Review of Applied Economics, Vol. 3, No. 1-2, (2007) : 25-48

THE TRADE-OFF BETWEEN LIQUIDITY AND PRECISION OF POSITION IN OPTION CONTRACTS

Alexander K. Koch * **& *Zdravetz Lazarov*** **

More liquid financial contracts are claimed to draw trading volume from contracts for which they are close substitutes. We provide the first analysis of how trading volume across existing financial contracts is affected by changes in the factors that govern the degree to which they are substitutes. Using data on DAX options with different strike prices, we identify these factors and their impact on the distribution of trades across contracts. The results are relevant for exchange design since they help gauge when options with different strike prices are good (bad) substitutes and the strike price grid should be coarse (fine).

***JEL classifications**: G10; G20; L15*

***Keywords:** Clustering; Exchange Design; Options.*

INTRODUCTION

A common contention is that more liquid financial contracts draw trading volume from contracts for which they are close substitutes. Working's (1953) famous example illustrates this point. In 1950, the Chicago Board of Trade introduced a future written on Pacific Northwest soft wheat, the risk of which could only be imperfectly hedged with an existing contract on hard wheat. Despite oering a better hedge for some market participants, the new contract failed because it could not divert enough trading volume from the more liquid hard wheat futures market. Since the two futures contracts were sufficiently close substitutes the liquidity advantage of the existing contract outweighed the hedging advantage that the new contract oered to some market participants. The impact of new financial contracts on trading volume in existing contracts has been studied in several papers (e.g., Silber 1981, and more recently Pennings *et al.* 2001). However, to our knowledge there exists no study of how trading volume across existing financial contracts is aected by changes in the factors that govern the degree to which they are substitutes. The first aim of our paper is to carry out such an empirical analysis. Specifically, we explore the causes of variations in trading volume across German DAX index options contracts with dierent strike prices. The advantage of using options data is that the empirical analysis automatically holds fixed many important variables such as contract regulations, market structure, time to maturity as well as the underlying asset's characteristics.

---

* Royal Holloway, University of London, Department of Economics, Egham, United Kingdom, *E-mail: Alexander.Koch@rhul.ac.uk*

** School of Accounting, Economics and Finance, Edith Cowan University, Joondalup, WA, Australia, *E-mail: z.lazarov@ecu.edu.au*

In DAX index contracts–as in many other option contracts–trading activity follows a saw-toothed pattern. 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. The explanation that we oer for this pattern of trade clustering is that investors face a trade-o between liquidity of their position and the precision with which the contract matches their investment objectives. When options with close strike prices are good substitutes, precision of position matters less and market participants concentrate their trades on a subset of contracts to generate liquidity. This leads to stronger clustering of trading activity: spikes in the saw-toothed volume pattern become more pronounced. In contrast, when options with close strike prices are bad substitutes, traders choose the contract that corresponds best to their needs and trading activity spreads more evenly across strike prices. As a result there is less clustering of trading activity and the saw-toothed volume pattern flattens out.

The second aim of our paper is to shed light on the precise factors that aect the demand for a diverse set of option contracts (many strike prices) rather than a narrow set (few strike prices). Knowledge of these factors is crucial for exchanges which need to design the grid size of option strike prices. 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 exchange. This might lead them to enter customized over-the-counter contracts despite the higher liquidity and transactions costs involved (e.g., Bartram and Fehle 2003). On the other hand, if the grid is too fine, overall volume on the exchange might decrease because trades spread across many strike prices, so that individual contracts witness little trading volume and become illiquid. In sum, if options with dierent strike prices are good substitutes the strike price grid should be coarser than if they are bad substitutes. Exchange regulations reflect such considerations to some extent. One example is that the rules governing the DAX contract create a strike price grid that is coarser for the less liquid long-term contracts than for those with short remaining time to maturity (as explained in more detail below). Another example is the change in regulations for OMX-index contracts implemented by the Swedish options exchange in 1997 and 1998 to obtain a coarser strike price grid (Nordén 2003).

The DAX index contract on Eurex is ideally suited for our investigation because it both is highly liquid and has very precise rules governing the set of available strike prices. Our analysis first identifies factors that potentially aect the degree of substitution between contracts: the options' open interest, the level of the index, time to maturity, the volatility of index returns, options' moneyness, and options' deltas. We then construct a measure of trade clustering which compares the trading activity of neighboring options. This allows us to analyze the trade-o between precision of position on the one hand, and liquidity (working through the factors above) on the other hand. Dierences in trade-clustering are explained as a result of market participants facing a dierent set of factors governing the degree to which contracts are substitutes.

Related Literature

The trade-o between liquidity and other contract characteristics is an important issue in derivatives markets, as illustrated above by Working's (1953) famous example. Nevertheless, only a few papers explicitly address the relation between liquidity and precision of position. 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. Pennings *et al.* (2001) show that, depending on its hedging eectiveness, the introduction of a new futures contract can either lead to an increase in trading volume of existing contracts (reinforcement) or to a decrease (cannibalism). Economides and Siow (1988) consider financial innovation in a model of spatial competition where the division of markets is limited by a liquidity externality. In a similar model, Fehle (2006) analyzes the implications for option exchange design of investors trading o standardization costs and liquidity/transaction costs. Judd and Leisen's (2002) theoretical model addresses the impact of the strike price grid on the demand for options with dierent strike prices, reflected by their open interest. However, their analysis does not include liquidity considerations.

Our paper also relates to the literature analyzing the phenomenon that transaction prices cluster on round price fractions in many financial markets (e.g., Grossman *et al.* 1997, Gwilym *et al.* 1998). Even though price and trade clustering are dierent phenomena, the line of reasoning for the existence of price clustering applies partly to the context of trade clustering. Harris (1991, 1994) suggests that traders aim to simplify negotiations by restricting trading to a discrete price set that is coarser than the price set available. In option markets for the same underlying asset and the same maturity date there typically exist several option contracts which dier only in their strike prices. In the spirit of Harris' negotiation hypothesis, traders might use discrete strike price sets which are coarser than the strike price set introduced by the exchange, facilitating negotiations which are along two dimensions for options, namely prices and strike prices. Ball *et al.* (1985) argue that the amount of information available in the market could determine the market participants' degree of price resolution. In line with this price resolution hypothesis, traders might 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.

The remainder of this paper is organized as follows. A description of the institutional features of the DAX index option market is given in Section 2. Section 3 explains what factors potentially aect clustering and formulates the hypotheses to be tested. Section 4 provides summary statistics of clustering in trading activity. The econometric analysis is contained in Section 5 and Section 6 provides robustness checks for our results. We discuss our findings in Section 7.

MARKET DESCRIPTION

DAX index options are European style and trade on the Eurex, which is an order driven electronic trading system that ranks orders and quotes by their price and time precedence. The underlying is the DAX 30 stock index and options have a contract value of five EUR per index point. On every trading day, the menu of available call and put options comprises eight dierent 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 finish in the succeeding two half-year expiration months (June and December).

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 Böorse 2000a, 2001). Thus, the exchange regulations segment the market for options with identical maturity dates 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 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 100/200-strike class contains all options that are either 100-or 200-strike options.

Overall, 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 B¨orse 2000a, 2001).

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.[1]

FACTORS AECTING CLUSTERING

Option trading activity tends to concentrate near the money and decrease with the distance from the at-the-money point (e.g., Kamara 1995). This decrease in volume, however, is not monotonic but follows a saw-toothed pattern in the case of DAX index options. Trading activity tends to 'cluster' on particular strike classes of options: volume on 200-strikes typically exceeds that on adjacent 100-strikes and volume on 200/100-strikes that on adjacent 50-strikes. We define as clustering an unequal distribution of trading activity in a pair of options with adjacent strike prices ("neighbors") in favor of the 200-strike or 100/200-strike contract, respectively. Our empirical analysis seeks to understand what factors explain dierences in the concentration of trading activity. These are captured by continuous measures of clustering based on the relative volume of trades across contracts in individual option pairs (see Section 5.1) and across contracts in dierent strike price categories on aggregate (see Section 6.2). For example, clustering is said to be stronger in pair 1 of neighboring option contracts than in pair 2 when the 200-strike contract in the former has three times the trading volume of the 100-strike neighbor, and in the latter twice that of the 100-strike counterpart.

Our hypothesis is that the distribution of trading activity across contracts with dierent strike prices (trade clustering) is driven by a trade-o between contract liquidity and the precision with which the contract matches the investor's objectives. When options with close strike prices are good substitutes, precision of position matters less and market participants concentrate their trades on a subset of contracts to generate liquidity. Such coordination can be achieved easily if some strike classes become focal. Option series originally are all 200-strike contracts (time to maturity greater than 12 months; see Section 2). When 100-strike options are introduced they have to compete with existing 200-strike contracts that have already accumulated open interest. This naturally would make the class of 200-strike options appear more attractive to market participants than 100-strike options, facilitating the coordination described above. Similarly, when 50-strike options are introduced they compete with 100/200-strike contracts which have established open interest, making the class of 100/200-strikes appear more attractive than 50-strikes. When two options with neighboring strike prices are close substitutes we argue that trading tends to concentrates on the more ‘attractive’ strike class. The extent of this depends on the following factors which aect 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 between options is (in terms of their strike prices) the higher the degree of substitution between them. Hence, all else equal, clustering should be positively related to the level of the DAX index.

Time to Maturity and Volatility of Index Returns: Many investors in option markets tend to close out their positions near maturity or exercise options since they are directional traders who pursue buy-and-hold strategies or because they implement hedging strategies that are based on holding options until maturity (e.g., see Bartram *et al.* 2004). These traders are interested in the index level at or near maturity. The accuracy with which market participants can predict the final index level decreases with increasing time to maturity and increasing volatility of the index returns. However, if investors’ predictions become less precise, small dierences in strike prices are less important to them.[3] Thus, clustering should increase with volatility and time to maturity.

Options’ deltas: The delta measures the sensitivity of the option’s price to changes in the index level. Market makers usually combine options with dierent 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, clustering should be inversely related to the dierence in options’ deltas.

Options’ moneyness: In option markets trading tends to concentrate around the at-themoney point and volume decreases for options that are farther away from the at-the-money point. Hence, we 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 aects the number of positions that will be closed out. Therefore, clustering should be higher for options that have similar open interest.

This leads to the following hypotheses:

Hypotheses: Clustering of trading activity in neighboring options should be positively (+)/ negatively (-) related to

($H^{+}1$) the level of the DAX index;

(H^+2) the volatility of the DAX index;

(H^+3) the options' time to maturity;

(H^+4) the options' distance from the at-the-money point;

(H^-5) the distance between options' deltas;

(H^-6) the dierence in options' open interest.

DATA AND SUMMARY STATISTICS

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.[4] To allow estimating with sucient precision the implied volatilities needed to obtain options' deltas, we exclude contracts with less than seven days to maturity (see Section 5.3).

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 are misleading because they do not account for the sequential nature of option introduction.[5]

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 deep out-of-themoney 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 contract in favor of the 100-strikes. 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 indeed 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 dierent maturity classes and dierent 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.

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

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			
$m \le -0.2$	0.31	0.78	0.46	0.43	4.24	7.42	6.46	11.58
$-0.2 < m \le -0.15$	0.85	1.13	0.83	0.50	13.61	18.08	12.82	17.89
$-0.15 < m \le -0.1$	2.15	1. 74	0.51	0.39	25.50	30.73	17.10	14.76
$-0.1 < m \le -0.05$	5.3:1	1.90	0.53	0.26	74.16	45.10	10.70	10.22
$-0.05 < m \le 0$	15.60	1.71	0.44	0.28	173.83	45.10	10.33	7. 28
$0 < m \le 0.05$	4.64	0.57	0.20	0.14	67.51	15.22	4.39	3.13
$0.05 < m \le 0.1$	0.30	0.11	0.05	0.07	9.86	3.25	1.91	1.47
$0.1 < m \le 0.15$	0.11	0.09	0.07	0.04	4.32	1.60	1.56	0.82
$0.15 < m \le 0.2$	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	0.4.0	0.49	0.2.3
	100-strike options				200-strike options			
$m \le 0.2$	1.16	1.74	1.24	1.48	1.84	3.01	2.88	4.78
$0.2 < m \le 0.15$	6.74	7.22	4.77	3.86	11.63	12.02	11.17	11.63
$0.15 < m \le 0.1$	17.61	18.86	9.28	5.34	21. 96	27.39	17.53	12.98
$0.1 < m \le 0.05$	60.46	35.78	11.67	5.65	74.71	50.58	17.52	11.87
$0.05 < m \le 0$	155.31	37.37	7.76	4.91	181.07	52.15	12.72	9.62
$0. < m \le 0.05$	61.30	12.49	3.00	1. 79	76.11'	18.:17	5.57	4.33
$0.1 < m \le 0.1$	6.55	2.01	0.65	0.54	10.63	3.54	2.03	1. 76
$0.1 < m \le 0.15$	2.04	0.73	0.37	0.27	4.07	1.51	1. 26	0.86
$0.15 < m \le 0.2$	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			
$m \le -0.2$	0.30	0.55	0.30	0.09	4.32	8.67	6.66	3.27
$-0.2 < m \le -0.15$	0.93	0.73	0.44	0.14	14.15	19.96	11.30	5.64
$-0.15 < m \le -0.1$	2.09	0.90	0.54	0.23	30.45	29.30	14.03	6.06
$-0.1 < m \le -0.05$	4.31	1.20	0.34	0.27	67.53	39.00	14.18	7.60
$-0.05 < m \le 0$	11.76	1.31	0.44	0.28	157.55	41.71	11.20	8.52
$0.1 < m \le 0.1$	5.69	0.70	0.28	0.25	72.20	18.86	6.16	5.30
$0.05 < m \le 0.1$	0.51	0.17	0.09	0.06	13.37	.5.58	2.67	2.64
$0.1 < m \le 0.15$	0.22	0.06	0.04	0.04	739	3.78	2.09	3.00
$0.1 < m \le 0.2$	0.16	0.09	0.03	0.01	6.43	3.24	1.77	1.21
$0.2 < m$	0.09	0.06	0.05	0.03	5.47	2.47	228	7.24
	100-strike options				200-strike options			
$m \le -0.2$	1.89	3.18	2.02	0.94	3.81	6.56	5.32	3.77
$-0.2 < m \le -0.15$	8.73	11.29	5.02	*2.'19*	13.86	20. 32	10.59	7.12
$-0.15 < m \le -0.1$	22.68	21.14	7.40	3.78	33.70	12.11	14.29	8.61
$-0.1 < m \le -0.05$	56.33	29.97	8.60	4.85	80.03	46.89	16.67	10.93
$0.05 < m \le 0$	136.22	13.44	8.16	5.69	173.84	50.37	14.70	11.76
$0 < m \le 0.05$	65.49	16.11	4.43	3.48	79.00	21.49	7.41;	6.94
$0.05 < m \le 0.1$	9.92	3.95	1.58	1.28	13.32	5.76	2.75	2.65
$0.1 < m \le 0.15$	4.19	1.80	1.06	0.66	6.23	3.11	1. 78	1.83
$0.15 < m \le 0.2$	2.65	1.1:1	0.65	0.53	4.70	1.77	1.31	1.46
$0.2 < m$	1.46.46	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

clustering increases with higher maturity classes and with larger absolute moneyness, in line with hypotheses H+3 and H+4. 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 from the raw data.

Table 2
Clustering statistics for Pairs of Options

	100/200 – vs 50-stirke options				*200 – vs 100-strike options*			
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
$0.05 < m \leq 0.1$	3.14	30.76	5.15	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
$0.15 < m < 0,2$	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
$0.15 < m < 0,2$	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,

To look at the eects 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 hypothesis $H^{+}2$. The eect 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 dierent from that for 200- versus 100-strike options. In the former case, clustering tends decrease with increasing index level, while in the latter case, the impact of an increase in the index level is in line with hypothesis $H^{+}1$.

Table 3
Clustering Statistics for Pairs of Options

	100/200- vs 50-strike options Maturity class				*200- vs 100-strike options Maturity class*			
Moneynessa	*1*	*2*	*3*	*4*	*1*	*2*	*3*	*4*
	Call options							
	Ratio of average trading volumes for trading days with volatility above/below the medianb							
m ≤–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 ≤–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 ≤–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 ≤–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 ≤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 ≤ 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 ≤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 ≤ 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 ≤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 medianb							
m ≤–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 ≤–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 ≤–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 ≤ –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 ≤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 ≤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 ≤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 ≤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 ≤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

contd. table 3

Moneyness[a]	100/200- vs 50-strike options, Maturity class 1	2	3	4	200- vs 100-strike options, Maturity class 1	2	3	4
Put options								
Ratio of average trading volumes for trading days with volatility above/below the median[b]								
m ≤–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 ≤–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 ≤–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 ≤–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 ≤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 ≤ 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 ≤ 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 ≤ 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 ≤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 ≤–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 ≤–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 ≤–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 ≤–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 ≤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 ≤ 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 ≤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 ≤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 ≤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.

The descriptive 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 eect of all factors and accounts for various econometric issues.

ANALYSIS OF CLUSTERING IN TRADING ACTIVITY

Pairwise Measure of Clustering

We start by introducing the measure of clustering which will be used in the main empirical analysis. Clustering of 100/200- versus 50-strikes is considered separately from that of 200- versus 100-strike options. The measure of clustering between the ith 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. Formally,

$$PC_{i,t}^{200/100} = \begin{cases} In\left(\frac{T_{i,t}^{200}}{T_{i,t}^{100}}\right) & if\ T_{i,t}^{200} \cdot T_{i,t}^{200} > 0 \\ not\ defined & otherwise, \end{cases} \tag{1}$$

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

$\mathcal{N}_t^{200-100}$: index set of neighboring 100- and 200-strike options on day t,

$T_{i,t}^{100}$: transaction volume on the 100-strike,

$T_{i,t}^{200}$: transaction volume on the 200-strike.

The measures for 100/200- versus 50-strikes are defined analogously. Note that the empirical analysis includes also all newly introduced options: such contracts simply give rise to new pairs of neighboring contracts. Table 4 reports summary statistics for the pairwise measure of clustering. The next sections describe our econometric procedure and present the estimation results.

ECONOMETRIC ISSUES

Even though the DAX contract is highly liquid compared to other options, many contracts are not traded on a given day (in particular away from the money). For this reason, any clustering measure will have censoring or truncation of the data. To see this, consider any measure of clustering based on a function $f(T^+, T^-)$ which depends on the trading activity in two neighboring options belonging to dierent 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^-. The 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. $f(0, 0) = (a + b)/2$. However, one needs to account for the resulting

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_t	2.57	2.51	1.63	1.40	2.64	2.54	1.56	1.27
Maximum PC_t	6.43	5.75	4.75	4.97	6.45	5.46	5.28	5.18
Minimum PC_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
# $T_t^{50} = 0$	16,747	18,709	17,518	19,517	15,736	19,003	17,728	19,513
# $T_t^{100/200}$	5,175	5,776	6,190	7,686	3,404	4,032	4,755	5,411
$T_t^{50} = T_t^{100/200} = 0$	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_t	0.30	0.41	0.51	0.69	0.33	0.47	0.53	0.64
Maximum PC_t	5.00	4.61	4.23	4.40	4.64	4.44	4.61	4.44
Minimum PC_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
# $T_t^{100} = 0$	10,235	9,946	9,346	14,784	7,236	6,983	7,527	13,511
# $T_t^{200} = 0$	7,944	8,066	7,378	10,200	5,351	5,441	5,406	8,855
# $T_t^{100} = T_t^{200} = 0$	6,808	7,094	6,297	8,874	4,265	4,453	4,208	7,285

mass points on a, b, and $(a + b)/2$ when regressing this measure on the factors that explain clustering.

We deal with these issues in a simple and robust manner by using a Heckman (1979) two-step estimation procedure. Our second-stage regression uses the pairwise measure of clustering, which is defined only when both options have non-zero trading volume, thus omitting the critical cases mentioned above. Exploiting the full sample of data we run first-stage probit regressions based on whether an option is traded (i.e., is included in the second-stage regression sample) or not. From these specifications we obtain a sample selection correction for the second-stage regression that controls for the potential bias introduced by omitting the zero-trading volume cases. The procedure is described in detail in Appendix A. In Section 6 we ascertain the robustness of our empirical results to the choice of measure and methodology by using a more coarse clustering measure based on aggregate trading volume across strike classes, which does not suer from truncation.

Regression Results

Results for the second-stage regressions are reported in Table 5. 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.

Moreover, significance levels are based on Newey and West (1987) heteroscedasticity and autocorrelation consistent standard errors.

Table 5
Second-Stage Regressions (Restricted Specifications)

Dependent Variable: PCt, 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	1.43***	1.15***	0.30***	0.15***
	(0.10)	(0.13)	(0.05)	(0.05)
PC_{t-1}	0.31***	0.34***	0.45***	0.34***
	(0.01)	(0.01)	(0.01)	(0.01)
ttm_t	0.03***	0.03***	$1.38E-3$***	$2.99E-3$***
	(4.12E-3)	(4.50E-3)	(4.35E-4)	(4.07E-4)
ttm^2_t	$-2.31E-4$***	$-1.89E-4$***		
	(5.49E-5)	(4.16E-5)		
PO_{t-1}	0.08***		0.13***	0.24***
	(0.01)		(0.01)	(0.02)
$ttm_t\, PO_{t-1}$		$8.46E-4$***	$-1.31E-3$***	$-1.43E-3$***
		(2.80E-4)	(2.45E-4)	(2.01E-4)
$\frac{1}{DAX_t}$		2,206.08***	–1,787.84***	–1,249.25***
		(415.90)	(249.71)	(206.20)
vol_t	1.74***	1.55***	0.68***	0.71***
	(0.34)	(0.38)	(0.10)	(0.09)
$D_t\, avm_t$[b]	–10.88***	–16.57***	–1.45***	–0.49**
	(2.33)	(1.70)	(0.47)	(0.21)
$(1-Dt)avm_t$	–5.35***		1.67***	0.46***
	(0.79)		(0.28)	(0.14)
$D_t\, avm^2_t$	152.11***	117.53***		
	(52.77)	(12.05)		
$(1-Dt)\, avm^2_t$		–20.63***		
		(3.28)		
$D_t\, avm_t\, ttm_t$		0.12***		
		(0.03)		
$(1-D_t)\, avm_t\, ttm_t$			–4.52***	
			(1.01)	
$2\frac{\lvert\delta_t^k-\delta_t^l\rvert_c}{\lvert\delta_t^k+\delta_t^l\rvert}$	–2.67***	–1.46***	–0.22***	–0.20***
	(0.78)	(0.54)	(0.07)	(0.08)
correction terms[d]				
R^2	0.35	0.36	0.27	0.19
Wald test[e]	32.70	30.85	30.27	28.55
	(0.11)	(0.08)	(0.09)	(0.05)
N	11,053	10,417	27,922	34,473

[a]Newey-West (1987) standard errors in parentheses; ***, **, and * indicate significance at 1 percent, 5 per cent level, and 10 per cent level, respectively.

[b]D_t: Dummy which takes on value one if the pair of options is in the money on day t. avmt: 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] $\delta_t^k\ (\delta_t^l)$: Delta of the option in the higher (lower) strike class $k \neq l$, $k, l \in \{50, 100/200, 100, 200\}$ on day t.

[d]Coefficients for the regressors interacted with the inverse Mills ratios computed from the corresponding first-stage estimations are not reported.

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

Evidence regarding hypothesis H+1 is mixed: the coecient on the inverse of the DAX index level $\left(\frac{1}{DAX_t}\right)$ has the correct sign for 200- versus 100-strike calls and puts; but for 100/200- versus 50-strikes it is either insignificant (calls) or negative (puts). The daily GARCH(1,1) DAX index volatility (volt) serves as a proxy for the impact of volatility on clustering. Its coecient is always positive and significant, in line with hypothesis H^+2. Similarly, there is clear support for hypothesis H^+3, the coecient on time to maturity ($ttm_{i,t}$) is always positive and significant, the eect being non-linear for 100/200- versus 50-strikes. Testing hypothesis H^+4, we allow for an asymmetric influence of moneyness on clustering depending on whether the option is in or out of the money. An option pair is labelled in the money (dummy $D_{i,t} = 1$) if the average moneyness[6] of the two options is positive. To account for possible nonlinear eects, we also include $avm^2_{i,t}$ and $avm_{i,t}$ interacted with time to maturity. The results are ambiguous. For 100/200- versus 50-strike options clustering decreases with higher average moneyness, contrary to H^+4. For the 200- versus 100-strike options the eect of moneyness on clustering in out-of-themoney options is positive, while for in-the-money options the eect is reversed. Hypothesis H^-5 on the eect of neighboring options' deltas on clustering is clearly supported because the coecients on the corresponding regressor is always significant and negative. As a regressor we use the absolute value of the arithmetic dierence between options' deltas, scaled by the average deltas of the two options to reduce potential moneyness eects on the measure. Option deltas are computed based on the Black and Scholes (1973) formula, using as input the implied volatility obtained by the procedure described in Appendix B. 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. To test hypothesis H^-6 we include the previous trading day's value of the measure of clustering for open interest, $PO_{i,t-1}$, and an interaction with $ttm_{i,t}$. Overall, the impact of $PO_{i,t-1}$ is significant and positive, as predicted. For 200- versus 100-strikes we find that the eect of open interest diminishes with shorter maturities.

ROBUSTNESS CHECKS

Potential Confounding "History" Eects

Clustering may be driven by past movements of the DAX index. For example, consider the part of options' life cycle during which only 200-strikes exist. If the index value changes significantly during that period then trading spreads across dierent 200-strike contracts, each of them being near the money only for a short period of time. In this case, 200-strikes do not have a large built-up liquidity advantage over 100- and 50-strike contracts at the time when these are introduced. Thus, all else equal, clustering should be low. In contrast, if the index does not move much in the period when only 200-strikes exist, then some of the older 200-strike options accumulate a large liquidity advantage over 100- and 50-strike options introduced later. Thus, all else equal, clustering should be high. In our regressions the lagged value of the pairwise measure and open interest control for this eect to some degree. However, to look at this in more detail, we re-estimate the restricted specifications of the regressions in Table 5 using only pairs of options which are simultaneously introduced so that no option has an initial liquidity advantage. As the results in Table 6 show, the regression coecients do not change signs.

Table 6
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-4)
ttm^2_t	$-2.73E-4$***	$-2.47E-4$***		
	(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-4)
$\frac{1}{DAX_t}$		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)
$D_t\, avm^2_t$	117.20**	89.13***		
	(57.55)	(25.77)		
$(1-D_t)\, avm^2_t$		–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)
correction terms[d]				
R^2	0.35	0.35	0.24	0.16
N	5,292	4,779	19,802	23,085

[a]Newey-West (1987) standard errors in parentheses; ***, **, and * indicate significance at 1 per cent, 5 per cent level, and 10 per cent level, respectively.
[b]D_t: Dummy which takes on value one if the pair of options is in the money on day t. avmt: 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] $\delta_t^k\ (\delta_t^l)$: Delta of the option in the higher (lower) strike class $k \neq l$, $k, l \in \{50, 100/200, 100, 200\}$ on day t.
[d]Coefficients for the regressors interacted with the inverse Mills ratios computed from the corresponding first-stage estimations are not reported.

This also allows us to control for another possible “history” eect. 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. Since this may lead clustering to fall with decreasing time to maturity, the interpretation of a negative coecient on time to maturity as supporting hypothesis $H^{+}3$ could be false. The regressions reported in Table 6 rule out this history-related liquidity advantage eect because it only looks at options that have the same age. The coefficients on time to maturity are roughly the same as in the earlier regressions, which clearly supports hypothesis $H^{+}3$.

Potential Confounds from the Estimation Methodology

As a control of our Heckman type estimation procedure we carry out an analysis using an alternative measure of clustering that is amenable to standard regression analysis. Both procedures should lead to similar qualitative results unless there are confounding effects introduced by the two-stage estimation strategy employed above. The alternative measure of clustering is based on daily aggregated trading volumes and has the advantage of not suering from truncation for the first maturity class.

The measure for the case of 100- versus 200-strike options is derived as follows. 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 eects of moneyness and dierences in deltas can not be isolated. The estimation results corroborate the evidence from our two-stage estimation procedure used earlier. The coecients reported in Table 7 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 coecients exhibit the same behavior for both the pairwise and the aggregate measures.

DISCUSSION AND CONCLUSION

Our empirical results provide robust evidence that trading patterns in DAX index options are consistent with the common contention that more liquid financial contracts draw trading volume from contracts for which they are close substitutes. In contrast to extant studies which consider the eect that introducing a new security has on other contracts (e.g., Pennings *et al.* 2001), our option data allows us to study how trading activity across existing financial contracts changes with variations in the factors that govern the degree to which they are substitutes. Moreover, focusing on options series which differ only in their strike price eliminates many potential confounding factors which arise in other settings, as characteristics such as market structure, time to maturity as well as the underlying asset’s characteristics do not vary across contracts.

Table 7

Regressions for Aggregate Measure of Clustering (Maturity Class 1, Restricted Specifications)

Dependent Variable: ACt, 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^2_t	$-1.81E-3$***	$-1.37E-3$***		
	$(2.44E-4)$	$(3.04E-4)$		
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 (1987) standard errors in parentheses; ***, ** and * indicate significance at 1 per cent, 5 per cent, and 10 per cent 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.

We confirm that clustering of trading activity is positively related to to the volatility of the DAX index and the options' time to maturity and inversely related to the distance between options' deltas as well as the dierence in options' open interest. In contrast, the effect of moneyness and the index level on clustering is ambiguous.

Overall, these eects can be quite strong, producing a pronounced saw-toothed pattern in option trading volume that has not been documented previously. For example, the ratio of trading volume on 100/200-strikes relative to 50-strikes which are close to the at-the-money point increases with time to maturity from around 12 to 20-40 and roughly doubles when comparing low to high volatility days.

From an efficiency perspective, product standardization and concentration of trading activity on one trading platform tend to be beneficial because of the liquidity externality trades create and lower transaction costs. However, option exchanges face competition from other trading venues and banks oering over-the-counter products. Therefore, it is crucial to design the characteristics of contracts to exploit liquidity externalities arising from product standardization while providing sucient product variety to limit order flow to competing contracts or over-the-counter markets. Our analysis provides a tool for identifying the crucial factors to consider in drawing up regulations that govern the strike price grid for option contracts. To promote the liquidity of its contracts an exchange should strive to implement a coarse grid size in situations in which options with close strike prices tend to be good substitutes. Otherwise, a finer strike price gradation is required to prevent traders from moving to customized over-the-counter products, despite higher liquidity and transactions costs that the latter involve.

APPENDIX

Estimation Procedure

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

We assume there exist three latent variables, $u_{i,t}^{50}, u_{i,t}^{100}$, and $u_{i,t}^{200}$, 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, . . . , T, and $i \in \mathcal{K}_t^k \subset IN$,

$$u_{i,t}^{k} \begin{cases} > 0 & if \;\; T_{i,t}^{k} \cdot T_{i,t-1}^{k} > 0, \\ \leq 0 & otherwise, \end{cases} \tag{2}$$

where $\mathcal{K}_t^k \subset IN$ is the index 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, . . . , T and $i \in \mathcal{N}_t^{200-100} \subset IN$,

$$\left(y_{i,t}, y_{i,t}^{*}\right) \begin{cases} = \left(PC_{i,t}^{200/100},\ PC_{i,t-1}^{200/100}\right) & if \;\; u_{q(i),t}^{100} \cdot u_{l(i),t}^{200} > 0 \\ \quad not\,observed & \quad otherwise, \end{cases} \tag{3}$$

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

The latent variables are assumed to depend as follows on three sets of regressors, $z_{i,t}^{100}, z_{i,t}^{200}$, and *xi,t*:

$$u^{100}_{q(i),t} = \left(z^{100}_{q(i),t}\right)' \gamma^{100} + \epsilon^{100}_{q(i),t} \tag{4}$$

$$u^{200}_{l(i),t} = \left(z^{200}_{l(i),t}\right)' \gamma^{200} + \epsilon^{200}_{l(i),t}, \quad i \in \mathcal{N}^{200-100}_t, \quad t = 2,\ldots.T, \tag{5}$$

$$y_{i,t} = y^{*}_{i,t}\beta^{*} + x'_{i,t}\beta + \nu_{i,t}. \tag{6}$$

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: $\text{var}\left(\epsilon^{100}_{q(i),t,}\right) \equiv \left(\epsilon^{200}_{l(i),t,}\right) \equiv 1.$

(iii) The conditional covariances between the residuals in Eqs. (4) and (6) and between the residuals in Eqs. (5) and (6) 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:

$$Cov\left[\epsilon^{k}_{j,t}, \nu_{i,t} \mid y^{*}_{i,t}, x_{i,t,} z^{200}_{q(i),t}, z^{200}_{l(i),t,}\ (y_{i,t,} y^{*}_{i,t})\ \text{observed}\right] = \left(b^{k}_{i,t}\right)' \xi^{k}, \tag{7}$$

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

With the above assumption we obtain as a straightforward extension of Heckman (1979)

$$PC^{200/100}_{i,t} = PC^{200/100}_{i,t-1}\beta^{*} + x'_{i,t}\beta + M^{100}_{i,t}\left(b^{100}_{i,t}\right)'\xi^{100} + M^{200}_{i,t}\left(b^{200}_{i,t}\right)'\xi^{200} + \omega_{i,t}, \tag{8}$$

where

$$\omega_{i,t} = \nu_{i,t} - E\left[\nu_{i,t} \mid u^{100}_{q(i),t} > 0,\ u^{200}_{l(i),t} > 0\right], \quad i \in \mathcal{N}^{200-100}_t, \tag{9}$$

$$M^{k}_{i,t} = \frac{\phi\left(\left(z^{k}_{j,t}\right)'\gamma^{k}\right)}{\Phi\left(\left(z^{k}_{j,t}\right)'\gamma^{k}\right)}, \ k \in \{100, 200\}, j \in \{q(i), l(i)\}, t = 2, \ldots, T, \tag{10}$$

where $\phi()$ and $\Phi()$ are the pdf and cdf of the standard normal distribution, respectively. Our estimation procedure assumes strict exogeneity of the regressors in (8).

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.[7] In the first-stage 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.8 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.

Table 8 reports summary statistics for the individual options in the dierent strike classes. Results from the probit estimation are summarized in Table 9. The dependent variable in the probit estimation is an indicator variable $I^k_{i,tt-1}$ that takes on the value one whenever the *i*th 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 set of regressors $z_{i,t}$ comprises time to maturity ($ttm_{i,t-1}$) and its square because the liquidity of options decreases with increasing time to maturity. 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 dummy variables for being in the money ($D_{i,t-1} = 1$) or not. Additionally, interaction effects between *ttm* and moneyness are included as well as open interest ($O_{i,t-2}$), its interaction with *ttm*, and GARCH(1,1) index volatility ($vol_{i,t-1}$). To control for possible autocorrelation in the residuals of the probit equation we also include the lagged dependent variable, $I^k_{i,t-2,t-3}$.

From these results we obtain the terms $M^k_{i,t}$ to include in the second-stage regressions to correct for potential selection bias. The set of regressors $b^k_{i,t}$ in the covariance terms $\left(b^k_{i,t}\right)' \xi^k$ includes all of the regressors in $x_{i,t}$ and most of those in $z^k_{.,t}$ 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 $b^k_{i,t}$ and not the corresponding variables from the probit equations.

Table 8
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 # trades> 0	12,011	21,821	34,034	12,278	26,221	40,321
Average # trades per day	1.73	13.24	12	1.51	13.63	13.55
Maximum # trades per day	422	577	892	302	623	905
Minimum # trades per day	0	0	0	0	0	0
Standard deviation	8.2	40.06	38.49	6.84	36.13	37.9

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.[9] That is,

Table 9
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***	–4.85*E*–3***	–0.05***	–0.01***	–0.01***
	(2.41*E*–3)	(5.53*E*–4)	(4.52*E*–4)	(2.79*E*–3)	(4.62E–4)	(5.41*E*–4)
ttm^2_{t-1}	1.90*E*–4***			2.02*E*–4***		
	(1.42*E*–5)			(1.57*E*–5)		
O_{t-2}[c]	1.50*E*–4***	6.49*E*–5***	2.87*E*–5***	4.64*E*–5**	1.05*E*–4***	5.16*E*–5***
	(1.34*E*–5)	(3.13*E*–6)	(1.27*E*–6)	(1.89*E*–5)	(6.93*E*–6)	(3.07*E*–6)
$ttm_{t-1}\ O_{t-2}$				1.35*E* – 6***	–3.45*E*–7***	–1.04*E*–7***
				(3.78*E*–7)	(7.38*E*–8)	(3.16*E*–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)		(3.21*E*–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)	(4.39*E*–3)
$D_{t-1}\ \lvert m_{t-1}\rvert\ ttm^2_{t-1}$	–3.53*E*–3***	5.05*E*–4***	3.33*E*–4***	–1.74*E*–3***		
	(7.30*E*–4)	(1.78*E*–4)	(6.84*E*–5)	(6.67*E*–4)		
$(1–D_{t-1})\ \lvert m_{t-1}\rvert\ ttm^2_{t-1}$	–1.04E–3***	–1.98*E*–4***	–1.97*E*–4***	–1.04*E*–3***	–2.76*E*–4***	–2.05*E*–4***
	(2.01*E*–4)	(4.64*E*–5)	(2.90*E*–5)	(1.99*E*–4)	(4.61*E*–5)	(2.21*E*–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 per cent, 5 per cent, and 10 per cent 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. $|m_{t-1}|$: (absolute) moneyness on day $t-1$, defined as $|strike/DAX_{t-1} - 1|$

[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.

$$F_{t,v}(T_F) = S_{t,v} e^{r(T_F - t)}, \tag{11}$$

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 (11) 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.[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 andWallmeier (2001), we approximate the smile on every trading day by fitting via OLS a smooth dierentiable spline function whose segments join at the at the money point. This allows for quadratic function segments for the in- and out-of-the-money ranges, respectively. Letting *M* = *strike/DAX*, we estimate

$$\sigma_{IV} = \alpha_0 + \alpha_1 M + \alpha_2 M^2 + D(\beta_0 + \beta_1 M + \beta_2 M^2) + \epsilon, \tag{12}$$

$$D = \begin{cases} 0 & \text{if} \quad \text{if strike below the at} - \text{the} - \text{money point}, \\ 1 & \text{if} \quad \text{if strike above the at} - \text{the} - \text{money point}. \end{cases}$$

ACKNOWLEDGEMENT:

We are grateful for helpful comments on this project from Bruno Biais, 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 and thank the Deutsche Börse AG for providing the data. An earlier version of the paper was entitled *Clustering of trading activity in the DAX index options market*.

NOTES

1. For an analysis of the link between prices in DAX futures and option markets see Schlag and Stoll (2005).
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 *et al.* (1985).
4. These are also the most liquid contracts. As in other derivatives markets, both trading volume and open interest decrease for longer maturity contracts (e.g., Bessembinder and Seguin 1992, Figlewski 1983).
5. It is a well know feature of options markets that liquidity (trading volume and open interest) is concentrated on the contracts nearest to expiry and our data are no exception, as will be discussed below.
6. Moneyness is defined as $m = 1 - \frac{strike}{DAX}$ for call options and as $m = \frac{strike}{DAX} - 1$ for put options.

7. 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.
8. Using the odd days only yields virtually the same results.
9. 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 diers 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 minor, however.
10. The interest rate convention for Libor rates is linear and, therefore, rates have to be converted to continuous compounding first.

REFERENCES

Ball, C. A., Torous, W. N. and Tschoegl, A. E. (1985), "The Degree of Price Resolution: The Case of the Gold Market", *Journal of Futures Markets* 5: 29–43.

Bartram, S. M., Brown, G. W. and Fehle, F. R. (2004), "International Evidence on Financial Derivatives Usage", Mimeo.

Bartram, S. M. and Fehle, F. R. (2003), "Alternative Market Structures for Derivatives", Mimeo.

Bessembinder, H. and Seguin, P. J. (1992), "Futures-Trading Activity and Stock Price Volatility", *Journal of Finance* 47: 2015–2034.

Black, F. and Scholes, M. (1973), "The Pricing of Options and Corporate Liabilities", *Journal of Political Economy* 81: 637–59.

Cuny, C. 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.

Deutsche Börse (2000b), German Stock Index Dax, Frankfurt am Main.

Deutsche Börse (2001), Conditions for Trading at Eurex Deutschland and Eurex Zürich, Frankfurt am Main.

Economides, N. and Siow, A. (1988), "The Division of Markets is Limited by the Extent of Liquidity (Spatial Competition with Externalities)", *American Economic Review* 78: 108–121.

Fehle, F. (2006), "Too Many Options? Theory and Evidence on Option Exchange Design", *Journal of Futures Markets* 26: 533–570.

Figlewski, S. (1983), "Hedging Performance and Basis Risk in Stock Index Futures", *Journal of Finance* 39: 657–669.

Grossman, S. J., Miller, M. H., Cone, K. R. and Fischel, D. R. (1997), "Clustering and Competition in Asset Markets", *Journal of Law and Economics* 40: 23–60.

Gwilym, O. A., Clare, A. and Thomas, S. (1998), "Extreme Price Clustering in the London Equity Index Futures and Options Markets", *Journal of Banking and Finance* 22: 1193–1206.

Hafner, R. and Wallmeier, M. (2001), "The Dynamics of DAX Implied Volatilities", *International Quarterly Journal of Finance* 1: 1–27.

Harris, L. (1991), "Stock Price Clustering and Discreteness", *Review of Financial Studies* 4: 389–415.

Harris, L. E. (1994), “Minimum Price Variations, Discrete Bid-Ask Spreads, and Quotation Sizes”, *Review of Financial Studies* 7: 149–78.

Heckman, J. J. (1979), “Sample Selection Bias as a Specification Error”, *Econometrica* 47: 153–61.

Judd, K. L. and Leisen, D. P. J. (2002), “Equilibrium Open Interest”, Mimeo.

Kamara, A. (1995), “Daily and Intradaily Tests of European Put-Call Parity”, *Journal of Financial & Quantitative Analysis* 30: 519–539.

Newey, W. and West, K. (1987), “A Simple Positive Semi-Definite, Heteroscedasticity and Autocorrelation Consistent Covariance Matrix”, *Econometrica* 55: 703–8.

Nordén, L. (2003), “Option Exchange Design: Concentration of Trading and Open Interest at the Swedish Index Options Market”, Mimeo.

Pennings, J. M. E. and Leuthold, R. M. (2001), “Introducing New Futures Contracts: Reinforcement versus Cannibalism”, *Journal of International Money and Finance* 20: 659–675.

Schlag, C. and Stoll, H. (2005), “Price Impacts of Option Volume”, *Journal of Financial Markets* 8: 69–87.

Silber, W. L. (1981), “Innovation, Competition, and New Contract Design in Futures Markets”, *Journal of Futures Markets* 1: 123–155.

Working, H. (1953), “Futures Trading and Hedging”, *American Economic Review* 43: 314–343.