23	2.61	1.41	2.88	3.45	3.49	1.78
Minimum	-1.05	-0.37	-0.90	-0.53	-2.97	-2.27	-2.08	-0.54
Std. Dev.	0.31	0.33	0.45	0.37	0.78	0.73	0.86	0.50
Observations	445	444	445	444	426	440	444	444
Total number of days	445	445	445	445	445	445	445	445
Truncated observations	0	1	0	1	19	5	1	1
Aggregate Measure of Clustering for 100- vs 200-strike Put Options								
Mean	0.26	0.39	0.49	0.58	0.61	0.69	0.92	0.72
Median	0.27	0.38	0.46	0.62	0.60	0.69	0.90	0.79
Maximum	1.58	0.76	2.41	1.47	3.43	6.27	3.92	1.42
Minimum	-1.05	-0.03	-0.63	-0.08	-1.61	-1.39	-2.14	-0.13
Std. Dev.	0.26	0.17	0.36	0.30	0.70	0.75	0.70	0.31
Observations	445	444	445	444	423	429	445	444
Total number of days	445	445	445	445	445	445	445	445
Truncated observations	0	1	0	1	22	16	0	1

Table 1: Summary Statistics for the Aggregate Measure of Clustering

Variable[a]	Maturity Class 1 and ttm<8				Maturity Class 1 and ttm>7			
	Call Options		**Put Options**		**Call Options**		**Put Options**	
	Full Model	Restricted	Full Model	Restricted	Full Model	Restricted	Full Model	Restricted
C	0.020	-0.006	-0.461	-0.247	0.927	0.273	0.616	1.984***
	(0.422)	(0.172)	(0.573)	(0.249)	(0.746)	(0.340)	(0.741)	(0.330)
ttm_t	0.256***	0.183***	0.314***	0.334***	-0.010	0.023***	0.051	
	(0.058)	(0.027)	(0.058)	(0.058)	(0.037)	(0.006)	(0.046)	
ttm_t^2	-0.012		-0.021***	-0.022***	0.001		-0.001	
	(0.008)		(0.008)	(0.008)	(0.001)		(0.001)	
AO_{t-1}	0.251		0.205		0.428***	0.429***	0.150	
	(0.182)		(0.203)		(0.135)	(0.142)	(0.158)	
$\frac{1}{DAX_t}$	-2820.167		-869.539		2744.295		12425.73***	14060.7***
	(2107.243)		(2037.849)		(1874.795)		(2898.517)	(2237.478)
vol_t	5.969***	6.486***	6.482***	6.733***	4.404***	5.716***	2.133	
	(0.950)	(0.568)	(1.131)	(0.991)	(1.368)	(1.057)	(1.586)	
$\left\lvert \frac{\partial \delta_t}{\partial strike} \right\rvert_{m=1}$					-307.389		-652.1***	-890.249***
					(187.929)		(220.108)	(107.637)
R^2	0.613	0.580	0.674	0.666	0.379	0.369	0.357	0.337
Wald Test[b]		7.204		2.369				5.259
		(0.066)		(0.306)				(0.262)

[a]Regression coefficients and Newey-West standard errors are reported. *** stands for 1 percent, ** for 5 percent, and * for 10 percent significance levels, respectively.

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

Table 2: Regressions: Aggregate Measure for 50- vs 100/200-Strikes (Maturity Class 1)

Variable[a]	Probit for Maturity Classes 2 - 4			
	Call Options		**Put Options**	
	Full Model	Restricted	Full Model	Restricted
C	3.008*** (0.774)	3.1*** (0.293)	2.426*** (0.771)	1.416*** (0.321)
ttm_t	-0.015* (0.009)	-0.008*** (0.001)	-0.010 (0.009)	
ttm_t^2	4.29E-5 (3.52E-5)		3.30E-5 (3.48E-5)	
$\frac{open50_{t-1}}{open100/200_{t-1}}$	0.735** (0.289)	0.585** (0.255)	0.151 (0.188)	
$open50_{t-1}$	2.26E-5** (8.92E-6)	2.92E-5*** (6.78E-6)	1.10E-5* (6.35E-6)	
$pairs_t$	0.011 (0.008)		0.025*** (0.008)	0.034*** (0.006)
$\frac{1}{DAX_t}$	-9878.327*** (2383.312)	-9580.027*** (1630.471)	-14697.81*** (2387.859)	-15849.520*** (1717.371)
vol_t	-1.127 (1.294)		-0.227 (1.210)	
$\left\|\frac{\partial \delta_t}{\partial strike}\right\|_{m=1}$	650.067 (593.640)		1579.597** (627.781)	2411.818*** 265.232
$\frac{1}{transactions_t}$	-16.581*** (3.138)	-18.271*** (2.995)	-15.825*** (2.918)	-14.917*** (2.825)
Wald Test[b]		9.334 (0.053)		5.104 (0.403)

[a]Regression coefficients and Huber-White standard errors are reported. *** stands for 1 percent, ** for 5 percent, and * for 10 percent significance levels, respectively.

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

Table 3: Selection Equations: Aggregate Measure for 50- vs 100/200-Strikes

Variable[a]	Regressions for Maturity Classes 2 - 4					
	Call Options			Put Options		
	Not corrected	Full Model	Restricted	Not corrected	Full Model	Restricted
C	1.863*** (0.624)	1.514** (0.705)	1.477*** (0.384)	1.882*** (0.63)	2.554*** (0.761)	2.119*** (0.266)
ttm_t	-0.023*** (0.008)	-0.007 (0.008)		-0.009 (0.008)	-0.006 (0.009)	
ttm_t^2	6.00E-5** (3.54E-5)	4.51E-5 (4.71E-5)		8.03E-6 (3.73E-5)	6.81E-5 (5.63E-5)	
AO_{t-1}	0.261*** (0.052)	0.321*** (0.084)	0.241*** (0.051)	0.193*** (0.068)	-0.06 (0.092)	0.082** (0.038)
$\frac{1}{DAX_t}$	8844.847*** (2035.211)	11412.92*** (2612.130)	10349.15*** (1711.222)	1822.37 (2548.413)	13362.12*** (3158.499)	13976.37*** (2024.914)
vol_t	2.382505** (1.06)	2.181** (1.169)	2.731*** (0.949)	3.927*** (0.918)	1.354 (1.68)	
$\left\|\frac{\partial \delta_t}{\partial strike}\right\|_{m=1}$	-436.689 (415.448)	-1190.059*** (435.542)	-1075.866*** (266.617)	290.971 (456.136)	-1108.831** (445.236)	-1062.419*** (232.333)
M_t		-7.758*** (1.649)	-9.344*** (1.477)		-4.201** (2.02)	-3.859*** (0.462)
$M_t ttm_t$		0.016** (0.009)	0.02*** (0.004)		-0.012 (0.012)	
$M_t AO_{t-1}$		-0.224 (0.179)			0.208 (0.139)	
$M_t \frac{1}{DAX_t}$		-4396.706 (6747.688)			12470.71** (6176.325)	
$M_t vol_t$		5.964** (2.653)	4.597** (1.869)		1.505 (3.267)	4.205** (1.918)
$M_t \left\|\frac{\partial \delta_t}{\partial strike}\right\|_{m=1}$		6676.570*** (1712.796)	6837.951*** (1378.459)		-1371.134 (2378.975)	
$M_t pairs_t$					-0.033 (0.02)	
$M_t \frac{open50_{t-1}}{open100/200_{t-1}}$		-2.003*** (0.635)	-1.661*** (0.610)			
$M_t open50_{t-1}$		1.45E-5 (2.78E-5)				
$M_t \frac{1}{transactions_t}$		9.054*** (2.66)	8.670*** (2.494)		11.343** (4.454)	5.963** (2.828)
R^2	0.166	0.329	0.326	0.13	0.32	0.313
Wald Test[b]			5.94 (0.312)			10.56 (0.228)

[a]Regression coefficients and Newey-West standard errors are reported. *** stands for 1 percent, ** for 5 percent, and * for 10 percent significance levels, respectively.

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

Table 4: Regressions: Aggregate Measure for 50- vs 100/200-Strikes

Variable[a]	Maturity Class 1 and ttm<8				Maturity Class 1 and ttm>7			
	Call Options		**Put Options**		**Call Options**		**Put Options**	
	Full Model	Restricted	Full Model	Restricted	Full Model	Restricted	Full Model	Restricted
C	0.179 (0.323)	0.042 (0.059)	0.254 (0.277)	0.236*** (0.039)	0.837*** (0.311)	0.321*** (0.06)	0.509** (0.226)	0.532*** 0.047
ttm_t	-0.079 (0.063)		-0.045 (0.065)		-0.021 (0.019)		-0.004 (0.014)	
ttm_t^2	0.009 (0.007)		0.003 (0.007)		3.16E-4 (8.08E-4)		2.94E-3 (1.08E-3)	
AO_{t-1}	0.324*** (0.122)	0.327*** (0.11)	0.077 (0.219)		0.306*** (0.076)	0.393*** (0.07)	0.064 (0.103)	
$\frac{1}{DAX_t}$	-170.905 (1840.309)		442.998 (1426.146)		-1531.549 (989.633)		-146.141 (742.612)	
vol_t	-0.021 (0.468)		-0.16 (0.526)		0.199 (0.463)		0.265 (0.471)	
$\left\lvert \frac{\partial \delta_t}{\partial strike} \right\rvert_{m=1}$					-229.061*** (93.04)	-193.683*** (39.732)	-191.412** (80.81)	-196.418*** (34.229)
R^2	0.099	0.078803	0.023	0	0.436	0.401	0.18	0.171
Wald Test[b]		0.042 0.479		2.945 (0.708)		8.065318 (0.089)		1.755 (0.882)

[a]Regression coefficients and Newey-West standard errors are reported. *** stands for 1 percent, ** for 5 percent, and * for 10 percent significance levels, respectively.

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

Table 5: Regressions: Aggregate Measure for 100- vs 200-Strikes (Maturity Class 1)

Variable[a]	Maturity Classes 1 - 4 and ttm>7			
	Call Options		Put Options	
	Full Model	Restricted	Full Model	Restricted
C	0.661*** (0.198)	0.735*** (0.156)	0.326** (0.164)	0.266** (0.113)
ttm_t	0.003 (0.003)	0.003*** (0.001)	0.003 (0.003)	0.004*** (0.001)
ttm_t^2	4.33E-06 (1.45E-05)		4.03E-06 (1.31E-05)	
AO_{t-1}	0.372*** (0.055)	0.380*** (0.055)	0.071 (0.061)	
$\frac{1}{DAX_t}$	-3566.273 (906.953)	-3223.770*** (765.0142)	-891.983 (843.271)	
vol_t	0.595 (0.398)		1.198*** (0.377)	1.04*** (0.308)
$\left\|\frac{\partial \delta_t}{\partial strike}\right\|_{m=1}$	-139.329 (87.195)	-157.65*** (56.709)	-192.067*** (71.280)	-225.79*** 51.663
R^2	0.297	0.295	0.184	0.174
Wald Test[b]		3.198 (0.202)		3.177 (0.365)

[a]Regression coefficients and Newey-West standard errors are reported. *** stands for 1 percent, ** for 5 percent, and * for 10 percent significance levels, respectively.

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

Table 6: Regressions: Aggregate Measure for 100- vs 200-Strikes (Maturity Classes 1-4)

Maturity Class	1		2		3		4	
	AC	AO	AC	AO	AC	AO	AC	AO
Pairwise Measure of Clustering for 50- vs 100/200-strike Call Options								
Mean	2.48	2.62	2.40	2.73	1.43	2.69	1.11	3.36
Median	2.50	2.50	2.48	2.67	1.39	2.65	1.10	3.41
Maximum	6.12	9.86	5.44	9.87	4.75	9.83	4.20	9.17
Minimum	-2.61	-2.39	-1.79	-3.53	-2.30	-3.56	-2.08	-5.28
Std. Dev.	1.27	1.38	1.22	1.60	1.19	1.95	1.15	1.95
Observations	6,121	13,649	3,329	11,611	1,206	7,764	904	9,207
Total number of pairs	14,300	14,300	12,967	12,967	9,795	9,795	11,377	11,377
Censored observations	8,179	651	9,638	1,356	8,589	2,031	10,473	2,170
Pairwise Measure of Clustering for 50- vs 100/200-strike Put Options								
Mean	2.51	2.72	2.44	2.88	1.49	2.74	1.13	3.45
Median	2.53	2.65	2.54	2.91	1.61	2.91	1.10	3.56
Maximum	6.22	8.50	5.24	9.43	4.94	9.87	4.57	9.58
Minimum	-1.90	-1.90	-2.77	-2.30	-2.94	-5.53	-3.78	-4.33
Std. Dev.	1.24	1.19	1.14	1.46	1.24	1.77	1.19	1.84
Observations	7,458	13,440	3,657	11,490	1,312	7,875	964	8,808
Total number of pairs	14,300	14,300	12,967	12,967	9,795	9,795	11,377	11,377
Censored observations	6,842	860	9,310	1,477	8,483	1,920	10,413	2,569
Pairwise Measure of Clustering for 100- vs 200-strike Call Options								
Mean	0.25	0.74	0.39	0.83	0.36	0.95	0.42	1.56
Median	0.21	0.56	0.30	0.71	0.29	0.88	0.29	1.42
Maximum	5.00	6.10	4.61	6.72	4.19	6.85	3.97	8.54
Minimum	-5.21	-4.69	-3.40	-6.67	-4.09	-6.42	-4.20	-4.47
Std. Dev.	1.10	1.17	1.04	1.32	1.17	1.51	1.25	1.42
Observations	6,489	9,947	5,408	8,956	3,577	6,762	3,882	10,406
Total number of pairs	10,160	10,160	9,386	9,386	7,405	7,405	10,754	10,754
Censored observations	3,671	213	3,978	430	3,828	643	6,872	348
Pairwise Measure of Clustering for 100- vs 200-strike Put Options								
Mean	0.20	0.62	0.42	0.79	0.42	0.84	0.40	1.06
Median	0.20	0.52	0.42	0.69	0.41	0.75	0.33	0.93
Maximum	4.61	7.23	4.30	7.23	3.99	7.95	4.44	5.98
Minimum	-4.47	-3.49	-4.08	-3.73	-3.74	-5.40	-3.85	-5.03
Std. Dev.	1.00	1.06	1.02	1.12	1.20	1.28	1.25	1.18
Observations	7,333	9,944	6,973	9,013	4,616	6,863	5,250	10,462
Total number of pairs	10,160	10,160	9,386	9,386	7,405	7,405	10,754	10,754
Censored observations	2,827	216	2,413	373	2,789	542	5,504	292

Table 7: Summary Statistics for the Pairwise Measure of Clustering

Variable[a]	50-strike Options				100-strike Options				200-strike Options			
	Call Options		Put Options		Call Options		Put Options		Call Options		Put Options	
	Full Model	Restricted	Full Model	Restricted	Full Model	Restricted	Full Model	Restricted	Full Model	Restricted	Full Model	Restricted
C	3.242***	3.306***	2.94***	2.991***	2.975***	2.953***	4.007***	4.031***	2.281***		2.592***	2.649***
	(0.098)	(0.091)	(0.092)	(0.072)	(0.120)	(0.117)	(0.122)	(0.118)	(0.096)		(0.100)	(0.092)
ttm_{it}	-0.039***	-0.04***	-0.042***	-0.042***	-0.021***	-0.021***	-0.036***	-0.036***	-0.023***		-0.025***	-0.025***
	(1.294E-4)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)		(0.001)	(0.001)
ttm^2_{it}	1.56E-4***	1.62E-4***	1.69E-4***	1.71E-4***	6.37E-5***	6.31E-5***	1.16E-4***	1.17E-4***	7.84E-5***		8.09E-5***	8.14E-5***
	(6.44E-6)	(5.86E-6)	(6.23E-6)	(5.97E-6)	(7.40E-6)	(7.38E-6)	(7.82E-6)	(7.48E-6)	(6.17E-6)		(5.94E-6)	(5.92E-6)
$O_{i,t-1}$	1.18E-4***	1.19E-4***	4.02E-5**	3.14E-5***	1.24E-4***	1.29E-4***	6.51E-5***	6.20E-5***	5.05E-5***		3.88E-5***	3.84E-5***
	(2.22E-5)	(1.21E-5)	(1.75E-5)	(9.96E-6)	(1.23E-5)	(5.86E-6)	(6.10E-6)	(4.04E-6)	(3.94E-6)		(3.32E-6)	(3.32E-6)
$ttmO_{i,t-1}$	1.80E-8		-2.10E-7		7.65E-8		-5.50E-8		3.63E-7***		-1.30E-7***	-1.27E-7***
	(4.37E-7)		(3.52E-7)		(1.41E-7)		(8.09E-8)		(4.45E-8)		(3.54E-8)	(3.53E-8)
$\frac{1}{DAX_{it}}$	-2678.073***	-2735.221***	195.384***		-4424.522***	-4396.766***	3563.405***	3564.088***	-2145.287***		7219.682***	7359.672***
	(422.602)	(420.51)	(399.339)		(459.815)	(456.645)	(456.893)	(456.902)	(348.226)		(363.566)	(357.469)
vol_{it}	-2.252***	-2.321***	-2.399***	-2.395***	2.351***	2.343***	0.525**	0.527**	3.052***		0.33*	
	(0.259)	(0.255)	(0.232)	(0.229)	(0.234)	(0.232)	(0.230)	(0.230)	(0.193)		(0.200)	
$\vert\delta_{it}\vert$	-2.228***	-2.228***	-2.479***	-2.475***	-1.88***	-1.878***	-3.761***	-3.76***	-1.255***		-3.067***	-3.066***
	(0.045)	(0.045)	(0.053)	(0.053)	(0.048)	(0.048)	(0.074)	(0.074)	(0.027)		(0.038)	(0.038)
$\vert\text{moneyness}_{it}\vert$	-1.200*		-5.891***	-5.585***	-5.627***	-5.655***	-6.778***	-6.729***	-3.308***		-2.772***	-2.755***
	(0.630)		(0.445)	(0.276)	(0.353)	(0.351)	(0.291)	(0.282)	(0.220)		(0.160)	(0.160)
$\frac{\vert\text{moneyness}_{it}\vert}{ttm_{it}}$	-153.201***	-201.369***	-38.637**	-54.307***	-77.554***	-75.897***	-189.731***	-192.52***	-146.159***		-188.829***	-189.494***
	(25.968)	(11.929)	(17.506)	(4.694)	(20.076)	(20.002)	(15.086)	(14.637)	(15.242)		(10.868)	(10.850)
$\frac{\vert\text{moneyness}_{it}\vert}{ttm^2_{it}}$	469.725**	800.689***	-122.628		403.115**	391.712**	773.821***	790.669***	806.965***		1012.672***	1017.343***
	(196.061)	(119.118)	(127.493)		(156.869)	(157.279)	(118.177)	(116.099)	(111.180)		(83.311)	(83.202)
Wald Test[b]		3.63		1.443		0.294		0.463				2.721
		(0.163)		(0.695)		(0.587)		(0.496)				(0.099)

[a]Regression coefficients and Huber-White standard errors are reported. *** stands for 1 percent, ** for 5 percent, and * for 10 percent significance levels, respectively.
[b]The p-value is reported in brackets below the Wald Test statistic which is distributed $\chi^2(q)$ under the null hypothesis that the model with q restrictions is true.

Table 8: Selection Equations: Pairwise Measure of Clustering for Maturity Classes 1-4

Variable[a]	50- vs 100/200-strike Options: Call Options level[b]	Call M_1	Call M_2	Put Options level[b]	Put M_1	Put M_2	100- vs 200-strike Options: Call Options level[b]	Call M_1	Call M_2	Put Options level[b]	Put M_1	Put M_2
C	0.588*** (0.110)		-0.429** (0.211)	0.871*** (0.165)	-0.916*** (0.119)		0.543*** (0.078)	2.253*** (0.234)	-2.537*** (0.318)	0.266*** (0.075)	0.758*** (0.221)	1.087*** (0.331)
ttm_{it}	0.025*** (0.002)	-0.018*** (0.001)	0.007*** (0.001)	0.02*** (0.003)	-0.013*** (0.002)	0.016*** (0.003)	0.003*** (0.001)	-0.024*** (0.004)	0.037*** (0.005)	0.009*** (0.001)		-0.012*** (0.002)
ttm^2_{it}	-2.03E-4*** (1.48E-5)	1.18E-4*** (9.06E-6)		-2.01E-4*** (1.88E-5)	1.06E-4*** (1.46E-5)	-6.27E-5*** (1.92E-5)		8.32E-5*** (1.63E-5)	-1.43E-4*** (2.39E-5)	-2.64E-5*** (5.39E-6)		3.12E-5*** (1.18E-5)
$AO_{i,t-1}$	0.225*** (0.013)	-0.07*** (0.010)	-0.112*** (0.014)	0.185*** (0.007)		-0.207*** (0.012)	0.31*** (0.025)	-0.157*** (0.021)		0.369*** (0.018)		-0.277*** (0.022)
$ttmitAO_{i,t-1}$							-0.001*** (3.16E-4)	0.001*** (2.90E-4)	-0.001*** (2.84E-4)	-0.002*** (2.43E-4)	0.001*** (1.82E-4)	
$\frac{1}{DAX_{it}}$	6126.993*** (480.917)		-4908.898*** (856.444)	7847.105*** (398.838)		-7192.45*** (789.941)	-2816.072*** (407.378)	1678.523*** (493.047)		-3418.605*** (324.833)	5802.128*** (1051.979)	-10861.49*** (1827.207)
vol_{it}	4.988*** (0.434)	-1.505*** (0.238)	-3.709*** (0.543)	4.138*** (0.516)	-0.979*** (0.307)				-2.87*** (0.319)	1.171*** (0.175)		-1.031** (0.419)
$\frac{\left\lvert m^1_{it}+m^2_{it}\right\rvert}{2}$ [c]	-9.67*** (0.513)	4.927*** (0.373)	6.169*** (0.561)		2.213*** (0.193)		2.177*** (0.308)	-1.567*** (0.293)		0.484*** (0.162)	-2.684*** (0.256)	
I^{md}_{it}	0.253*** (0.028)	-0.082*** (0.028)	-0.102** (0.051)	0.491*** (0.050)	-0.259*** (0.051)	0.493*** (0.101)	0.549*** (0.030)	-0.558*** (0.059)	-0.464*** (0.099)	0.579*** (0.025)	-1.122*** (0.059)	
$I^m_{it}\frac{m^1_{it}+m^2_{it}}{2}$				-3.152*** (0.650)	1.509*** (0.451)	-1.583** (0.777)	1.472*** (0.330)		2.703*** (0.351)	-0.756*** (0.231)		6.259*** (0.538)
$I^{\delta}_{it}\left\lvert ln\left(\delta^1_{it}\right)-ln\left(\delta^2_{it}\right)\right\rvert$ [e]	-4.379*** (0.318)			-5.777*** (0.669)	1.947** (0.873)	-20.743*** (2.251)	-0.568*** (0.080)	3.143*** (0.429)	-2.37*** (0.563)	-0.261** (0.109)		-3.218*** (0.608)
$\left(1-I^{\delta}_{it}\right)\left\lvert ln\left(\delta^1_{it}\right)-ln\left(\delta^2_{it}\right)\right\rvert$	-8.222*** (0.584)	3.307*** (0.558)		-6.812*** (0.412)		-16.992*** (1.653)	-0.626*** (0.088)	2.651*** (0.381)	-5.409*** (0.639)	-0.763*** (0.114)	4.333*** (0.435)	-13.646*** (0.906)
$O_{i,t-1}$ [f]					6.73E-5*** (6.31E-6)	1.17E-4*** (7.83E-6)		1.46E-4*** (1.93E-5)	-8.39E-5*** (1.00E-5)		4.41E-5*** (9.41E-6)	
$ttmO_{i,t-1}$								-9.22E-7*** (1.51E-7)			-2.27E-7*** (1.04E-7)	
$\frac{\lvert m_{it}\rvert}{ttm}$ [g]			-66.7*** 8.716		-90.836*** (14.492)	214.647*** (23.128)		-182.362*** (33.380)	257.365*** (36.306)			-16.566*** (5.876)
$\frac{\lvert m_{it}\rvert}{ttm^2}$		120.226*** (26.156)			701.385*** (0.003)	-1300.699*** (179.663)		1323.59*** (260.547)	-1817.787*** (301.392)			
R^2		0.481			0.447			0.235			0.19	
Wald Test[h]		23.832 (0.161)			18.783 (0.094)			8.937 (0.348)			19.289 (0.154)	

[a]Regression coefficients and Newey-West standard errors are reported. *** stands for 1 percent, ** for 5 percent, and * for 10 percent significance levels, respectively.
[b]level: coefficients on regressors x_i, M_1: coefficients $M^1_{it}\xi^1_{it}$ on regressors b^1_{it}, and M_2: coefficients $M^2_{it}\xi^2_{it}$ on regressors b^2_{it} in equation (21).
[c]The absolute value of the average moneyness of the two neighboring options.
[d]Dummy which takes on value one if the option from the lower strike class (i.e. 50- or 100-strike) has a larger absolute moneyness than its counterpart (i.e. 100/200- or 200-strike) and zero otherwise.
[e]Dummy which takes on value one if the option from the lower strike class (i.e. 50- or 100-strike) has a larger absolute delta than its counterpart (i.e. 100/200- or 200-strike) and zero otherwise.
[f]The option's open interest for the appropriate strike class (50-, 100-, or 200-strike) which was used in the corresponding probit equation.
[g]The option's absolute moneyness for the appropriate strike class (50-, 100-, or 200-strike) which was used in the corresponding probit equation.
[h]The p-value is reported in brackets below the Wald Test statistic which is distributed $\chi^2(q)$ under the null hypothesis that the model with q restrictions is true.

Table 9: Regressions: Pairwise Measure of Clustering for Maturity Classes 1-4

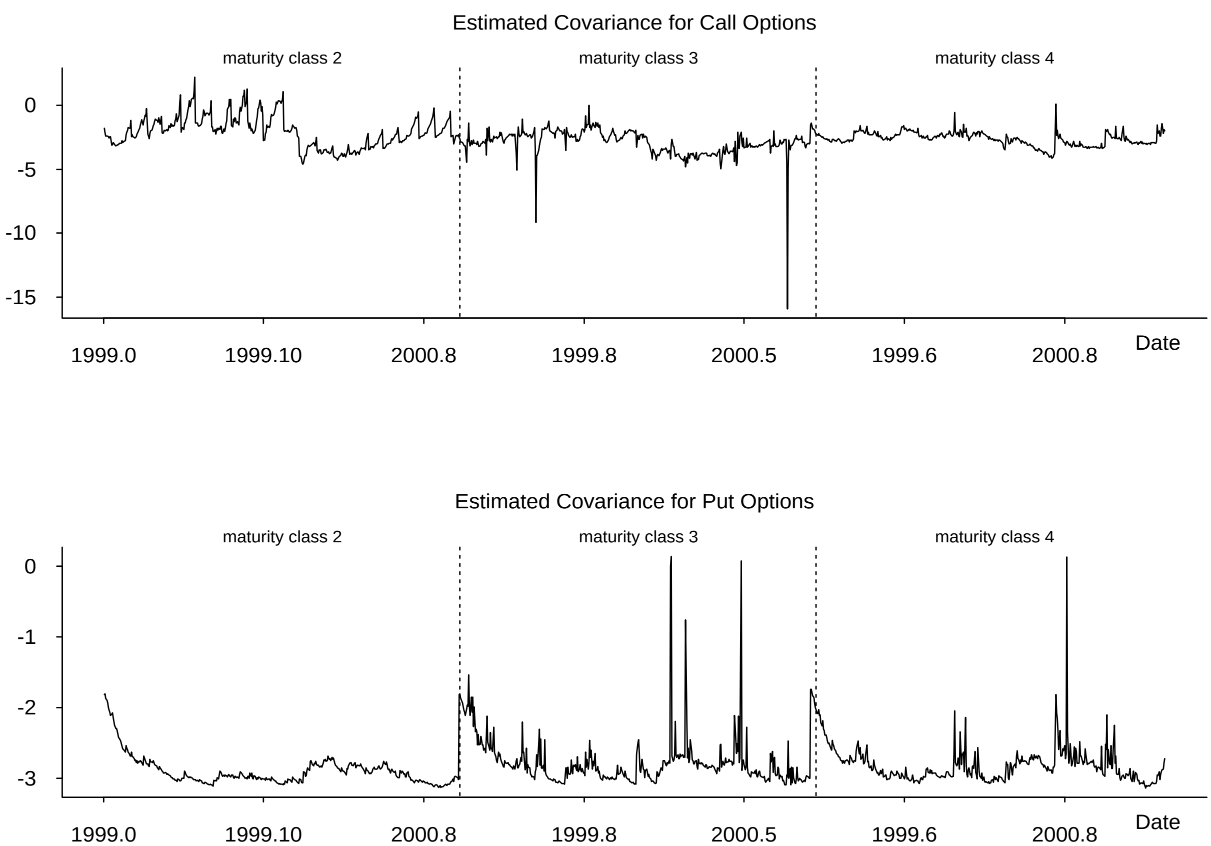

Figure 1: Covariance Between Residuals in the Regression and the Selection Equation