Data description

In this appendix we provide a description of the dataset used in the application. Section 1 concerns the cross-sectional data, while Section 2 concerns the historical data.

1 Cross-sectional data

We first describe the raw dataset. Then, we discuss the data processing steps that we have followed to obtain the dataset used for XMM estimation.

The raw dataset is a subset of daily data of OptionMetrics for European call and put options on the S&P 500 index from June 1, 2005 to June 30, 2005, that is 22 trading days. For any quoted maturity and strike, the raw dataset reports the daily volume of the call and put contracts and the closing bid/ask quotes. The daily volume is the total number of contracts traded within the day. This variable is used to select the options that are actively traded. More precisely, in the application we only consider the options with daily trading volume larger than 4000 contracts.

This raw dataset is treated to get a transformed dataset suitable for XMM estimation. We compute an option price as average between closing bid/ask quotes. Then, we apply four preprocessing steps. (i) First, we eliminate options with time-to-maturity larger than 250 days. This step avoids kernel regression involving long histories of returns $Y_t = (X_{t+1}, ..., X_{t+\bar{h}})'$ with $\bar{h} > 250$, for which too few observations $T - \bar{h}$ are available when $T = 1000$. (ii) Second, we eliminate the options with extreme moneyness strikes. More precisely, we eliminate the calls (resp. the puts) with time-to-maturity h and log-discounted moneyness strike above the 99% quantile (resp. below the 1% quantile) in the distribution of cumulated index excess returns at horizon h. These options are eliminated since the corresponding sample local moment restrictions are numerically very close to zero for any value of the parameters, making less robust the estimation. (iii) Third, when both the call and put options for the same moneyness strike and time-to-maturity are actively traded, we retain the put if the discounted moneyness strike is smaller than 1, and the call otherwise. This step avoids introducing arbitrage opportunities in the set of observed option prices, and is compatible with the procedure suggested by Aït-Sahalia and Lo (1998). (iv) Finally, we eliminate the option price observations that violate the no-arbitrage restrictions on the price curves, namely that the option price curve is monotonic and convex with respect to the moneyness strike, for any given time-to-maturity.

To each option corresponds a riskfree yield with the same time-to-maturity, derived by a spline interpolation from T-bond price data available on Datastream.

2 Historical data

The historical data consists of the daily excess return and realized volatility of the S&P 500 index for the period from June 1, 2001, to June 30, 2005. The S&P 500 index daily excess return is the difference between the S&P 500 index daily log-return and the 1-day riskfree yield. The former is computed from the S&P 500 index level, and the latter from T-bond price data by spline interpolation. Data are collected from Datastream. The daily realized volatility of S&P 500 index is computed with a frequency of 30 minutes, after a preliminary trimming to eliminate outliers.