North American Journal of Economics and Finance 31 (2015) 1–26

Contents lists available at ScienceDirect

North American Journal of Economics and Finance

Implied volatility and the risk-free rate of return in options markets☆

Marcelo Bianconi[a,*], Scott MacLachlan[b], Marco Sammon[c]

[a] Department of Economics, Tufts University, USA
[b] Department of Mathematics and Statistics, Memorial University of Newfoundland, Canada
[c] Federal Reserve Bank of Boston, USA

ARTICLE INFO

Article history:
Received 15 May 2014
Received in revised form 14 October 2014
Accepted 16 October 2014
Available online 4 November 2014

JEL classification:
G13
C63

Keywords:
Re-pricing options
Forecasting volatility
Seemingly unrelated regression
Implied volatility

ABSTRACT

We numerically solve systems of Black–Scholes formulas for implied volatility and implied risk-free rate of return. After using a seemingly unrelated regressions (SUR) model to obtain point estimates for implied volatility and implied risk-free rate, the options are re-priced using these parameters. After repricing, the difference between the market price and model price is increasing in time to expiration, while the effect of moneyness and the bid-ask spread are ambiguous. Our varying risk-free rate model yields Black–Scholes prices closer to market prices than the fixed risk-free rate model. In addition, our model is better for predicting future evolutions in model-free implied volatility as measured by the VIX.

☆ This research is based on Sammon's thesis at Tufts University, winner of the Linda Datcher-Loury award. We thank Dan Richards and the Tufts Department of Economics for funding the research data and Tufts University for use of the High-Performance Computing Research Cluster. Any errors are our own. The views expressed here are solely those of the authors and do not necessarily reflect official positions of the Federal Reserve Bank of Boston or the Federal Reserve System.

* Corresponding author at: Tufts University, Department of Economics, 111 Braker Hall Medford, MA 02155, USA. Tel.: +1 617 627 2677; fax: +1 617 627 3917.
E-mail addresses: Marcelo.Bianconi@tufts.edu (M. Bianconi), smaclachlan@mun.ca (S. MacLachlan), Marco.Sammon@bos.frb.org (M. Sammon).
URL: http://www.tufts.edu/ mbiancon (M. Bianconi).

http://dx.doi.org/10.1016/j.najef.2014.10.003
1062-9408/

1. Introduction

There are investment firms that pay people to sit outside of factories with binoculars to count the number of trucks going in and out. Investors do this because each truck contains information and they believe having this information before anyone else gives them an advantage in financial markets. Every day, several thousands of options are traded and each trade contains information. In the same way that information is gained by watching trucks, there must be a way to capture information by observing the options market. We develop a model to do just that.

Ever since Black and Scholes (1973), both academics and finance practitioners have used it to garner information from the options market. Implied volatility is calculated by inverting the Black–Scholes formula, given the market price of an option. Option-implied stock market volatility even became a tradable asset when the Chicago Board Options Exchange launched the CBOE Market Volatility Index (VIX) in 1993. Becker, Clements, and White (2007), among others find that the VIX does not contain information relevant to future volatility beyond that available from model based volatility forecasts.[1] Alternatively, Hentschel (2003) shows that estimating implied volatility by inverting the Black–Scholes formula is subject to considerable imprecision when option characteristics are observed with plausible errors.[2]

The risk-free rate is needed to calculate implied volatility by inverting the Black–Scholes formula and it is usually approximated with the yield on Treasury bills. The risk-free rate is important, as it determines the no-arbitrage condition. Treasury bills yields, however, may not capture the risk-free rate implicitly used by market participants to price options for a number of reasons. People buying and selling Treasury bills probably have different funding costs than those trading options and Treasury bill yields are influenced by Federal Reserve asset purchasing programs and as a result, may not reflect market forces.

Setting the risk-free rate equal to Treasury bill yields complicates the interpretation of implied volatility, as it then contains information on investor expectations for both the risk-free rate and the underlying security's volatility. This problem can be solved by setting up a system of two Black–Scholes formulas in two unknowns and solving simultaneously for implied volatility and implied risk-free rate. We believe allowing the risk-free rate to vary better isolates the implied volatility implicit in options prices.

Empirically, implied volatility and implied risk-free rate differ among options on the same underlying security with different strike prices. We build on the methods of Macbeth and Merville (1979) and Krausz (1985), using a seemingly unrelated regressions (SUR) model to calculate point estimates of at-the-money implied volatility and implied risk-free rate for each underlying security. These point estimates are used in the Black–Scholes formula to re-price the options.

We examine the impact of moneyness, time to expiration and the size of the bid-ask spread on the difference between market prices and model-based Black–Scholes prices. We find that as time to expiration increases, the difference between market and model prices increases. In almost every regression specification, the coefficients on moneyness and moneyness squared have opposing signs. We believe this is explained by the volatility smile.

The difference between implied volatility calculated using a fixed risk-free rate and the same quantity calculated with a varying risk-free rate increases over the sample period, indicating that the additional information becomes more important as the sample period progresses. The correlation between the two measures of implied volatility is positive across all leads and lags. We believe the varying risk-free rate model better fits the data because it yields a smaller average difference between the market price and the model-based Black–Scholes price.

The model outlined above extracts additional information from the options market. We measure the marginal impact of allowing the risk-free rate to vary on the volatility smile and the accuracy of VIX

[1] See also Canina and Figlewski (1993) and Christensen and Prabhala (1998).

[2] Also, Jiang and Tian (2005) find that model-free implied volatility is a more efficient forecast of future realized volatility than model based implied volatility. On the general relationship between risk management and financial derivatives, see Hammoudeh and McAleer (2013).

forecasting. The volatility smile changes shape when using the simultaneous solution method because there is a balancing effect between the risk-free rate and implied volatility. In Fig. 9, there is a pattern for the implied risk-free rate across strikes that is the inverse of the pattern for implied volatility. This balancing is not enough to get rid of the volatility smile, so the problem remains unresolved. For forecasting the VIX, our measure of implied volatility is superior to traditional implied volatility. This result holds both in-sample and out-of-sample, as measured by the Diebold–Mariano test.

Finally, we examine potential trading strategies based on the discrepancy between Black–Scholes prices and market prices and based on the predictability of the VIX. The simultaneous and fixed risk-free rate solutions yield alternative relative performances in the sample period.

The paper is organized as follows. Section 2 discusses the simultaneous solution for implied volatility and the implied risk-free rate. Section 3 goes over the at-the-money adjustment using the seemingly unrelated regressions model. Section 4 discusses the data used in our analysis and the output of our numerical solution. Section 5 examines factors that explain the difference between model prices and market prices. Section 6 investigates the marginal effect of allowing the risk-free rate to vary in several finance problems, while Section 7 overviews potential trading strategies and Section 8 concludes. An appendix presents additional information on our numerical solution as well as performance and sensitivity analysis.

2. Simultaneous solution for implied volatility and implied risk-free rate

This section begins with a review of the Black–Scholes formula and implied volatility, followed by a review of the literature on simultaneous solutions for implied volatility and implied risk-free rate. We then provide a description of our algorithm for finding the simultaneous solution.

2.1. Black–Scholes formula and implied volatility

Black and Scholes (1973) created the following model for pricing a European call option:

$$CallPrice = \phi(d_1)S - \phi(d_2)Ke^{-r\tau} \tag{1}$$

where $d_2 = \frac{\ln\left(\frac{S}{K}\right)+\left(r-\frac{\sigma^2}{2}\right)\tau}{\sigma\sqrt{\tau}} = d_1 - \sigma\sqrt{\tau}$, S is the spot price of the underlying security, $\phi(.)$ is the normal CDF, K is the strike price, r is the risk-free rate of return, σ is the volatility of returns of the underlying asset and τ is the option's time to expiration.

For any call traded on an exchange S, K and τ are known, but σ and r, which are meant to be forward looking, cannot be observed directly. Finance practitioners applying Black–Scholes to price options approximate the risk-free rate with the annualized yield on Treasury bills and approximate future volatility with past volatility.

Implied volatility for the underlying asset's returns can be calculated if the option's market price is known. After deciding on an appropriate value for r, it is a case of one equation and one unknown. There is no closed-form solution for implied volatility, so an optimization routine is needed. When solving for implied volatility using Newton's method, the goal is to minimize the quadratic function $[C^* - C(\sigma_n)]^2$ given S, K, and τ from market data and r from Treasury bill yields, where C^* is the market price for the call and $C(\sigma)$ is the Black–Scholes formula evaluated at σ. Solving for σ is useful, as it captures investor sentiment about the volatility of the underlying asset, but we believe it still leaves out important information.

The true risk-free rate is not observable in the market, so it would be better if both σ and r could be extracted from options data. This eliminates the need to approximate the risk-free rate with the Treasury bill yields, which may not accurately capture option traders' expectations of changes in the discount rate. The following sections discuss methods for finding both σ and r.

2.2. Simultaneous solutions for implied volatility and implied risk-free rate

Pairs of call options on the same underlying security with the same time to expiration and different strike prices are needed to solve simultaneously for implied volatility and implied risk-free. This yields

a system of two equations and two unknowns, which is solved for the parameters of interest. Several authors, such as Krausz (1985), O'Brien and Kennedy (1982) and Swilder (1986) use various methods to find simultaneous solutions for σ and r. We build on their models, using modern mathematics software packages which allow for the use of larger datasets and more precision in the estimates for σ and r. Appendix A discusses why an optimization routine is needed to find this simultaneous solution.

The goal of a simultaneous solution is to solve both $C_1(\sigma, r) = C_1^*$ and $C_2(\sigma, r) = C_2^*$ where C_1^* and C_2^* are the calls' market prices and $C_1(\sigma, r)$ and $C_2(\sigma, r)$ are the first and second calls priced with the Black–Scholes formula evaluated at σ and r. Given that σ and r both enter non-linearly into the Black–Scholes formula, these parameters cannot be solved for directly. Krausz's algorithm picks a starting point and adjusts σ and r by small increments, $\delta\sigma$ and δr, until a solution to the system is found.

To determine the change in σ and r, Krausz's algorithm solves the following system for $\delta\sigma$ and δr:

$$C_1(\sigma + \delta\sigma, r + \delta r) = C_1^* \quad \text{and} \quad C_2(\sigma + \delta\sigma, r + \delta r) = C_2^* \tag{2}$$

To simplify the problem, Krausz uses the following first order Taylor approximation, which is valid for small $\delta\sigma$ and δr:

$$C_i(\sigma_n + \delta\sigma, r_n + \delta r) \approx C_i(\sigma_n, r_n) + \frac{\partial C_i}{\partial \sigma}(\sigma_n, r_n)\delta\sigma + \frac{\partial C_i}{\partial r}(\sigma_n, r_n)\delta r \tag{3}$$

Given the Taylor approximation, Eq. (2) is rewritten in matrix form and solved for $\delta\sigma$ and δr:

$$\begin{bmatrix} \delta\sigma \\ \delta r \end{bmatrix} = \frac{1}{\frac{\partial C_1}{\partial \sigma} \times \frac{\partial C_2}{\partial r} - \frac{\partial C_1}{\partial r} \times \frac{\partial C_2}{\partial \sigma}} \begin{bmatrix} \frac{\partial C_2}{\partial r} & -\frac{\partial C_1}{\partial r} \\ -\frac{\partial C_2}{\partial \sigma} & \frac{\partial C_1}{\partial \sigma} \end{bmatrix} \begin{bmatrix} C_1^* - C_1(\sigma_n, r_n) \\ C_2^* - C_2(\sigma_n, r_n) \end{bmatrix} \tag{4}$$

If there is a solution to this system, $\delta\sigma$ and δr are added to σ and r. This process of finding $\delta\sigma$ and δr and adding them to σ and r is repeated until the assumptions of the Black–Scholes model are violated or the desired level of precision is reached. This method is computationally expensive, as it requires evaluating four derivatives of the Black–Scholes formula at each step. In addition, this method may not always find a solution as it relies on the assumption that a σ and r pair exists such that $C_1(\sigma, r) = C_1^*$ and $C_2(\sigma, r) = C_2^*$, which may not be the case for noisy real-world data.

2.3. Proposed method for finding a simultaneous solution

Rather than set up two equations in two unknowns, we propose a single equation to be minimized for σ and r with an optimization routine:

$$F(\sigma, r) = \frac{1}{(C_1^*)^2}(C_1^* - C_1(\sigma, r))^2 + \frac{1}{(C_2^*)^2}(C_2^* - C_2(\sigma, r))^2 \tag{5}$$

This equation has advantages over the two equations model discussed above. First, it has a mechanism for weighing the difference between the Black–Scholes price and the market price. Without this, the solution for σ and r will be biased toward minimizing $C_2^* - C_2(\sigma, r)$ if the second option in the pair is demonstrably more expensive. In addition, this method has a lower failure rate because it does not rely on the assumption that an exact solution exists such that $F(\sigma, r) = 0$. The function F is a measure of the quality of our solutions and is sensitive to moneyness, bid-ask spread and time to expiry.

A number of different algorithms in MATLAB were tested for minimizing this equation. We focus on algorithms built into the *fmincon* function while alternative optimization models are discussed in Appendix B. Within the *fmincon* function, we experimented with two algorithms: sequential quadratic programming (SQP) and interior point. In both cases, the algorithms actually run slower when we provide gradient and Hessian information, so we use derivative-free approaches. These methods approximate the gradient using finite differences and use the Broyden–Fletcher–Goldfarb–Shanno (BFGS) method to approximate the Hessian. Both SQP and interior point use similar approaches at each step of the optimization, but they implement range constraints differently.

The interior point algorithm tries to find a point where the gradient is equal to zero, but it weighs the quality of solutions by how close they are to the range constraints. We impose the restriction that σ and r must be between zero and one, because values outside of this range are empirically unrealistic. When looking for a minimum, this algorithm chooses points that are opposite the direction of the gradient until it reaches a balance between getting the gradient close to zero and staying far enough away from the edge of the feasible set.

The SQP algorithm takes a second-order Taylor Series approximation of the function to be minimized. The quadratic approximation is minimized directly, similar to the linearizations used in Newton's method, allowing iterative improvement of the approximate minimization of the non-linear and non-quadratic function F. Unlike interior point, this algorithm does not discount the quality of solutions where σ and r are close to the edge of the feasible set. The algorithm continues taking these Taylor Series approximations to adjust σ and r, until the gradient of F is sufficiently near zero.

Interior point is, on average, faster than SQP, even though it requires more calculations per iteration. In addition, interior point achieved more accurate solutions, with smaller average values of F.[3] Both of these algorithms are faster and more accurate than the algorithm developed by Krausz (1985) and two benchmark alternatives: a brute-force approach that directly samples $F(\sigma, r)$ on an evenly spaced mesh of 200 values of σ and r, for a total of 40,000 points; and an algorithm that applies a classical Newton's method to directly minimize F by satisfying its first-order optimality conditions. The final algorithm works as follows:

1 The starting values of σ and r for each pair of options is $(0.5, r_t)$ where r_t is the Treasury bill yield on that day.
2 The interior point algorithm finds a simultaneous solution for σ and r, starting at the point determined in step 1.
3 The *patternsearch* algorithm, which is another derivative-free method in MATLAB's optimization toolbox, is run to minimize F, starting from the point found in step 2, to find another possible solution for σ and r.
4 Starting at the point found in step 3, the interior point algorithm is run again to minimize F and find a third possible solution for σ and r.
5 The algorithm compares the three values of F from steps 2, 3 and 4, choosing the (σ, r) pair which yields the smallest F.

3. At-the-money adjustment

After running the algorithm discussed in Section 2, we obtain an implied volatility and implied risk free rate for each pair of options. In order to proceed with re-pricing the options, we need a point estimate for implied volatility and implied risk free rate for each underlying security. This section reviews others' approaches and discusses our method for weighing the individual estimates.

Empirically, options on the same underlying security with different strike prices have different implied volatilities. Given that implied volatility is supposed to be a measure of volatility for the underlying security, an adjustment is needed to extract a single point estimate for this parameter.

We believe these point estimates should not be calculated using an arithmetic average, given the presence of two common phenomena for options: volatility skew and volatility smile. For call options, Volatility skew is when implied volatility is highest for in-the-money options and decreases steadily as strike prices increase. The volatility smile is when volatility is lowest for at-the-money options and it increases as options become deeper in-the-money or farther out-of-the-money. An example of a volatility smile is presented in Fig. 1.

To adjust for the volatility skew and the volatility smile, Macbeth and Merville (1979) run the following ordinary least squares (OLS) regression:

$$\sigma_{jkt} = \phi_{0kt} + \phi_{1kt} M_{jkt} + \varepsilon_{jkt} \tag{6}$$

[3] See Appendix B for a detailed comparison of speed and accuracy between these two algorithms.

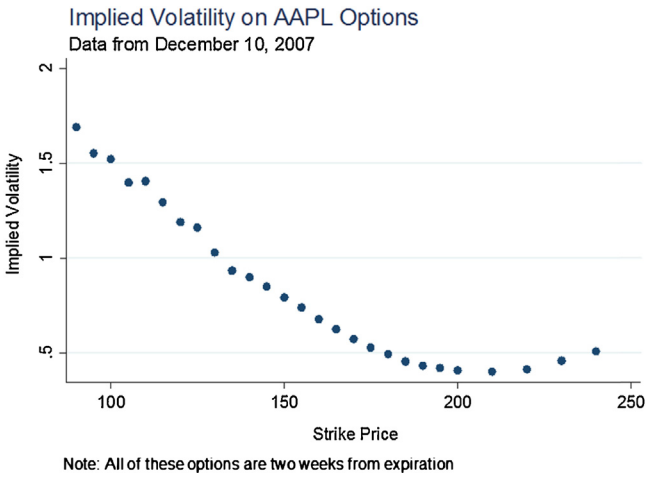

Fig. 1. Volatility smile for AAPL options. Implied volatility is different for each strike price. It is highest for in-the-money options, decreasing steadily as strike price increases until it hits the closing price of 194.21, where it begins to increase again.

where σ_{jkt} is the model implied volatility for an option j, on security k, at time t, ε_{jkt} is the error term and $M_{jkt} = \frac{S_{kt} - X_{jk}e^{-r\tau}}{X_{jk}e^{-r\tau}}$ is a measure of moneyness. S_{kt} is the price of the underlying security, $X_{jk}e^{-r\tau}$ is the present value of the option's strike price and r is the Treasury bill yield. M_{jkt} equals zero when, on a present value basis, an option is at-the-money, so the estimate $\widehat{\phi_{0kt}}$ is the implied at-the-money volatility. Essentially, $\widehat{\phi_{0kt}}$ is the part of implied volatility that cannot be explained by variation in moneyness.

Krausz (1985) adapts this technique to his simultaneous solution for σ and r. He runs an OLS regression to adjust each parameter:

$$\sigma_{jkt} = \phi_{0kt} + \phi_{1kt}M_{jkt} + \varepsilon_{jkt} \tag{7}$$

$$r_{jkt} = \rho_{0kt} + \rho_{1kt}M_{jkt} + \epsilon_{jkt} \tag{8}$$

where σ_{jkt} and r_{jkt} are the model implied values for implied-volatility and risk-free rate for an option, j on security, k at time, t. In addition, $M_{jkt} = \frac{S_{kt} - X_{jk}e^{-r^*\tau}}{X_{jk}e^{-r^*\tau}}$ where S_{kt} is the price of the underlying security, $X_{jk}e^{-r^*\tau}$ is the present value of the option's strike price and r^* is the average model implied risk-free rate across all securities on a given date. In this model, ε_{jkt} and ϵ_{jkt} are error terms. The at-the-money implied volatility and risk free rate for each security are $\widehat{\phi_{0kt}}$ and $\widehat{\rho_{0kt}}$. As in the Macbeth and Merville (1979) model, $M_{jkt} = 0$ when, on a present value basis, an option is at the money.[4]

3.1. Proposed at-the-money adjustment

Using the average value of r in the calculation of M_{jkt}, r^*, causes an endogeneity problem because r_{jkt} is on both sides of Eq. (8). In addition, the fact that these regressions are run separately omits the

[4] While pairs of options are assigned the same values for σ and r, these equations are estimated at the option level, rather than option pair level. This is because in each pair, the options have different strikes, and as a consequence, different values for M_{jkt}.

simultaneity of the solution for σ and r. We rewrite M_{jkt} to isolate r. First, we move all terms that contain r to the left hand side of the equation:

$$M_{jkt} = \frac{S_{kt} - X_{jk}e^{-r_{jkt}\tau}}{X_{jk}e^{-r_{jkt}\tau}} \rightarrow X_{jk}e^{-r_{jkt}\tau}(1 + M_{jkt}) = S_{kt} \tag{9}$$

Then, we take the natural log of each side and solve for M_{jkt}:

$$\ln(1 + M_{jkt}) = -\ln(X_{jk}) - \ln(e^{-r_{jkt}\tau}) + \ln(S_{kt}) \tag{10}$$

For $M_{jkt} \approx 0$, we have that $\ln(1 + M_{jkt}) \approx M_{jkt}$, so we can approximate Eq. (10) as

$$M_{jkt} = -\ln(X_{jk}) + r_{jkt}\tau + \ln(S_{kt}) \tag{11}$$

Substituting this into Krausz's Eqs. (7) and (8) gives:

$$\sigma_{jkt} = \phi_{0kt} + \phi_{1kt}(\ln(S_{kt}) - \ln(X_{jk}) + r_{jkt}\tau) + \varepsilon_{jkt} \tag{12}$$

$$r_{jkt} = \rho_{0kt} + \rho_{1kt}(\ln(S_{kt}) - \ln(X_{jk}) + r_{jkt}\tau) + \epsilon_{jkt} \tag{13}$$

An additional adjustment is required because r_{jkt} is still on both sides of Eq. (13). This equation is solved explicitly for r_{jkt} as follows:

$$r_{jkt} = \frac{\rho_{0kt}}{1 - \rho_{1kt}\tau} + \frac{\rho_{1kt}}{1 - \rho_{1kt}\tau}(\ln(S_{kt}) - \ln(X_{jk})) + \epsilon'_{jkt} \tag{14}$$

As with the Macbeth and Merville (1979) model, the constant terms, ϕ_{0kt} and ρ_{0kt} represent the at-the-money implied volatility and risk-free rate for the underlying security because $\ln(S_{kt}) - \ln(X_{jk}) + r_{jkt}\tau = 0$ for options expiring at the money. We explicitly identify ρ_{0kt} as follows: We start by taking a linear approximation of $\frac{1}{1-\rho_{1kt}\tau}$. We rewrite $\frac{1}{1-\rho_{1kt}\tau}$ as the geometric series $1+\rho_{1kt}\tau+(\rho_{1kt}\tau)^2+\ldots$ for $\rho_{1kt}\tau < 1$, which given the units is safe to assume. While a higher order approximation would give more accurate results, this must be weighed against the additional computational cost. A first order approximation yields:

$$r_{jkt} = \rho_{0kt}[1 + \rho_{1kt}\tau] + \rho_{1kt}\tau[1 + \rho_{1kt}\tau](\ln(S_{kt}) - \ln(X_{jk})) + \epsilon''_{jkt} \tag{15}$$

which is simplified as follows:

$$r_{jkt} = \beta_0 + \beta_1\tau + \beta_2(\ln(S_{kt}) - \ln(X_{jk})) + \beta_3(\ln(S_{kt}) - \ln(X_{jk})) \times \tau + \epsilon''_{jkt} \tag{16}$$

In this equation, $\beta_0 = \rho_{0kt}$, is the model-implied at-the-money risk-free rate.

Another improvement in our at-the-money adjustment is that it accounts for the simultaneity of σ and r by using a seemingly unrelated regressions (SUR) model to estimate Eqs. (12) and (16). A SUR model requires that the error terms across the regressions are correlated, which makes sense for our data, given that every σ and r pair is extracted from two options on the same underlying security with the same time to expiration.

The implications of using the SUR model are more evident when considering shocks that enter the system. Without using SUR, a shock in the error term for the risk-free rate regression has no impact on the volatility regression. With SUR, a shock in the error term for the risk-free rate regression also shocks the error term for implied volatility regression and vice versa. Bliss and Panigirtzoglou (2004) explain that, "risk preferences are volatility dependent". Given the close relationship between implied risk-free rate and implied volatility, it makes sense that a shock affecting one of these equations should also affect the other. Finally, if there is correlation between the error terms in the two equations, using the SUR yields smaller standard errors for the estimated coefficients.[5]

[5] SUR only achieves smaller standard errors if two conditions are met: (1) There is a correlation between the error terms in each equation and (2) the two equations have different independent variables. In our model, the right hand side is different for each equation and we believe the standard errors are correlated, so there should be efficiency gains.

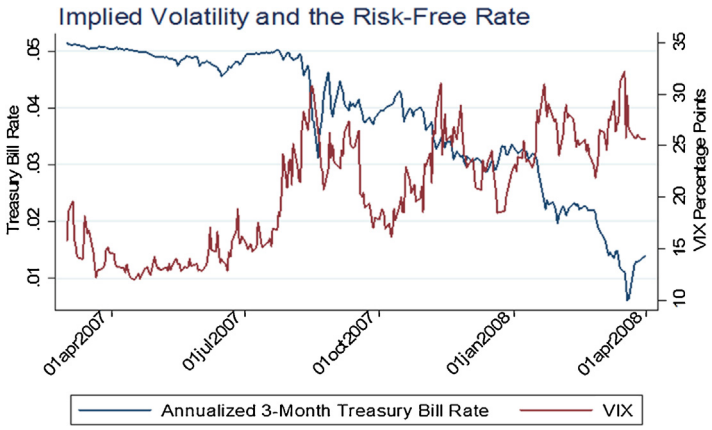

Fig. 2. Trends over sample period. Our model is estimating option implied values for the risk-free rate and volatility. We believe the VIX and 3-month Treasury bills are market-based benchmarks for these quantities. Volatility is increasing and the risk-free rate is declining over the sample period.

4. Data

4.1. Data sources and description

The options dataset is from http://www.historicaloptiondata.com and it contains end of day quotes on all stock options for the U.S. equities market. This includes all stocks, indices and ETFs for each strike price and time to expiration. Data on the VIX, Treasury bills and other market indices is collected from the Federal Reserve Bank of St. Louis. Our empirical analysis uses data from March 2007 to March 2008. Given that we are interested in implied volatility and the risk-free rate, Fig. 2 shows the evolution of these two quantities, as measured by the VIX[6] and 3-month Treasury bills. The risk free rate is trending downward, while implied volatility is increasing.

4.2. Restrictions on the data

We restrict the data using a procedure similar to that of Constantinides, Jackwerth, and Savov (2012). First, the interest rate implied by put-call parity is computed. The equation for put-call parity is solved algebraically for $r_{Put-Call}$ as follows:

$$r_{Put-Call} = \frac{-\ln\left(\frac{S+P-C}{K}\right)}{t} \tag{17}$$

All the observations with values of $r_{Put-Call}$ that do not exist or are less than zero are dropped. Constantinides et al. (2012) removed these options because negative or nonexistent values for $r_{Put-Call}$ suggest that the options are mispriced. All of the puts are then removed from the dataset, as we only analyze call options. Other procedures in line with Constantinides et al. are the removal of options with bid prices of zero and options with zero open interest. Options with zero volume for a given day are allowed to remain in the dataset.[7]

[6] The new VIX is a model free calculation of volatility based upon the prices of S&P500 index options and it does not rely on the Black–Scholes framework. See e.g. CBOE Volatility Index – VIX (2009).

[7] Open interest is the total number of option contracts that have been traded, but not yet liquidated. Volume is the number of option contracts traded on a given day.

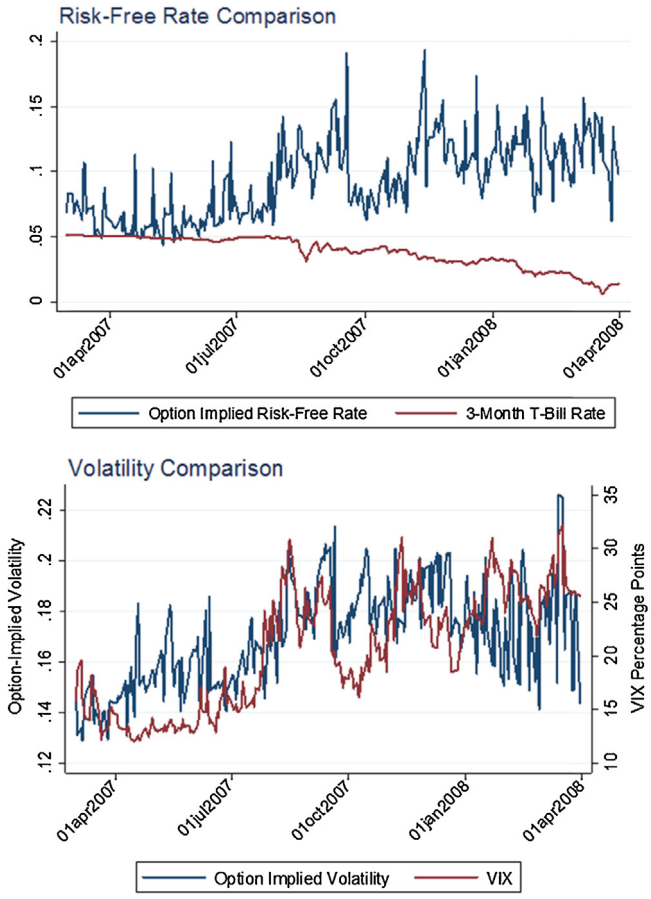

Fig. 3. Comparison between market and model. The top pane of this figure is the daily average risk-free rate implied by options, which increased during the sample period, while the yield on Treasury bills declined. The bottom pane of this figure presents our measure of implied volatility, which sometimes follows a similar trend to the VIX, but does not track it well. These averages are constructed with data that has not been adjusted with the SUR model.

4.3. Descriptive statistics

Fig. 3 graphs the evolution of average option-implied volatility and risk-free rate for SPX options against their market benchmarks.[8] In this analysis, and all future analysis, calculations involving implied volatility and implied risk-free rate are restricted to observations where these parameters take values between zero and one. As discussed in Section 2, the algorithm that solves for the initial values of σ and r already sets these boundaries. We need to make this restriction again, however, because the SUR model does not place restrictions on $\widehat{\phi_{0kt}}$ and $\widehat{\rho_{0kt}}$.

The average option-implied risk-free rate does not track Treasury bill rates and the average option-implied volatility does not track the VIX index. The lack of relationship between the series calculated using our methodology and the benchmark series indicates to us that Treasury bill yields do not accurately represent the risk-free rate and we are getting new information through the simultaneous solution.

[8] There are two commonly traded S&P 500 index options, SPX and SPY. SPX options are based on the entire basket of underlying securities and are settled in cash while the SPY options are based on an ETF and settled in shares.

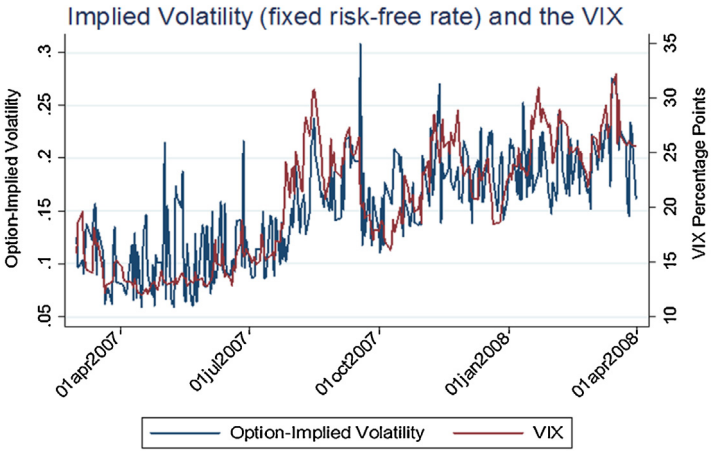

Fig. 4. Implied volatility calculated with fixed risk-free rate. The average daily option implied volatility calculated with a fixed-risk free rate tracks the VIX better than our measure of implied volatility.

Table 1
Descriptive statistics under alternative restrictions.

	No restrictions	Restrict M'	Restrict time to expiration	Restrict bid-ask spread	Restrict F
Average risk-free rate	0.0932	0.0944	0.0883	0.0907	0.0946
Average ATM risk-free rate	0.1237	0.1214	0.1218	0.1210	0.1228
Average implied volatility	0.3536	0.3365	0.3437	0.3504	0.3509
Average ATM implied volatility	0.3767	0.3645	0.3797	0.3708	0.3735

In Fig. 4, we compare implied volatility, calculated with the risk-free rate fixed at the yield on 3-month Treasury bills, to the VIX index. Unlike Fig. 3, these two series follow a similar trend. This makes sense, because according to the CBOE website: "The risk-free interest rate, R, is the bond-equivalent yield of the U.S. T-bill maturing closest to the expiration dates of relevant SPX options." Both of these calculations use the same fixed risk free rate and as a result, have similar values for implied volatility.

Table 1 presents descriptive statistics for all options. We experimented with several different restrictions on the data using moneyness, time to expiration, the size of the bid-ask spread, and F. We define:

$$M'_{jkt} = \frac{S_{kt} - X_{jk}^{-r_T\tau}}{X_{jk}^{-r_T\tau}} \tag{18}$$

where r_T is the yield on 3-month Treasury bills. Our restriction on M' is to within one standard deviation of its mean, excluding far in the money and out of the money options. The restriction on time to expiration removes options expiring fewer than 90 days in the future. We define:

$$Spread_{jkt} = \frac{Ask_{jkt} - Bid_{jkt}}{Midpoint_{jkt}} \tag{19}$$

and the restriction on the spread is to within one standard deviation of its mean, excluding options with large spreads. The restriction on F is to values smaller than $\frac{1}{100}$.

In almost all cases, the restrictions do not make a significant difference in the variables' averages. The only obvious change is for the risk-free rate before the at-the-money adjustment, when restricting time to expiration or the bid-ask spread. This implies that options closer to expiration or with larger spreads have higher option-implied risk free rates. This is in line with mispricing that occurs when options are close to expiration or are thinly traded.

Table 2
Descriptive statistics under alternative restrictions cont.

	No restrictions	Restrict M'	Restrict time to expiration	Restrict bid-ask spread	Restrict F
Moneyness	0.2117	0.1112	0.2407	0.2626	0.2288
Moneyness squared	0.3418	0.0708	0.4202	0.3741	0.3541
Time to expiration	0.6161	0.5606	0.8442	0.6246	0.6271
TTE squared	0.7369	0.6320	1.0686	0.7492	0.7539
Spread	0.1792	0.1742	0.1709	0.1020	0.1626
Spread squared	0.0995	0.0894	0.0902	0.0196	0.0828
F	0.0072	0.0066	0.0039	0.0044	0.0002
Y	0.0231	0.0214	0.0315	0.0233	0.0229
# Obs	15,700,000	13,600,000	10,800,000	14,000,000	15,000,000

5. Determinants of underpricing/overpricing

This section analyzes factors that explain the difference between model-based Black–Scholes prices and market prices.

5.1. Macbeth and Merville regressions

Macbeth and Merville (1979) solved for implied volatility and after making the at-the-money adjustment to this parameter, they re-price the options using using the Black–Scholes formula. They then examine the determinants of differences between market prices and Black–Scholes prices. They run a regression of $Y = Call_{Market} - Call_{B-S}$ on moneyness and time to expiration. We propose a similar model, but redefine Y to be a relative difference:

$$Y = \frac{|Call_{Market} - Call_{B-S}|}{UnderlyingPrice} \tag{20}$$

and add additional terms to their regression:

$$Y_{jkt} = \alpha_1 + \beta_1 M'_{jkt} + \beta_2 M'^2_{jkt} + \beta_3 \tau_{jkt} + \beta_4 \tau^2_{jkt} + \beta_5 Spread_{jkt} + \beta_6 Spread^2_{jkt} + \varepsilon_{jkt} \tag{21}$$

where Y_{jkt}[9] is the relative difference between the market price and the Black–Scholes model price for an option j, on security k, at time t, M'_{jkt} is defined in Eq. (18), τ_{jkt} is the time to expiration, *Spread* is defined in Eq. (19) and ε_{jkt} is an error term.

Table 2 presents the averages across all options for key variables in each specification. The restrictions on moneyness, time to expiration, the spread and F are the same as those described in Section 4.

Table 3 presents the regression results. In almost every specification, the coefficient on moneyness is negative, while the coefficient on moneyness squared is positive. This makes sense, given the volatility smile discussed in Section 3. This pattern does not hold for the specification that restricts moneyness, which implies that there are be far in the money and out of the money options that are biasing the regression results. The coefficients on time to expiration and time to expiration squared are both positive across all specifications, indicating that as options get far from expiration, their model prices diverge from market prices Finally, the coefficient on spread is negative, but the coefficient on spread squared is positive, indicating an ambiguous effect.

5.2. Other explanations for the differences between model prices and market prices

We do not adjust for dividends, even though many of the underlying securities in the data are dividend paying. Merton (1973) presented an adjustment to the Black–Scholes formula for underlying

[9] We also tried using a percentage difference, $\frac{Call_{Market} - Call_{B-S}}{Call_{B-S}} \times 100$, as the dependent variable, but this function gets unrealistically large for options with Black–Scholes prices close to zero.

Table 3
Macbeth and Merville regressions. Estimated model is $Y_{jkt} = \alpha_1 + \beta_1 M'_{jkt} + \beta_2 M'^2_{jkt} + \beta_3 \tau_{jkt} + \beta_4 \tau^2_{jkt} + \beta_5 Spread_{jkt} + \beta_6 Spread^2_{jkt} + \varepsilon_{jkt}$ where Y_{jkt} is the relative difference between the market price and the Black–Scholes model price for an option j, on security k at time t. M'_{jkt} is a measure of moneyness, τ_{jkt} is the time to expiration and $Spread_{jkt}$ is a measure of the size of the bid-ask spread.

	No restrictions	Restrict M'	Restrict time to expiration	Restrict bid-ask spread	Restrict F
Moneyness	−0.0107***	−0.00771***	−0.0122***	−0.00959***	−0.00797***
Moneyness squared	0.000504***	−0.00330***	0.000588***	0.000450***	0.000334***
Time to expiration	0.0359***	0.0368***	0.0342***	0.0372***	0.0350***
TTE squared	0.00127***	0.00143***	0.00200***	0.000215***	0.00132***
Spread	−0.0306***	−0.0302**	−0.0410**	−0.0129***	−0.00445***
Spread squared	0.0279**	0.0307**	0.0398**	0.0245***	0.00271***
Constant	0.00507***	0.00379***	0.00693***	0.00314***	0.00221***
Observations	16,016,475	13,847,564	10,824,043	14,370,826	15,016,942

** 5% level.
*** 1% level.

Table 4
Loss function regressions. Estimated model is $f_{jkt} = \alpha_1 + \beta_1 M'_{jkt} + \beta_2 M'^2_{jkt} + \beta_3 \tau_{jkt} + \beta_4 \tau^2_{jkt} + \beta_5 Spread_{jkt} + \beta_6 Spread^2_{jkt} + \varepsilon_{jkt}$ where f_{jkt} is the relative difference between the market price and the Black–Scholes model price for an option j, on security k at time t. M'_{jkt} is a measure of moneyness, τ_{jkt} is the time to expiration and $Spread_{jkt}$ is a measure of the size of the bid-ask spread.

	No restrictions	Restrict M'	Restrict time to expiration	Restrict bid-ask spread
Moneyness	−1.260***	−2.648***	−0.646***	−0.933***
Moneyness squared	0.0732***	1.832*	0.0349***	0.0587***
Time to expiration	8.908***	8.515***	10.91***	2.971***
TTE squared	−3.003***	−2.915***	−3.821***	−0.961***
Spread	6.515***	5.433**	12.62***	−0.969
Spread squared	9.674***	9.724***	12.80***	26.51***
Constant	−3.040***	−2.701***	−5.404***	−0.584***
Observations	16,016,475	13,841,834	10,824,043	14,254,017

* 10% level.
** 5% level.
*** 1% level.

securities that pay dividends. In his model, option price is decreasing in dividend yield. Not accounting for this makes the average Y in our analysis larger than it should be if an adjustment for dividends were made, owing to the fact that the model prices on average exceed the market price.

In addition to the omission of the dividend adjustment, we believe there are certain types of options which are predisposed to low quality solutions, as measured by the size of $f_{jkt} = \frac{1}{(C^*)^2}(C^* - C(\sigma, r))^2$. We run the following regression for several specifications:

$$f_{jkt} = \alpha_1 + \beta_1 M'_{jkt} + \beta_2 M'^2_{jkt} + \beta_3 \tau_{jkt} + \beta_4 \tau^2_{jkt} + \beta_5 Spread_{jkt} + \beta_6 Spread^2_{jkt} + \varepsilon_{jkt} \tag{22}$$

and the results are presented in Table 4. In every specification, the coefficients on spread and spread squared are positive, indicating that as the spread gets larger, f gets larger. We believe this is because a large spread indicates illiquidity and mispricing at the bid price, ask price or both, and as a result Black–Scholes does not price these options well.

The magnitude and significance of the regression coefficients show that there are systematic ways in which solution quality is related to covariates in Eq. (21). Looking back at Table 3, the only significant difference between the "no restrictions" and "restrict F" specifications is the coefficients on spread and spread squared, which become economically smaller in the restricted specification.

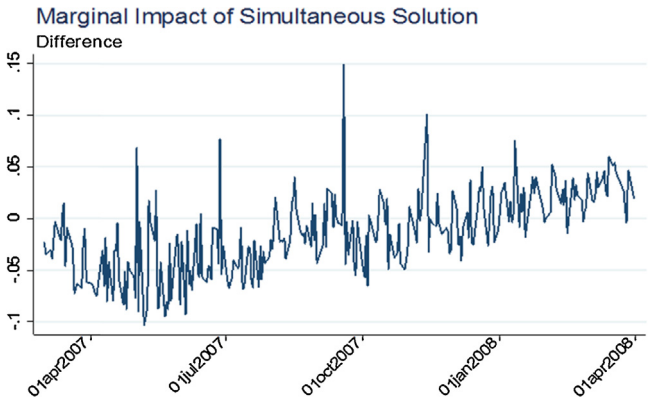

Fig. 5. IV(Fixed-r)–IV(Simultaneous r). As can be seen in Figs. 3 and 4, both measures of implied volatility are increasing over the sample period. The implied volatility calculated with a fixed risk-free rate, however, is increasing faster, so the difference increases between March 2007 and March 2008.

6. Marginal effect of allowing risk-free interest rate to vary

Finding the simultaneous solution and making the at-the-money adjustment yields more information. We explore whether or not this additional information is useful. The following section compares the new and old information sets and applies the new information to several finance problems.

Fig. 5 plots the difference between implied volatility calculated using a fixed r and the same quantity calculated with r allowed to vary for SPX options. The difference increases over the sample period, indicating that the additional information becomes more important as the sample period progresses.

Fig. 6 presents the cross correlation function between implied volatility calculated using a fixed r and the same quantity calculated with r allowed to vary. As expected, they are positively correlated across all leads and lags at about 50%, with a peak at the contemporaneous correlation of over 50%.

6.1. Volatility smile

In practice, implied volatility is different across options with different strike prices. We explore the impact of allowing the risk-free rate to vary by comparing the volatility smile in the simultaneous

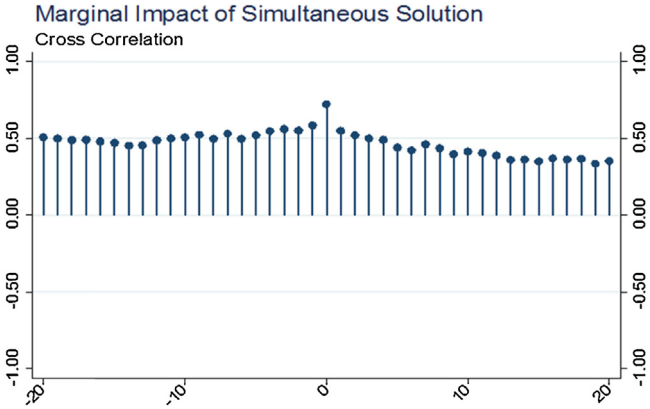

Fig. 6. Cross correlation. The cross correlation is positive across all leads and lags, with a peak at the contemporaneous correlation.

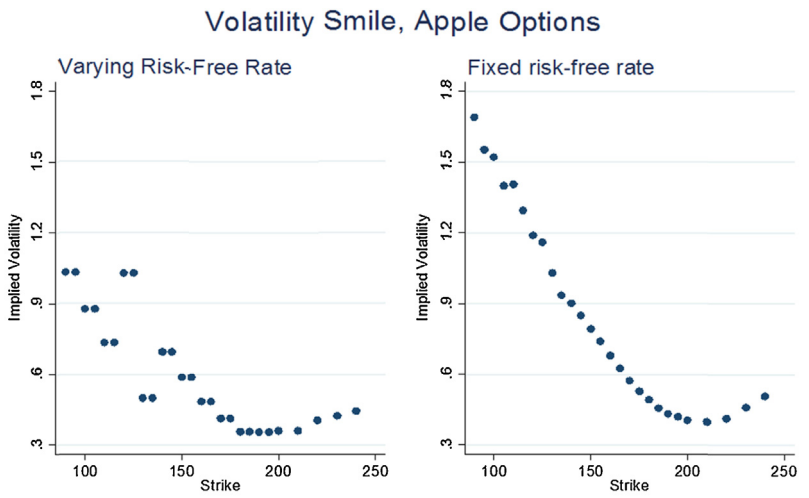

Fig. 7. Volatility smile comparison. The volatility smile looks similar when using the simultaneous solution and the fixed risk-free rate solution. The main difference is for far in-the-money options and this is the result of restricting implied volatility to values between zero and one.

risk-free rate and fixed risk-free rate models. Fig. 7 shows this comparison for AAPL options. Allowing r to vary does not change the pattern much, as the volatility smile still exists.

Fig. 8 is a similar plot for GOOG options. Although the volatility skew is still apparent in the graph on the left, the pattern is less clearly defined when r is allowed to vary.

The volatility smile looks different when using the simultaneous solution because there is a balancing effect between the risk-free rate and implied volatility. In Fig. 9, there is a pattern for the implied

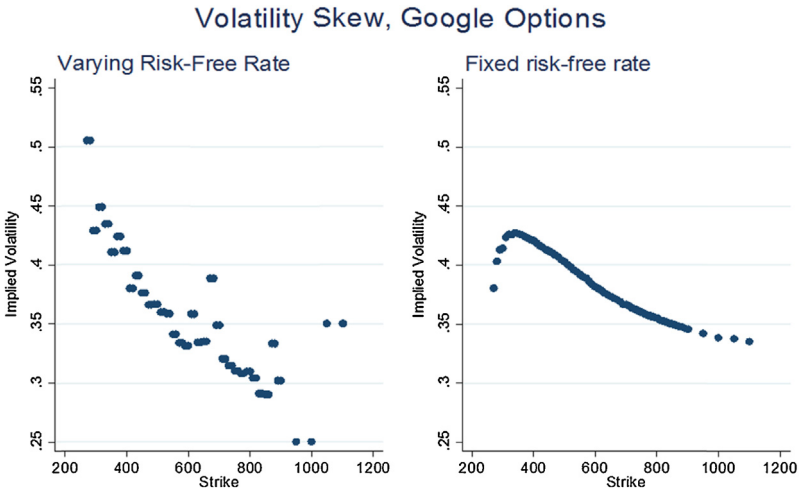

Fig. 8. Volatility skew comparison. The volatility skew looks steeper using the simultaneous solution. Unlike the example with AAPL options, this cannot be explained by our restrictions on values for implied volatility.

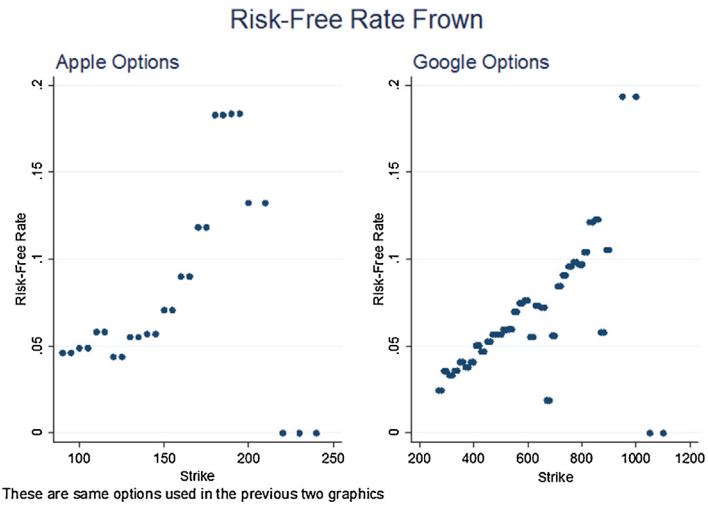

Fig. 9. Balancing risk-free rate. Unlike implied volatility, which is lowest at-the-money, implied risk-free rate is highest at-the-money.

risk-free rate across strikes that is the inverse of the pattern for implied volatility. This balancing effect, however, is not enough to get rid of the volatility smile, so the problem remains unresolved.

6.2. *Difference between market price and Black–Scholes price*

The evolution of the absolute value of the difference between the market price and the Black–Scholes price is presented in Fig. 10. In both cases, we obtain these prices by making the at-the-money adjustment and re-pricing the options. The simultaneous solution is generally better at matching market prices than the model with a fixed risk-free rate. The varying risk-free rate model at a minimum better fits the data and possibly provides better estimates of implied volatility.

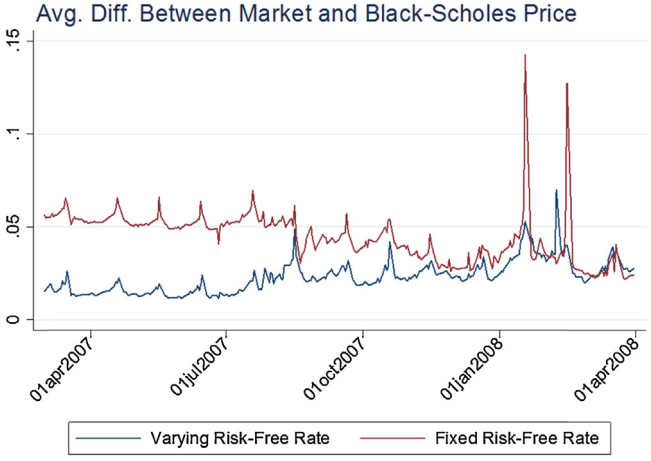

Fig. 10. Accuracy of repricing. For the majority of the sample period, the simultaneous risk-free rate model does a better job of matching model prices to market prices. This changes at the end of the sample, when the fixed risk-free rate model better matches market prices.

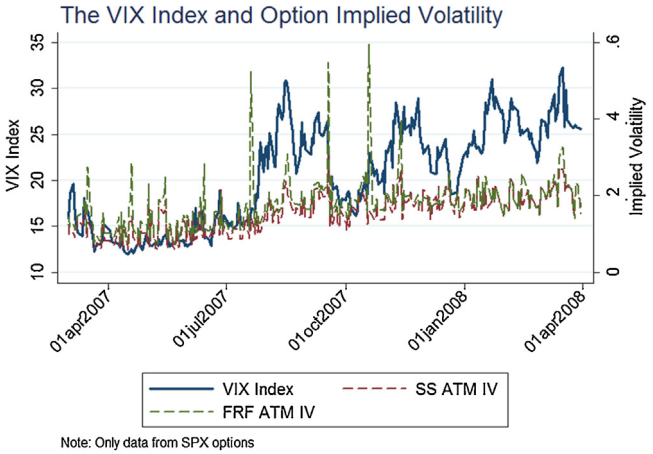

Fig. 11. Comparing volatility measures. SS stands for simultaneous solution and FRF stands for fixed risk-free rate. These variables all increase over the sample period, but at varying rates.

6.3. Predicting volatility

In order to compare the accuracy of the VIX prediction under alternative risk-free rate assumptions, we estimate in-sample and out-of-sample forecasts of the VIX using implied volatility on SPX options. We only use SPX options because the VIX uses these exclusively to calculate the volatility of the S&P 500 index. Fig. 11 shows the VIX and the at-the-money implied volatilities for varying and fixed risk-free solutions. In each specification, implied volatility is the average across all SPX options traded on that day.

The econometric model is obtained via time series identification as:

$$VIX_t = \alpha + \beta_1 Vix_{t-1} + \beta_2 Vix_{t-2} + \gamma_0 ATMVolatility_t + \gamma_1 ATMVolatility_{t-1} + \gamma_2 ATMVolatility_{t-2} + \epsilon_t \quad (23)$$

where ATM Volatility is the at-the-money implied volatility for the two alternative cases. The in-sample predictions refer to the forecast error ϵ_t while the out-of-sample predictions are obtained by calibrating the model on data from March 2007 to October 2007 and calculating the dynamic predictions for the remaining periods. In both cases, we use the Diebold and Mariano (1995) test to determine which forecast is more accurate. The in-sample and out-of-sample results, presented in Table 5, are both evidence in favor of the joint implied volatility and risk-free rate model. We tested several alternatives, including static forecasts, and the results are similar and robust.[10]

For the purposes of forecasting the VIX, the traditional implied volatility model is inferior to the joint implied volatility and implied risk-free rate model we present in this paper.[11]

7. Potential trading strategies

7.1. Re-pricing of options strategy

The first trading strategy we examine is based on the re-pricing of options after the at-the-money adjustment. We believe the model price should be a more accurate measure of long-run value than

[10] We also tried removing the contemporaneous at-the-money implied volatility term, and our model still yields superior out-of-sample forecasts.

[11] The regressions and additional models will be posted on a website with additional materials and are available upon request.

Table 5
Diebold–Mariano tests. For both in-sample and out-of-sample forecasts, we reject the null hypothesis of no difference at the 1% level.

Series	MSE (mean squared error) over 262 obs.
In-sample	
At the money implied volatility (varying risk-free rate)	1.825
At the money implied volatility (fixed risk-free rate)	2.585
Difference	−0.7594
Diebold–Mariano $S(1) = -2.899$ (p value = 0.0037)	Reject null of equal forecasts in favor of joint implied volatility and risk-free rate
Series	MSE (mean squared error) over 99 obs.
Out-of-sample	
At the money implied volatility (varying risk-free rate)	12.77
At the money implied volatility (fixed risk-free rate)	47.39
Difference	−34.62
Diebold–Mariano $S(1) = -4.889$ (p value < 0.0001)	Reject null of equal forecasts in favor of joint implied volatility and risk-free rate

the market price. Given this, if there is a discrepancy between market and model prices, we defer to the Black–Scholes price.

We define a strategy to exploit the prices differences in the data:

- Drop all observations for which the model price differs from the market price by more than 20%. In our opinion, these options are mispriced and there is probably an unusual event or a data anomaly that can explain these differences.[12]
- Drop observations with fewer than 90 days to expiration, because options close to expiration can have unusual price fluctuations.
- Drop observations where the difference between market and model prices is zero.
- For the remaining options: If the market price exceeds the Black–Scholes price, write that call option and hedge this position by buying the underlying security. If the market price is lower than the Black–Scholes price, buy the option.

It is possible to hedge the long call position by shorting the stock, but we avoid shorting for simplicity. In addition, any trading strategy will incur transaction costs, but again for simplicity these ignored. Finally, options are bought at the bid price and sold at the ask price, but using the bid or ask instead of the midpoint does not make a substantive difference, as is discussed in Appendix C.

We implement the trading strategy for March 1st, 2007. There were 502 SPX options traded that day which met the selection criteria. The return for each option is calculated and added to the return on the stock if the position is hedged. We then calculate the average return for the strategy.[13] We excluded options expiring on March 21st 2008, because we could not get a price for the S&P 500 index on that day and thus could not calculate the exercise value of those options. The average position return is about 38%, and the standard deviation is about 39%, both of which are economically large. Implementing this same strategy on November 30th, 2007 yields an average return of about −80%, with a standard deviation of about 29%, suggesting that the strategy's performance is dependent on the day selected. These returns should be measured on a risk-adjusted basis. A popular measure for this is the Sharpe ratio: $SharpeRatio = \frac{\overline{r_p} - r_f}{\sigma_p}$ where $\overline{r_p}$ is the expected portfolio return, r_f is the risk-free rate of return and σ_p is the standard deviation of portfolio returns. For the March 1st 2007 data, the risk-free rate as measured by 3-month Treasury bills is about 5%, making the Sharpe ratio less than one.

[12] We considered restricting F, but after dropping options whose model prices and market prices differ by more than 20%, almost every observation already has a small F.
[13] We do not weight each position by the degree of mispricing, but that is another possible strategy.

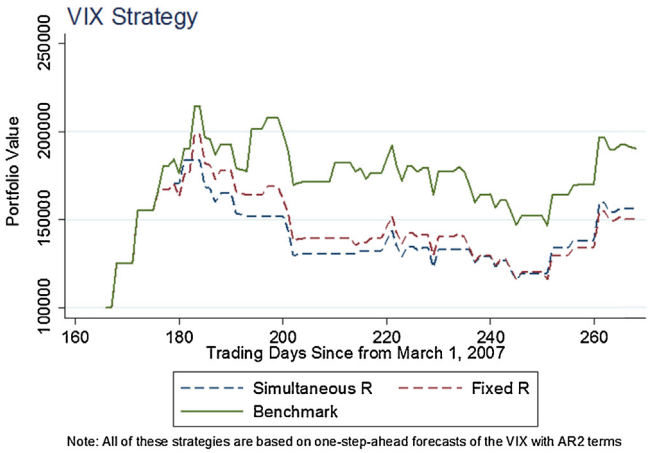

Fig. 12. Comparing trading strategies. For the first 20 trading days, all of the portfolios have the same performance, but in later periods, the strategies differentiate themselves.

The standard deviation would be lower if we could hedge the long call positions, as many positions go to zero as they expire out-of-the money.

We implement the same strategy, but instead of re-pricing the calls with the at-the-money implied volatility from the simultaneous solution, we use the at-the-money implied volatility calculated with a fixed risk-free rate. We evaluate the marginal impact of the additional information from a simultaneous solution by comparing the results of each trading strategy. For the March 1st 2007 data, there is an average position return of about 24%, and a standard deviation of about 7.5%. On a risk-adjusted basis, this is superior to the performance of the simultaneous solution strategy, but as mentioned above, these results are sensitive to the days chosen for analysis.

7.2. VIX prediction strategy

Another potential trading strategy is based on predicting the VIX index and making trades based on its expected future value. For this strategy, we use the econometric model in Eq. (23) to obtain one-step-ahead forecasts of the VIX index and define a strategy as follows:

- A position initiated during a given trading day must be closed before the end of that trading day. We assume positions held overnight do not collect interest.
- If the next period predicted VIX is higher than the current level of the VIX, put the entire portfolio into shares of a product that closely tracks the VIX and sell them at the end of the next trading day.[14] We assume there is no tracking error between these products and the index itself.
- If the next period predicted VIX is lower than the current level of the VIX, keep the entire portfolio in cash for the next trading day.

While it would be possible to short the VIX when the next period predicted value is lower than its current value, this is avoided for simplicity. In addition, transaction costs are ignored.

To implement the strategy, the model described in Eq. (23) is calibrated using the first 160 trading days in our sample. The VIX forecasts after the first 160 trading days are iterative, so the model is re-estimated for every forecast using all data points before the one to be predicted. The hypothetical portfolio starts with $100,000. A benchmark strategy includes only the first and second lag of the VIX in Eq. (23), while the competitor strategies include measures of model based at-the-money implied

[14] There are a variety of exchange traded products designed to track VIX futures including NYSE ARCA: CVOL and VXX.

volatility. Fig. 12 shows the value of this hypothetical portfolio over time. The simultaneous and fixed risk-free rate solutions yield alternative relative performances in the sample period. In particular, the simultaneous risk-free rate model dominates the fixed risk-free model in the later periods but the reverse occurs in the earlier periods.

8. Summary and conclusions

We implement an algorithm that solves systems of Black–Scholes formulas for implied volatility and implied risk-free rate. For each underlying security, point estimates of at-the-money implied volatility and implied risk-free rate are calculated using a seemingly unrelated regressions (SUR) model. These point estimates are used to re-price the options using the Black–Scholes formula. We examine the impact of moneyness, time to expiration and size of the bid-ask spread on the difference between market prices and model-based Black–Scholes prices.

We find that across most of our specified restrictions, the coefficient on moneyness is negative, while the coefficient on moneyness squared is positive. This is explained by volatility skew and the volatility smile. The coefficients on time to expiration and time to expiration squared are both positive across all specifications, indicating that as options get far from expiration, their model prices diverge from market prices.

We examine the marginal impact of allowing the risk-free rate to vary in terms of the volatility smile and the accuracy of market volatility predictions. The difference between implied volatility calculated using a fixed r and the same quantity calculated with r allowed to vary increases over the sample period, indicating that additional information becomes more important as the sample period progresses. The varying risk-free rate model better fits the market data and potentially provides better estimates of implied volatility, as it is better able to minimize the difference between model-based Black–Scholes prices and market prices. For the purposes of forecasting the VIX, our model is superior to the traditional implied volatility model. Finally, we outline two potential trading strategies based on our analysis. The first uses the discrepancy between Black–Scholes prices and model prices. We compare this strategy's risk-adjusted return to a similar strategy using implied volatility calculated with a fixed risk-free rate. The other strategy is based on predicting the VIX index. In both cases, the simultaneous and fixed risk-free rate models yield alternative relative performances in the sample period.

There are several avenues for future research that we think are fruitful. We believe it is important to improve the accuracy of the at-the-money adjustment by using more nonlinear terms in the SUR model. Expanding the sample period to more recent years is needed to test the robustness of our results. Finally, an information index based on implied volatility calculated with a variable risk-free rate could be used in parallel to the VIX as a measure of market volatility. In general, we believe options prices present important information on expectations that is essential for market participants and policy makers.

Appendix A. The need for an optimization routine

This section justifies the need for an optimization routine to minimize: $F(\sigma, r) = \frac{1}{(C_1^*)^2}(C_1^* - C_1(\sigma, r))^2 + \frac{1}{(C_2^*)^2}(C_2^* - C_2(\sigma, r))^2$ Given that the Black–Scholes formula is monotonically increasing in both σ and r, picking starting values and moving against the gradient of F until a minimum is reached would be more straightforward than the optimization routine we propose. This, however, will not always work, as is shown in the analysis of a pair of Agilent Technologies (NYSE:A) call options. On 3/1/2007, the stock was trading at \$31.44 and both options were 16 days from expiration. The calls had strike prices of \$27.50 and \$30.00 and had bid-ask midpoints of \$4.08 and \$1.73.

Fig. A.1 is a plot of F for this pair of options. Looking at the plot of F, it is hard to see the global minimum, so we make small values of F appear large on the vertical axis by plotting $-log(F)$ in Fig. A.2. This plot shows two things: (1) F is not monotonically increasing in σ and r and (2) there are several local minima surrounding the global minimum.

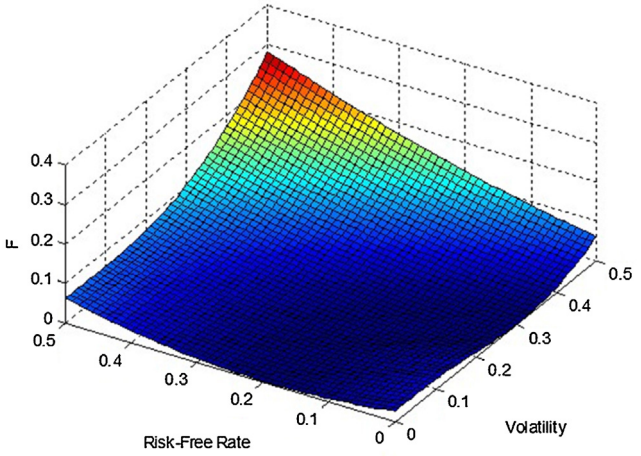

Fig. A.1. Plot of F.

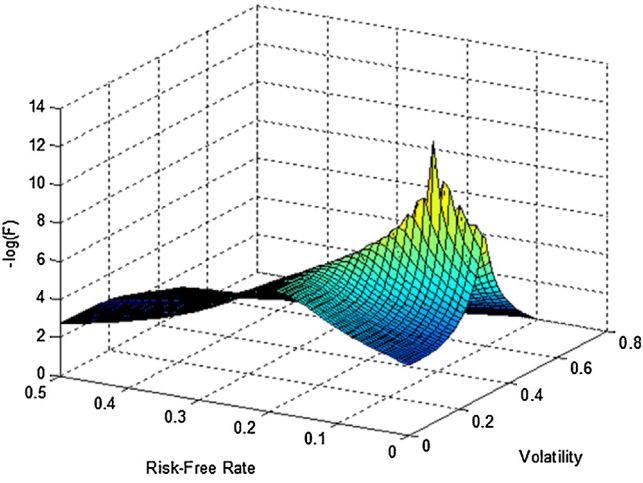

Fig. A.2. Plot of $-\log(F)$.

To dig deeper into this issue, F is divided into two pieces: $f_1 = \frac{1}{(C_1^*)^2}(C_1^* - C_1(\sigma, r))^2$ and $f_2 = \frac{1}{(C_2^*)^2}(C_2^* - C_2(\sigma, r))^2$. Plots of these functions have the same issue as the plot of F: the minimum is hard to see. To resolve this issue, the plots of $-log(f_i)$ are shown below in Figs. A.3 and A.4. The lack of monotonicity of f_i in σ and r is apparent in both figures. There are multiple local minima in each of these plots because as $C_i - C(\sigma, r)$ gets close to zero, it is possible to increase σ by a small amount and decrease r by a small amount and keep $C(\sigma, r)$ about the same.

To finalize this analysis, in Fig. A.5 we superimpose $-log(f_1) + 5$ on $-log(f_2)$. We add 5 to $-log(f_1)$ to make it easier to distinguish the two functions. The figure mostly in yellow is $-log(f_1) + 5$ while the figure mostly in blue is $-log(f_2)$. In the region near $\sigma = 0.3$ and $r = 0.1$, both of the functions have multiple local minima, which is explains why F has multiple local minima in this region as well.

Given the existence of multiple local minima and the lack of monotonicity in σ and r, an optimization routine is needed to minimize F.

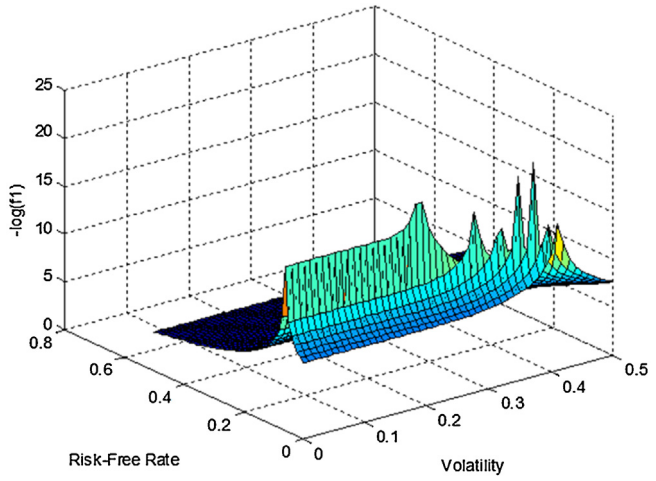

Fig. A.3. Plot of $-\log(f1)$.

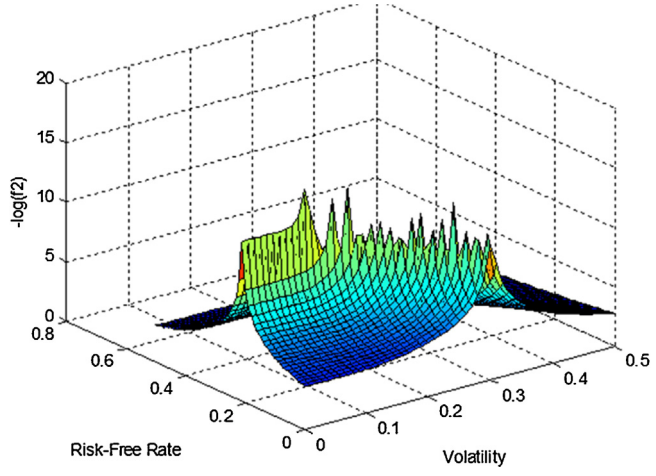

Fig. A.4. Plot of $-\log(f2)$.

Appendix B. Alternative optimization methods

B.1. Alternative optimization algorithms

While our primary algorithm for finding σ and r uses the *fmincon* function built into MATLAB's optimization toolbox, there are other algorithms that can minimize: $F(\sigma, r) = \frac{1}{(C_1^*)^2}(C_1^* - C_1(\sigma, r))^2 + \frac{1}{(C_2^*)^2}(C_2^* - C_2(\sigma, r))^2$.

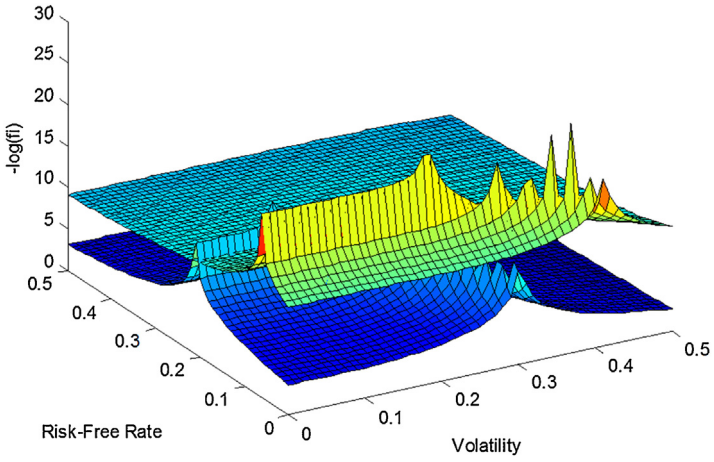

Fig. A.5. Plot of $-\log(f1)+5$ superimposed on $-\log(f2)$.

The optimization toolbox also contains a function designed to solve nonlinear least-squares problems called *lsqnonlin*. The input for *lsqnonlin* is a vector which contains the square roots of the functions to be minimized. A potential input for *lsqnonlin* to minimize *F* as described above is:

$$\mathrm{F}(\sigma, \mathrm{r}) = \begin{bmatrix} \dfrac{C_1^* - C_1(\sigma, r)}{C_1^*} \\ \dfrac{C_2^* - C_2(\sigma, r)}{C_2^*} \end{bmatrix} \tag{24}$$

To minimize the sum of squares of the functions contained in **F**, the algorithm starts at particular values for σ and r and approximates the Jacobian of the vector **F** using finite differences. Then, it solves a linearized least-squares problem to determine how to adjust σ and r. A trust-region method is used to control the size of these changes at each step: if the proposed change in σ and r gets the sum of squares in F closer to zero, it is used. Otherwise, σ and r are perturbed by a small amount and the algorithm solves the linearized least-squares problem at the new values for σ and r.

This process is repeated until a sufficiently small sum of squares has been achieved, or the maximum number of allowed iterations is reached. A limit is set on the maximum number of iterations because for some pairs of calls, there are no values of σ and r that yield a sufficiently small sum of squares.

The *lsqnonlin* algorithm is faster than both the interior point and SQP algorithms built into the *fmincon* function. This is because *lsqnonlin* gains efficiency from that fact that the problem is known to be a minimization of squares, which allows the algorithm to make assumptions that *fmincon* cannot.

Table B.1 compares the performance of *lsqnonlin* to both the interior point and SQP algorithms built in *fmincon* using options from March 2007. These averages are based on using the bid-ask midpoint as the representative price for each option.

Lsqnonlin is many times faster than SQP and is more than three times as fast as interior point. The average *F* is smallest for interior point, but *lsqnonlin* is not far behind. Finally, all three algorithms fail to find solutions for the same pairs of options. This leads us to believe that the inability to find a

Tabel B.1
Speed and accuracy comparisons. *lsqnonlin* is the fastest algorithm, but it is not as accurate as interior point.

	Algorithm:	Average seconds per pair	Average *F*	%Solutions found
Fsolve	Lsqnonlin	0.0269	0.0123	96.51%
	SQP	0.1744	0.1483	96.51%
	Interior point	0.0831	0.0075	96.51%

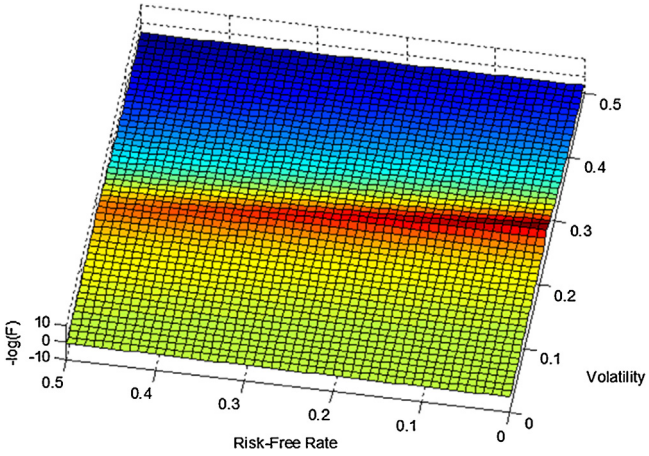

Fig. B.1. Plot of $-\log(F)$.

solution is not an algorithm-specific problem, but rather an issue where some pairs of options are so mispriced that it is impossible to pick a σ and r pair to make F sufficiently small.

B.2. Alternative function specifications

One of the issues with minimizing F is the somewhat random nature by which our optimization routine selects r. The randomness arises because there is usually a large range of r values for which a chosen value of σ makes F small. The issue of randomness does not apply to σ as there is usually a smaller range of σ values for which a chosen r value makes F small.

Fig. B.1 below shows the surface of $-log(F)$ for a pair of Agilent Technologies (NYSE:A) call options, with the red areas indicating where F is close to zero. This shows that the range of possible σ values for which F is small is between 0.2 and 0.3, while the same range for r values is between 0 and 0.4. The range for r is 4 times as large as the range for σ. Fig. B.2 presents a different view of the same

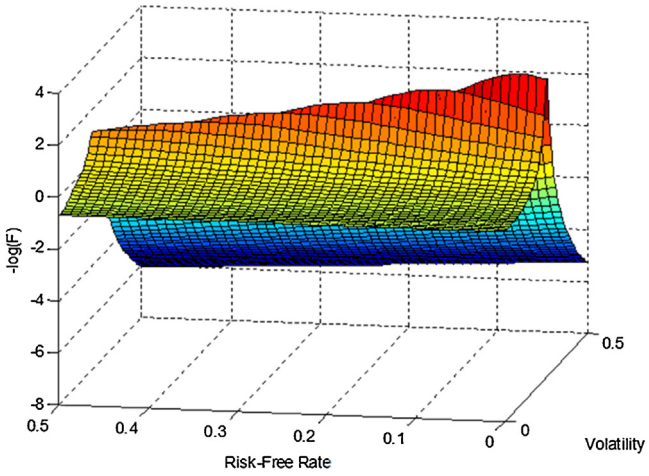

Fig. B.2. Plot of $-\log(F)$ rotated.

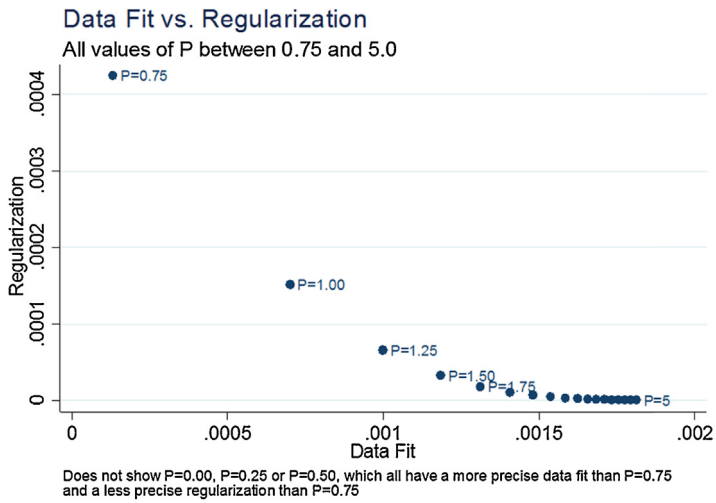

Fig. B.3. Determining appropriate size for *P*. Setting $P = 1.25$ achieves the right balance between minimizing the regularization and data fit terms.

object, which shows several local minima along the red ridge which an optimization routine might accidentally pick as the best value for minimizing *F*.

If the optimization routine selects one of the local minima, the chosen σ is not far from the σ at the global minimum, but the value of r could be drastically different. To address the issue of r's randomness, we can add a term to *F* as follows:

$$F(\sigma, r) = \frac{1}{(C_1^*)^2}(C_1^* - C_1(\sigma, r))^2 + \frac{1}{(C_2^*)^2}(C_2^* - C_2(\sigma, r))^2 + P(r - r^*)^2 \tag{25}$$

where r^* is the benchmark value for r and P is the degree by which values of r are penalized as they deviate from r^*. We use the annualized three-month Treasury bill rate for r^*. The justification for this is that it is reasonable to believe that r should be close to r^*, but this must be balanced against minimizing the difference between the market price and the Black–Scholes price for each pair of calls. To decide on a value for P that balances these demands, the average size of $\frac{1}{(C_1^*)^2}(C_1^* - C_1(\sigma, r))^2 + \frac{1}{(C_2^*)^2}(C_2^* - C_2(\sigma, r))^2$, the data fit, is plotted against $(r - r^*)^2$, the regularization term, for all values of P between 0.75 and 5.00 in increments of 0.25 using data from March 2007. The result is shown in Fig. B.3.

Tabel B.2
Speed and accuracy comparisons. *lsqnonlin* with $P = 1.25$ is the fastest, but it is still not as accurate as interior point, even adjusting for the fact that *F* now has an extra term.

	Algorithm:	Average seconds per pair	Average *F*
Lsqnonlin	$P = 1.25$	0.0246	0.0213
	$P = 0.00$	0.0269	0.0123
Fsolve	SQP	0.1744	0.1483
	Interior point	0.0831	0.0075

Tabel B.3
Impact of adding regularization. by regularizing r with *lsqnonlin*, its average size nearly doubles.

Algorithm:	P	Average r	Average $(r\text{-}r^*)^2$
Lsqnonlin	1.25	0.1556	0.0068
	0.00	0.0847	0.0288

Based on Fig. B.3, $P = 1.25$ is near the inflection point of the curve created by the points, making it the best choice for balancing the data fit and regularization terms. Table B.2 compares the performance of the *lsqnonlin* algorithm with the input in Eq. (26) against several benchmarks:

$$\mathrm{F_2}(\sigma, \mathrm{r}) = \begin{bmatrix} \dfrac{C_1^* - C_1(\sigma, r)}{C_1^*} \\ \dfrac{C_2^* - C_2(\sigma, r)}{C_2^*} \\ \sqrt{1.25}(r - r^*) \end{bmatrix} \tag{26}$$

Setting $P = 1.25$ actually makes the algorithm faster. This could mean that the global minimum is normally near r^*, so nudging r towards r^* speeds up the optimization routine. The average value of F is larger, but this is no surprise, given that an additional term has been added. If we remove the impact of adding $1.25(r - r^*)^2$ to F, the average is 0.0145, which is slightly bigger than the average for $P = 0.00$.

The next thing to consider is the impact of setting $P = 1.25$ on r. Table B.3 shows how this affects the average r and the average squared difference between r and the annualized three-month Treasury bill rate.

As expected, setting $P = 1.25$ gets r closer to the Treasury bill rate, but at the expense of almost doubling the average size of r. We conclude that any regularization of r is going to lose important information, so it is better to solve for r without any restrictions.

Appendix C. Using the bid/ask prices instead of the bid-ask midpoint

C.1. Summary statistics

Throughout the entire paper, the bid-ask midpoint is used as the representative price for each option. While this is an easy way to resolve the fact that a bid-ask spread exists, the validity of this technique is better determined by looking at the paper's results using the bid and ask prices themselves. Table C.1 presents summary statistics for the average option-implied risk-free rate and option-implied volatility. The implied volatility specification fixes the risk-free rate at the yield of the three-month Treasury bill.

Both option-implied risk-free rates and volatilities are on average higher for the ask specification than the midpoint specification. This is because the Black–Scholes price is increasing in σ and r, and the ask price is at least as large, if not larger, than the midpoint price. This means the average σ and r must be at least as large in the ask specification as they are in the midpoint specification.

Tabel C.1
Summary statistics.

	Specification			
	Bid	Ask	Midpoint	Implied volatility
Average risk-free rate	0.0876	0.1117	0.0932	
Average ATM risk-free rate	0.1207	0.1577	0.1237	
Average implied volatility	0.3332	0.3538	0.3536	0.3953
Average ATM implied volatility	0.3520	0.3835	0.3767	0.4110

Tabel C.2
Macbeth and Merville regressions.

	Specification			
	Bid	Ask	Midpoint	Implied volatility
Moneyness	0.206***	0.498***	0.392***	0.0328***
Time to expiration	−3.178***	−3.368***	−3.091***	−5.024***
Spread	−2.294***	−2.330***	−2.199***	−4.318***
Constant	−0.396***	−0.415***	−0.379***	0.984***
Observations	15,650,630	15,951,644	16,016,475	16,461,704

*** $p < 0.01$.

C.2. Macbeth and Merville regression

We also want to examine the regression results using the bid and ask prices instead of the bid-ask midpoint. Table C.2 presents a simplified version of the Macbeth and Merville (1979) regressions from Section 5 for all specifications. The Implied Volatility specification uses the midpoint price, while fixing the risk-free rate at the yield on 3-month Treasury Bills. The same restrictions apply to these regressions that were described in Section 4.

The largest deviation among specifications occurs in the coefficient on moneyness, which is more than twice as large for the ask specification as it is for the bid specification. Generally speaking, however, the results are similar, so we believe it is safe to use the bid-ask midpoint, instead of the bid and ask prices, without losing important information.

References

Becker, R., Clements, A. C., & White, S. (2007). Does implied volatility provide any information beyond that captured in model-based volatility forecasts? *Journal of Banking and Finance*, *31*, 2535–2549.

Black, Fischer, & Scholes, Myron. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, *81*(3), 637.

Bliss, Robert R., & Panigirtzoglou, N. (2004). Option-implied risk aversion estimates. *The Journal of Finance*, *59*(1), 407–446.

Canina, L., & Figlewski, S. (1993). The informational content of implied volatility. *Review of Financial Studies*, *6*, 659–681.

CBOE Volatility Index – VIX. (2009). *Chicago Board Options Exchange, Incorporated*.

Christensen, B. J., & Prabhala, N. R. (1998). The relation between implied and realized volatility. *Journal of Financial Economics*, *50*, 125–150.

Constantinides, G., Jackwerth, J. C., & Savov, A. (2012). *The puzzle of index option returns*. The University of Chicago, Booth School., working paper, page 9, February 2012.

Diebold, Francis, & Mariano, Roberto. (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics*, *13*(3), 253–263.

Hammoudeh, S., & McAleer, M. (2013 August). Risk management and financial derivatives: An overview. *The North American Journal of Economics and Finance*, *25*, 109–115.

Hentschel, L. (2003). Errors in implied volatility estimation. *Journal of Financial and Quantitative Analysis*, *38*, 779–810.

Jiang, G., & Tian, Y. (2005). Model-free implied volatility and its information content. *Review of Financial Studies*, *18*, 1305–1342.

Krausz, J. (1985). Option parameter analysis and market efficiency tests: A simultaneous solution approach. *Applied Economics*, *17*, 885–896.

Macbeth, J., & Merville, L. (1979). An empirical examination of the Black–Scholes call option pricing model. *The Journal of Finance*, *34*, 1173–1186.

Merton, R. (1973). Theory of rational options pricing. *Bell Journal of Economics and Management Science*, *4*, 141–183.

O'Brien, T. J., & Kennedy, W. F. (1982). Simultaneous option and stock prices: Another look at the Black–Scholes model. *The Financial Review*, *17*(4), 219–227.

Swilder, S. (1986). Simultaneous option prices and an implied risk-free rate of interest: A test of the Black–Scholes Models. *Journal of Economics and Business*, *38*(2), 155–164.