Evaluating VaR Models with Desk-Level Data

Jeremy Berkowitz (Houston)
Peter Christoffersen (McGill)
Denis Pelletier (NC State)

www.christoffersen.ca

Research Questions

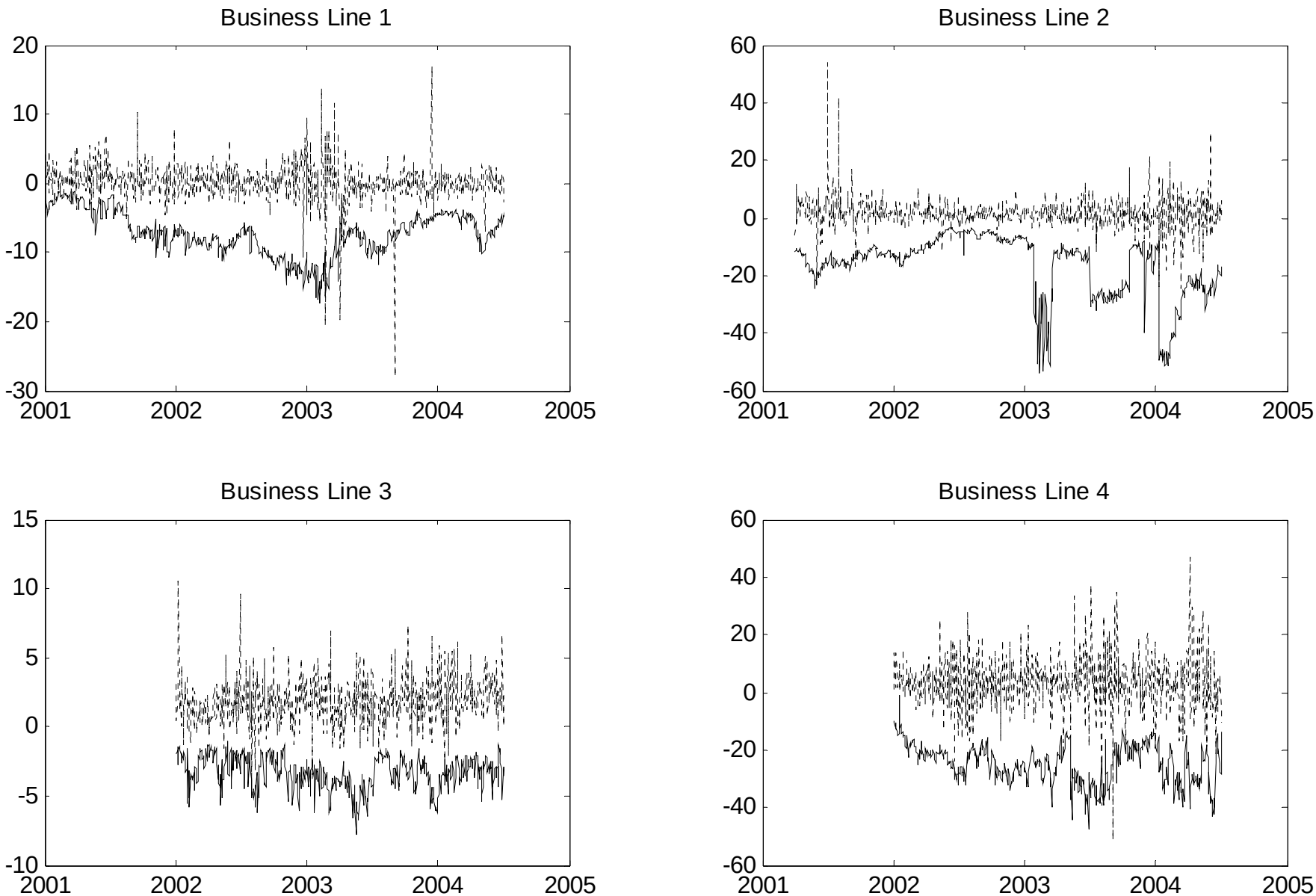
- What do actual daily desk-level P/Ls and VaRs look like?
- The model-free Historical Simulation (HistSim) methodology seem to be gaining popularity in banks. Is this a positive development?

  have any chance of rejecting HistSim VaRs in realistic sample sizes?
- If so which test should we use?
- Can we rely on the asymptotic critical values of these tests?

Selected Literature

- Kupiec (JD, 1995)
- Christoffersen (IER, 1998)
- Berkowitz (JEC, 2001)
- Berkowitz and O’Brien (JF, 2002)
- Engle and Manganelli (JBES, 2004)
- Christoffersen and Pelletier (JFEC, 2004)

Desk-Level P/L and VaR



P/L and VaR: Descriptive Statistics

P/Ls

	Desk 1	Desk 2	Desk 3	Desk 4
Number of Observations	873	811	623	623
Mean	0.1922	1.5578	1.8740	3.1562
Standard Deviation	2.6777	5.2536	1.6706	9.2443
Skewness	-1.7118	1.5441	0.5091	-0.1456
Excess Kurtosis	24.2195	19.8604	2.0060	3.6882

VaRs

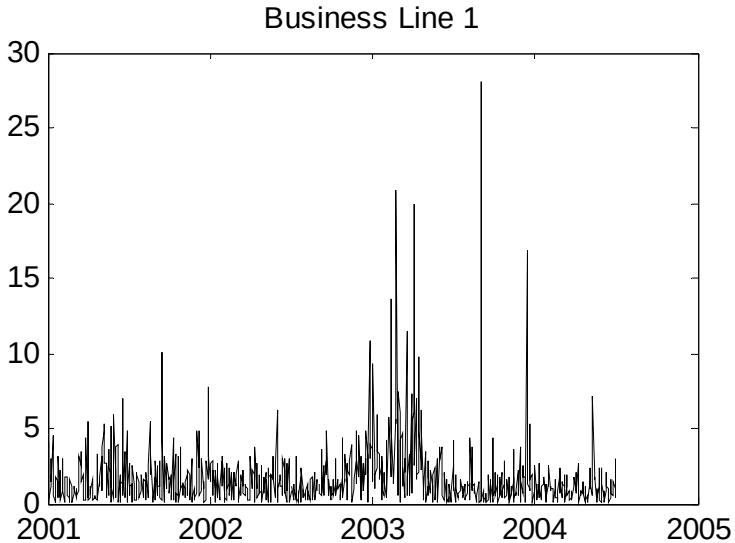
	Desk 1	Desk 2	Desk 3	Desk 4
Number of Observations	873	811	623	623
Mean	-7.2822	-16.3449	-3.2922	-24.8487
Standard Deviation	3.1321	10.5446	1.1901	6.6729
Skewness	-0.3038	-1.3746	-0.6529	-0.3006
Kurtosis	-0.1525	1.6714	-0.0133	-0.1211
Observed Number of Hits	9	5	1	4
Expected Number of Hits	9	8	6	6

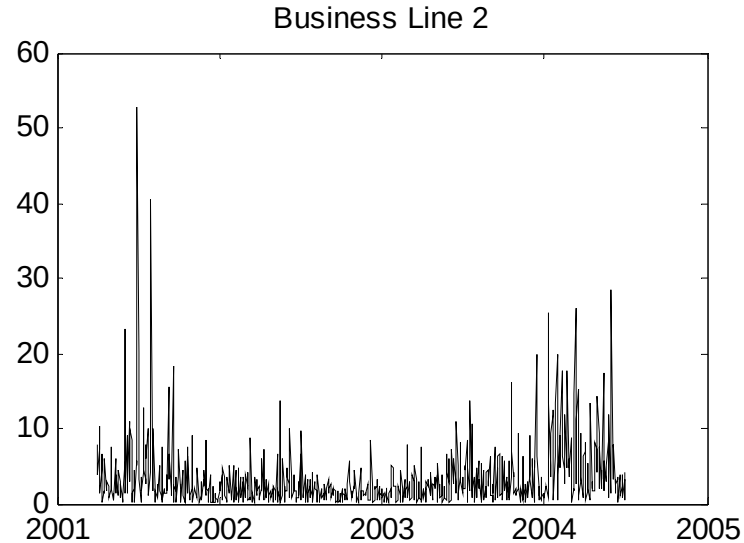
Key Features

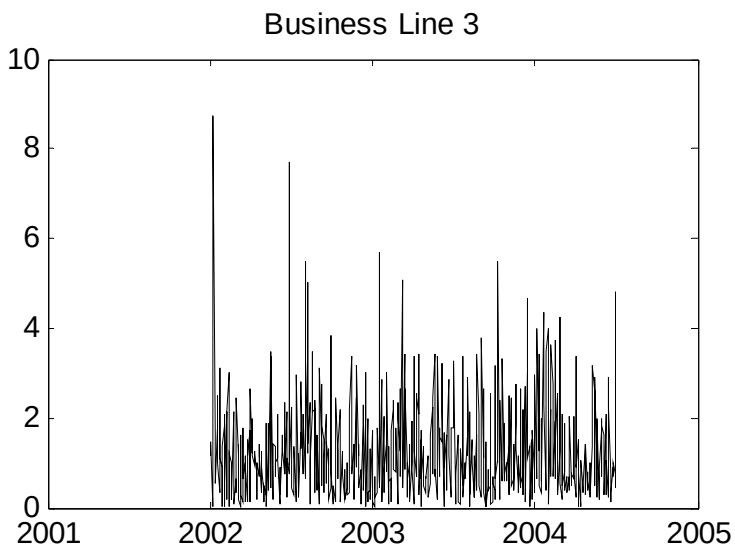
- Non-normality
- Sometimes erratic VaR behavior
- Sometimes dramatic violations
- Volatility dynamics?

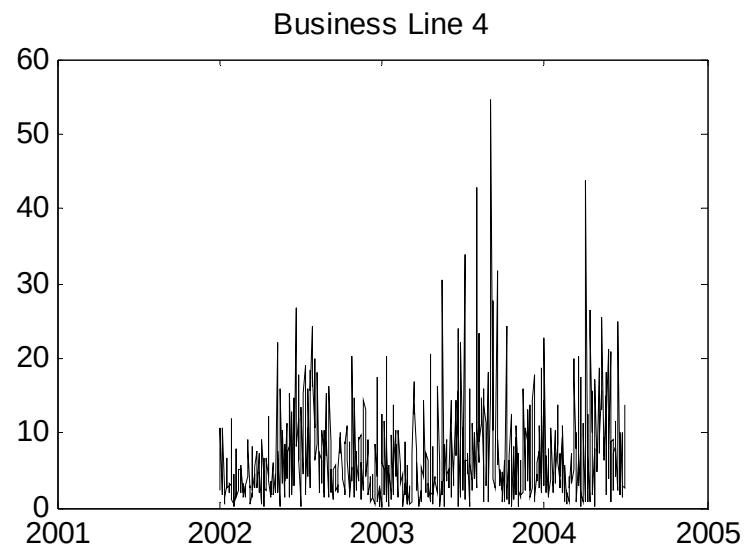
- Consider a plot of the demeaned absolute P/Ls.

Demeaned Absolute P/Ls









General Backtesting Hypothesis

- The VaR must satisfy

$$F\left(R_{t+1} < -VaR_t \mid \Omega_t\right) = p$$

- By definition the VaR only allows us to test the conditional probability of a violation and not for example the magnitude of violations.
- Conditional testing is crucial.

Types of Tests

- We will test various features of the hit sequence defined as

$$I_{t+1} = \begin{cases} 1, & if \;\; R_{t+1} < -VaR_t(p) \\ 0, & else \end{cases}$$

- Conditional moment tests
- Duration tests
- Spectral density tests

Conditional Moment Tests

- Test the hypothesis

$$E\left[(I_{t+1} - p) \otimes Z_t\right] = 0$$

- Conditional information:
- Markov: $Z_t = I_t$
- LB(1): $Z_t = I_t$
- LB(5): $Z_t = \{I_t, I_{t-1}, \ldots, I_{t-4}\}$
- Caviar: $Z_t = \{I_t, VaR_t\}$

Duration Based Tests

- Define the duration between two violations

$$D_i = t_i - t_{i-1}$$

- Null hypothesis: No duration dependence, that is, a flat hazard equal to *p*. The probability of getting a hit tomorrow does not depend on the time since the last hit.
- Alternative hypothesis: The Weibull (continuous) and the Geometric (discrete). Allow for duration dependence.
- Intuition: Omitted volatility dynamics will result in clustered violations, that is, a higher probability of a hit soon if a hit happened recently and vice versa.

Spectral Based Tests

- The spectral density function is defined by

$$f(\omega) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \gamma_k e^{-ik\omega}$$

- For the VaR hit sequence we should have

$$f(\omega) = \frac{1}{2\pi} p(1-p)$$

- Test distance between the two using Cramer von Mises (CVM) and Kolmogorov-Smirnov (KS) statistics.

Size of Asymptotic Tests

- We first want to assess the ability of the backtester to rely on the asymptotic normal property of the tests considered.
- We do a Monte Carlo generating i.i.d. hit sequences with the correct hit rate, p.
- Consider various sample sizes and p=1% and 5% respectively.
- Nominal size is set to 10%.

Table 2: Size of 10% Asymptotic CC Tests

1 % VaR

Sample	LB(1)	LB(5)	Markov	Weibull	Geometric	Caviar	KS	CVM
250	0.0253	0.0999	0.0497	0.1103	0.5306	0.0505	0.0390	0.0517
500	0.0440	0.1336	0.0676	0.1759	0.2332	0.1153	0.0664	0.1120
750	0.0669	0.1650	0.0663	0.1616	0.1582	0.1426	0.0923	0.1241
1000	0.0763	0.1465	0.0759	0.1569	0.1186	0.1207	0.0944	0.1247
1250	0.1022	0.1458	0.0550	0.1276	0.1106	0.1214	0.1120	0.1401
1500	0.1005	0.1309	0.0637	0.1273	0.0954	0.1116	0.1117	0.1372

5 % VaR

Sample	LB(1)	LB(5)	Markov	Weibull	Geometric	Caviar	KS	CVM
250	0.0805	0.1080	0.1280	0.1336	0.0976	0.1053	0.1020	0.1056
500	0.0675	0.1009	0.1284	0.1252	0.0762	0.1010	0.0907	0.0830
750	0.0685	0.1018	0.1659	0.1400	0.0678	0.0964	0.0965	0.0863
1000	0.0891	0.0965	0.2085	0.1423	0.0718	0.0970	0.0951	0.0950
1250	0.0920	0.0925	0.1607	0.1490	0.0608	0.0954	0.0956	0.0982
1500	0.0866	0.0978	0.1515	0.1596	0.0630	0.1029	0.0949	0.0974

Power Study

- We want to assess the ability of the tests to reject a HistSim(250) VaR when the P/L's exhibit volatility dynamics.
- We model the P/L process using a GARCH-t(d), which exhibits volatility dynamics and conditional fat tails.
- Estimating the model on the four observed business line P/Ls confirm the suspicion of excess kurtosis and volatility dynamics.
- Finite sample critical values using Dufour (2004)

Table 3: P/L GARCH Model Parameters and Properties

	Desk 1	Desk 2	Desk 3	Desk 4
d	3.808	3.3183	6.9117	4.7017
θ	-0.245	0.5031	-0.9616	0.0928
β	0.7495	0.9284	0.8728	0.9153
α	0.1552	0.0524	0.0261	0.0723
ω	0.5469	0.2154	0.2127	1.6532
Variance Persistence	0.9140	0.9941	0.9230	0.9882
Unconditional Stdev	2.5220	6.0233	1.6624	11.8478
LogL	-1360.76	-1781.25	-825.87	-1855.98
LogL (HomoSked.)	-1401.64	-1843.49	-831.46	-1877.73
P-value	0.0000	0.0000	0.0108	0.0000

Test Power for 1% VaR

Business Line 1

Sample	LB(1)	LB(5)	Markov	Weibull	Geometric	Caviar	KS	CVM
250	0.1958	0.3201	0.1862	0.1431	0.3258	0.1340	0.3273	0.3305
500	0.2293	0.4200	0.1914	0.1439	0.2642	0.3360	0.4112	0.3563
750	0.2997	0.4755	0.1903	0.1474	0.2758	0.4090	0.4607	0.4075
1000	0.3709	0.5185	0.1678	0.1820	0.3418	0.4810	0.5083	0.4725
1250	0.4335	0.5635	0.1874	0.2280	0.3704	0.5360	0.5421	0.5027
1500	0.4463	0.6027	0.2018	0.2444	0.4098	0.6500	0.5852	0.5638

Business Line 2

Sample	LB(1)	LB(5)	Markov	Weibull	Geometric	Caviar	KS	CVM
250	0.2312	0.2320	0.2110	0.1373	0.3654	0.1420	0.2777	0.2662
500	0.2201	0.2960	0.1896	0.1557	0.3678	0.3040	0.3146	0.2689
750	0.2372	0.3319	0.1806	0.1795	0.3866	0.3390	0.3407	0.2810
1000	0.2806	0.3614	0.1726	0.2181	0.4214	0.3750	0.3895	0.3229
1250	0.2808	0.4001	0.1604	0.2652	0.4752	0.4260	0.3807	0.3265
1500	0.2795	0.4231	0.1604	0.3041	0.5068	0.4830	0.4263	0.3707

Test Power for 1% VaR (cont.)

Business Line 3

Sample	LB(1)	LB(5)	Markov	Weibull	Geometric	Caviar	KS	CVM
250	0.0769	0.1170	0.0725	0.0794	0.1372	0.1020	0.1127	0.1158
500	0.0677	0.1532	0.0625	0.0739	0.0814	0.2000	0.1276	0.1078
750	0.0899	0.1603	0.0529	0.0537	0.0546	0.2420	0.1264	0.1122
1000	0.1059	0.1462	0.0359	0.0507	0.0438	0.3280	0.1374	0.1209
1250	0.1305	0.1273	0.0390	0.0493	0.0470	0.4060	0.1373	0.1295
1500	0.1472	0.1260	0.0311	0.0421	0.0380	0.5470	0.1498	0.1473

Business Line 4

Sample	LB(1)	LB(5)	Markov	Weibull	Geometric	Caviar	KS	CVM
250	0.2502	0.2637	0.2338	0.1593	0.4058	0.1720	0.3132	0.3016
500	0.2400	0.3367	0.2142	0.1839	0.4144	0.3270	0.3824	0.2984
750	0.2803	0.3824	0.2038	0.2118	0.4288	0.3500	0.4030	0.3216
1000	0.3326	0.4193	0.2019	0.2671	0.5026	0.4350	0.4494	0.3753
1250	0.3170	0.4580	0.1955	0.3432	0.5438	0.4960	0.4546	0.3919
1500	0.3293	0.5095	0.2018	0.3889	0.5968	0.5630	0.4881	0.4272

Research Questions Revisited

- What do actual daily desk-level P/Ls and VaRs look like? **Volatility dynamics and non-normal P/Ls. Sometimes erratic VaRs.**
- Model-free Historical Simulation (HistSim) VaRs seem to be gaining popularity in banks. Is this a positive development? **Not if volatility dynamics are present in P/Ls.**
- If volatility dynamics are present in P/L's do we have any chance of rejecting HistSim VaRs in realistic sample sizes? **Surprisingly, yes.**
- If so which test should we use? **Geometric Duration and Caviar.**
- Can we rely on the asymptotic critical values of these tests. **Not generally.**

Broader Policy Recommendations

- On Backtesting
  - Power is reasonable for 500 and more observations
  - Conditional backtesting is crucial
- On VaR Computation
  - HistSim allows for non-normality, but ignores volatility dynamics.
  - More frequent historical data updating and longer sample are important for future improvements.
  - Augment HistSim VaR with GARCH filter on aggregate P/L. “Filtered HistSim”
  - Report 1%,2%,3%,4%,5% VaRs to get a sense of the shape of the P/L distribution tail.

Table 7: Backtesting Actual VaRs from Four Business Lines

Business Line 1

	LB(1)	LB(5)	Markov	Weibull	Geometric	Caviar	KS	CVM
Test Value	0.0955	0.4830	0.1961	1.0138	1.2896	2.9296	18.7481	2.4382
P-Value	0.4601	0.5513	0.9628	0.6619	0.3758	0.3189	0.3241	0.3948

Business Line 2

	LB(1)	LB(5)	Markov	Weibull	Geometric	Caviar	KS	CVM
Test Value	0.0315	0.1594	1.4576	3.6338	3.8377	4.1389	11.5003	1.3503
P-Value	0.8252	0.8379	0.3199	0.2354	0.1251	0.2339	0.4671	0.5618

Business Line 3

	LB(1)	LB(5)	Markov	Weibull	Geometric	Caviar	KS	CVM
Test Value	0.0016	0.0083	6.8487	NaN	NaN	27.7656	70.3651	70.3651
P-Value	0.9920	0.9919	**0.0176**			**0.0124**	**0.0533**	**0.0242**

Business Line 4

	LB(1)	LB(5)	Markov	Weibull	Geometric	Caviar	KS	CVM
Test Value	0.0263	38.5720	0.9752	4.4235	4.9972	4.1304	19.6270	5.2362
P-Value	0.7845	**0.0093**	0.3687	0.1720	**0.0602**	0.2010	0.1815	0.1801