



Institutional Members: CEPR, NBER and Università Bocconi

WORKING PAPER SERIES

**Learning to smile: Can rational learning
explain predictable dynamics in the implied
volatility surface?**

Alejandro Bernales, Massimo Guidolin

Working Paper n. 565

This Version: December, 2015

IGIER – Università Bocconi, Via Guglielmo Röntgen 1, 20136 Milano –Italy
<http://www.igier.unibocconi.it>

The opinions expressed in the working papers are those of the authors alone, and not those of the Institute, which takes non institutional policy position, nor those of CEPR, NBER or Università Bocconi.

Learning to smile: Can rational learning explain predictable dynamics in the implied volatility surface?*

Alejandro Bernales

The Universidad de Chile (Centro de Economía Aplicada and Centro de Finanzas)

Massimo Guidolin

Bocconi University, IGIER, CAREFIN

Abstract

We develop a general equilibrium asset pricing model under incomplete information and rational learning in order to understand the unexplained predictability of option prices. In our model, the fundamental dividend growth rate is unknown and subject to breaks. Immediately after a break, there is insufficient information to price option contracts accurately. However, as new information arrives, a representative Bayesian agent recursively learns about the parameters of the process followed by fundamentals. We show that learning makes beliefs time-varying and generates predictability patterns across option contracts with different strike prices and maturities; as a result, the implied movements in the implied volatility surface resemble those observed empirically.

Keywords: option pricing, rational learning, Bayesian updating, implied volatility, predictability.

JEL classification: G12, D83.

* The authors would like to thank one anonymous referee, Michael Brennan, Mario Cerrato Gaetano Gaballo, Christian Hellwig, Stuart Hyde, Jean-Stéphane Mésonnier, Bruce Lehmann, and Alex Taylor for their comments on earlier versions of the paper. Additionally, we would like to thank seminar/session participants at Banque de France, Central Bank of Chile, the 2012 Financial Management Association Conference in Atlanta, the 2012 French Economic Association Conference in Paris, the 2013 Market Microstructure and Nonlinear Dynamics workshop in Evry, and Rouen Business School. The authors are grateful for the computational resources proportioned by the computing grids MACE and MAN2 at the University of Manchester and especially for the useful help from Simon Hood and Michael Croucher in running algorithms in the grids (MACE is part of the Mechanical, Aerospace & Civil Engineering School, while MAN2 is part of the British North Western Grid.) Alejandro Bernales acknowledges financial support from Fondecyt project 11140628 and the Institute for Research in Market Imperfections and Public Policy (ICM IS130002).

1 Introduction

Contrary to the constant volatility assumption of the Black and Scholes' (1973) model (henceforth, BS), the volatilities implicit in option contracts change over time. Moreover, it is well known that at least from a statistical perspective, strong predictability patterns exist in implied volatilities and option prices (e.g., Harvey and Whaley, 1992; Heston and Nandi, 2000; Gonçalves and Guidolin, 2006; Konstantinidi, Skiadopoulou, and Tzagkaraki, 2008; and Christoffersen, Heston, and Jacobs, 2009). Additionally, and also in sharp contrast with BS' assumptions and pricing results, the volatilities implicit in option contracts written on the same underlying asset systematically differ across strike prices and expiration dates. These cross-sectional implied volatilities are known as the implied volatility surface (IVS) (e.g., Rubinstein, 1994; Dumas, Fleming, and Whaley, 1998; Das and Sundaram, 1999). Historically, while BS' constant volatility assumption was initially believed to characterize market option prices reasonably well (e.g., Rubinstein, 1985), since the 1987 market crash, data have been found to be inconsistent with BS because of the presence of persistent implied volatility smiles/skews and term structures. Furthermore, and similar to the behavior of the implied volatility of a single option contract, there is evidence of predictable dynamics in the shape characteristics of the IVS (e.g., Gonçalves and Guidolin, 2006; Chalamandaris and Tsekrekos, 2010).

Despite this widespread and compelling evidence of dynamics in the volatilities implicit in traded options, there are few equilibrium pricing models based on first principles (i.e., from simple and generally accepted assumptions concerning preferences and the stochastic process of fundamentals driving asset prices), such ubiquitous patterns of dynamic predictability in the IVS.¹ The main goal of our research is to fill this gap by developing a rather standard and yet novel and powerful equilibrium model, in which the rational learning by the investors explains the predictable dynamics in option prices and in the corresponding IVS.

We develop a discrete-time endowment, Lucas-type economy in which a representative agent trades in a risk-free one-period bond, in a stock, and in a set of option contracts with

¹ A handful of exceptions are, however, discussed below. Researchers have proposed econometric models for the IVS and tested whether these may support profitable, out-of-sample trading strategies (e.g., Dumas, Fleming, and Whaley, 1998; Gonçalves and Guidolin, 2006; Fengler, 2009; Kim and Lee, 2013).

different strike prices and expiration dates. The stock pays out an infinite stream of real dividends that evolve according to a geometric random walk; however, the mean dividend growth rate, g_t , is subject to infrequent (and always observable) breaks where time periods between breaks follow a memory-less stochastic process. In a scenario in which a break takes place, the new mean dividend growth rate is drawn from a continuous univariate density $g_{t+1} \sim G(\cdot)$ defined on the support $[g_d, g_u]$. Even though breaks are observable, g_t is unknown to the agent, who recursively obtains incomplete information about the mean growth rate by observing independently distributed but noisy daily dividend realizations. The agent efficiently uses these signals following a rational Bayesian updating (learning) process.² Immediately after a break, historical information is scarce and this makes signals potentially unreliable; as a result, drastic revisions of beliefs concerning the new post-break value of g_t become likely. As long as no new stochastic breaks occur (and given their infrequent nature, this is likely), these initial large updates in beliefs gradually decline as the agent endogenously learns as new information arrives. Nevertheless, learning never disappears completely, even asymptotically, because its strength is destined to be revived after a new break hits the mean growth rate. Therefore, breaks in the mean growth rate induce two main effects on all assets. Firstly, breaks in g_t impact the stochastic evolution of future dividends, affecting the pricing of all securities directly. Secondly, breaks modify the quantity and reliability of the information that the agent has access to regarding the mean dividend growth rate g_t , hence breaks change the speed and intensity with which the investor updates her beliefs. Moreover, given that the learning process produces dynamic effects in beliefs, this process of recursive belief adjustments is responsible for a corresponding, highly nonlinear dynamic in options prices and in the associated IVS.³

Financial markets and the economy are subject to continuous changes that force investors into an ever progressing process of learning regarding fundamentals. There are numerous

² Because under rather general conditions that are satisfied under our simple set-up with observable breaks, it can be shown that the application of Bayes' rule to the learning problem is equivalent to rational updating (Bray and Kreps, 1987; Guidolin and Timmermann, 2007). Below we discuss Bayesian and rational learning as if the two terms are interchangeable.

³ As a first step, we assume that the volatility, σ , in the geometric random walk followed by dividends is constant, in order to obtain the simplest setting that allows us to observe the effects on options of learning about g_t , which is in line with Timmermann (1996, 2001). However, later we extend the model by allowing the dividend volatility to vary following a GARCH (1,1).

breaks in economic fundamentals reported in the literature, such as in the parameters of the dividend process and in real GDP growth (e.g., Bai, Lumsdaine, and Stock, 1998; Timmermann, 2001; Granger and Hyung, 2004). Breaks in fundamentals could be due to permanent technological innovations, or shifts in tax codes, monetary policy, or stock market participation, among other possibilities. However, when there are breaks in economic fundamental, agents optimally follow a rational belief updating mechanism as they need to understand the new market conditions. Therefore, we propose a simple and yet powerful model based on the interaction between rational learning and infrequent structural breaks to explain documented but not currently well understood features of the way options are priced.

Through an extensive set of simulations of price options with strikes and maturities determined according to the same (listing and delisting) rules that are followed in established option markets, we show that Bayesian learning induces dynamic patterns in option prices and implied volatilities (henceforth, IVs) that are consistent with what is reported in the empirical literature. We find that learning produces different dynamics in the IVs across strike prices and time-to-maturities, and thus induces movements in the shape of the IVS. We also show that learning produces serial correlation and volatility clustering in IVs, as well as in (measures of) the slope and curvature of the IVS.⁴ For instance, we report strong predictability patterns on the slope and curvatures on the moneyness and maturity dimensions measured by ARCH LM tests (both with one and three lags). This means that when levels, slope or convexity of the IVS become variable over time, this instability tends to persist over time. Nevertheless, ARCH effects are weaker in the case of the slope and curvature indices measured with respect to moneyness, although 10% statistical significance is preserved for at least 25% of the simulations. We compare the results of our simulations, using a range of IVS predictability measures, with option market data concerning S&P 500 index options and a number of equity options traded in the U.S. markets to show that our incomplete information model to a large extent generates the same predictable features described by traded option prices.

⁴ Predictability patterns in the level, slope, and curvature of the IVS have already been reported in studies using S&P 500 index options (e.g., Gonçalves and Guidolin, 2006), as well as individual equity options (e.g., Bernales and Guidolin, 2014).

The closest papers related to ours are David and Veronesi (2002), Guidolin and Timmermann (2003, henceforth GT), and Shaliastovich (2015). These researchers explore the effects of learning on option prices, measured at a certain point in time (i.e., they mostly perform static analyses), to explain the different IVs across strike prices and maturity dates that define the IVS. These studies show that learning induces asymmetric slopes and curvatures in the IVS, which of course are results we also demonstrate. However, our focus is predominantly on providing a rational, asset pricing-based, explanation for the movements over time and the predictability in the IVS, besides calibrating the shape of the IVS itself. For instance, and differently from these earlier papers, our focus is devoted to calibrating and explaining autocorrelations in IVs, the volatility clustering of IVs, and the predictability patterns in slopes and curvatures of the IVS. In particular, David and Veronesi (2002) introduce a continuous-time model in which the dividend drift follows a two-state regime-switching process. In their model, investors' uncertainty about the current state of the economy induces cross-sectional IV skews and systematic shapes in the term structure of the IVS. In our paper, we work in discrete time and with rare, infrequent breaks, not regimes, while our focus is distinctively on the IVS and its dynamic features.⁵ Guidolin and Timmermann (2003) present a discrete-time equilibrium model in which the mean dividend growth rate evolves between two states in a binomial lattice with an unknown but recursively updated state probability. However, in Guidolin and Timmermann's work, learning effects vanish asymptotically as time deterministically flows, because investors eventually achieve complete knowledge of the unknown state probability.⁶ Moreover, our model is more general than a simple binomial lattice and, although less theoretical results can be precisely documented, its calibrated versions give more realistic predictions than in Guidolin and Timmermann. Shaliastovich (2015) introduces a discrete-time long-run risk type model in which the unobservable consumption growth rate has to be learned via a "recency"-biased updating procedure. In his paper, expected consumption growth and its uncertainty are time-varying, while uncertainty is subject to jumps. Compared to his paper,

⁵ In this sense, the most closely related papers are Timmermann (2001) and Guidolin (2006), where infrequent breaks are modeled and empirically estimated, but the goal is simply to explain the features of the realized distribution of stock returns, such as the equity premium, volatility clustering, excess kurtosis, etc.

⁶ Although Guidolin and Timmermann (2003) perform a dynamic analysis, they only examine the weekly fit of their model over time. Therefore, they do not study specifically whether their model may generate predictable dynamics in option prices and the associated IVS.

we use a simpler model with no long run risks or jumps, but restrict the investor to rationally learning about the mean growth rate since the most recent structural break.

There are a few studies that are somewhat related to our research, although they do not specifically investigate the effects of learning on predictable option IV dynamics. Our research has links to studies that examine structural breaks in economic fundamentals in relation to the effects of learning on stock prices and their return process (e.g., Pastor and Stambaugh, 2001; Timmermann, 2001). Lettau and Van Nieuwerburgh (2008) document that the stock return predictability puzzle can be explained by breaks in economic fundamentals. They show that in-sample financial ratios and future returns are significantly related; however, in real time this relation cannot be exploited due to the occurrence of infrequent breaks. They also report that return predictability is mainly affected by the uncertainty induced in the estimation of the new fundamental value after breaks, whereas the uncertainty generated by the detection of breaking dates is less critical. Ederington and Lee (1996) and Beber and Brandt (2006, 2009) show that macroeconomic events or news at both expected and unexpected times increase IVs, while they decline when uncertainty is resolved. Donders, Kouwenberg, and Vorst (2000), Dubinsky and Johannes (2006), and Ni, Pan, and Poteshman (2008) present similar results to Ederington *et al.*'s, although they mostly focus on the effects of earnings announcement dates on IVs.

The rest of the paper is organized as follows. In Section 2, we present the model, and in Section 3 we describe the simulations and results. We document the nature of our qualitative results (i.e., the fact our model “can do the job” requested of it), and then perform a quantitative calibration to show that the framework may re-produce standard econometric evidence on the predictability of the IVs. In Section 4, we report a model extension by allowing the dividend volatility to vary. Concluding remarks appear in Section 5.

2 The model

In Section 2.1 we price options when information is complete, there is no learning, but there are breaks in the process of the fundamentals. Section 2.2 extends the pricing framework

when there are breaks but information is incomplete so that rational learning occurs and therefore affects prices.

2.1 Option pricing under breaks and complete information

We consider a representative agent discrete-time endowment economy as in Lucas (1978). This economy contains three types of assets: a one-period zero-coupon default free bond, B_t , in zero net supply; a stock with net supply normalized at one, S_t ; and a set of redundant call option contracts, $Call_t(K, \tau)$, which are European-style with underlying asset priced at S_t , strike price K , and time-to-maturity τ . The stock pays out an infinite stream of real dividends, D_t ; however the mean (continuously compounded) dividend growth rate, $g_t \equiv \ln(D_t/D_{t-1})$, is subject to unpredictable breaks. The time periods between breaks follow a memory-less geometric process parameterized by π ; and thus the number of breaks in a given period follows a binomial distribution.⁷ We assume that when a break occurs, the new mean dividend growth rate is obtained from a continuous univariate density, $g_{t+1} \sim G(\cdot)$, defined on the support $[g_d, g_u]$. In addition, net of the break dynamics, dividends evolve according to a geometric random walk with constant process with constant volatility process with constant volatility σ and drift $\mu_t + 1$ with constant volatility σ and drift μ_{t+1} ,

$$\ln\left(\frac{D_{t+1}}{D_t}\right) = \mu_{t+1} + \sigma \varepsilon_{t+1}, \quad (1)$$

in which the innovation term, ε_{t+1} , is homoscedastic and serially uncorrelated; however, μ_{t+1} changes over time since it is related to g_{t+1} by $1 + g_{t+1} = \exp(\mu_{t+1} + \sigma^2/2)$.⁸

We assume a perfect, frictionless, and complete capital market: there are no taxes, no transaction costs, unlimited short sales possibilities, perfect liquidity, and no borrowing or

⁷ Shaliastovich (2008) uses a continuous-time Poisson process in his discrete-time learning model to describe jumps in the uncertainty over time, and thus time periods between jumps follow a memory-less exponential process that is also in continuous-time. This kind of set-up is common in the literature. However, we prefer to be consistent with our discrete-time model; thus, we use a discretized version of the Poisson and exponential processes. which are the binomial and the memory-less geometric processes, respectively.

⁸ We assume that σ is constant to obtain the simplest setting to analyze the impact of learning on the dynamics of option prices. This is consistent with earlier work by Timmermann (1996, 2001) who, with reference to equilibrium equity prices, shows that the investors' learning regarding only the mean dividend growth rate is sufficient to induce excess volatility and volatility clustering in stock returns, even though the volatility of the dividend random walk process is constant. However, we extend our model setup in Section 4 by allowing the dividend volatility to vary.

lending constraints. As discussed in Brennan and Cao (1996), it is market completeness that makes options redundant assets. The representative agent has preferences described by a standard power utility function,

$$u(C_t) = \begin{cases} \frac{C_t^{1-\alpha} - 1}{1-\alpha} & \alpha \geq 0 \\ \ln C_t & \alpha = 1 \end{cases} \quad (2)$$

where C_t is real consumption and α corresponds to the constant coefficient of relative risk aversion (CRRA). We assume that dividends represent the unique source of income of this representative agent. As is typical in a Lucas-type model, dividends are perishable and consumed when they are received at any time $t+k$ (i.e., $C_{t+k} = D_{t+k}$). Therefore the agent maximizes the discounted value of her expected stream of future utility choosing assets' holdings and subject to a standard budget constraint,

$$\begin{aligned} & \max_{\{D_{t+k}, w_{t+k}^S, w_{t+k}^B\}} E_t \left[\sum_{k=0}^{\infty} \beta^k u(D_{t+k}) \right] \\ \text{s. t. } & C_{t+k} + w_{t+k}^S S_{t+k} + w_{t+k}^B B_{t+k} \leq w_{t+k-1}^S (S_{t+k-1} + D_{t+k-1}) + w_{t+k-1}^B \end{aligned} \quad (3)$$

where $\beta \equiv 1/(1+\rho)$, ρ is the subjective impatience rate, and w_{t+k}^S (w_{t+k}^B) are the shares of stocks (bonds) in her portfolio. Since call option contracts are redundant assets in zero endogenous net supply, option holdings do not affect the agent's optimization because they fail to appear in her budget constraint. Therefore, option holdings do not affect stock and bond prices. Consequently, Euler equations are obtained for the stock and the bond by standard dynamic programming methods (see Pliska, 1997):

$$S_t = E_t \left[\beta \left(\frac{D_{t+1}}{D_t} \right)^{-\alpha} (S_{t+1} + D_{t+1}) \right] \quad (4)$$

$$B_t = E_t \left[\beta \left(\frac{D_{t+1}}{D_t} \right)^{-\alpha} \right] \quad (5)$$

where $Q_{t+1} = \beta(D_{t+1}/D_t)^{-\alpha}$ is the pricing kernel defined as the intertemporal marginal rate of substitution multiplied by the subjective discount factor.

In this section, we assume complete knowledge of the parameters appearing in the process for real dividends. This means that both μ_t and σ are known. Of course, μ_t remains time-varying so that knowledge of μ_t does not imply it is identical to μ_{t+1} . However, the

occurrence of breaks is also assumed to be observable. Moreover, under complete information (CI), we assume that the distribution from which log-growth rates are drawn, $g_{t+1} \sim G(\cdot)$ with support $[g_d, g_u]$, is also known to the representative investor. Under these simplifying restrictions, although the Euler conditions in (4) and (5) appear to be standard in the literature, solving them as difference equations in the presence of infrequent breaks but complete information yields non-trivial expressions for equilibrium stock and bonds prices, presented in Proposition 1.

PROPOSITION 1 (Complete Information): *Assuming that the mean growth rate g_t is subject to breaks, and that when a break occurs (with probability π), the new mean dividend growth rate is drawn from a given univariate density $g_{t+1} \sim G(\cdot)$ with support $[g_d, g_u]$, where $1 + \rho > (1 + g_u)^{1-\alpha}$, then the stock and bond prices under complete information, S_t^{CI} and B_t^{CI} , are:*

$$S_t^{CI} = \frac{D_t}{1+\rho-(1-\pi)(1+g_t)^{1-\alpha}} \left\{ (1-\pi)(1+g_t)^{1-\alpha} + \pi \left(\frac{I_1+(1-\pi)I_2}{1-\pi I_3} \right) \right\} = D_t \Psi(g_t), \quad (6)$$

where:

$$I_1 = \int_{g_d}^{g_u} (1+g_{t+1})^{1-\alpha} dG(g_{t+1})$$

$$I_2 = \int_{g_d}^{g_u} \frac{(1+g_{t+1})^{2-2\alpha}}{1+\rho-(1-\pi)(1+g_{t+1})^{1-\alpha}} dG(g_{t+1})$$

$$I_3 = \int_{g_d}^{g_u} \frac{(1+g_{t+1})^{1-\alpha}}{1+\rho-(1-\pi)(1+g_{t+1})^{1-\alpha}} dG(g_{t+1});$$

moreover:

$$B_t^{CI} = \frac{1}{(1+\rho)} \left\{ (1-\pi)(1+g_t)^{-\alpha} + \pi \int_{g_d}^{g_u} (1+g_{t+1})^{-\alpha} dG(g_{t+1}) \right\}, \quad (7)$$

in which the one period risk-free interest rate is defined as $r_t^{CI} \equiv 1/B_t^{CI} - 1$.

Proof: See Appendix A.

Proposition 1 has a number of implications. The ex-dividend (real) stock prices are first order homogeneous in dividends and are affected by breaks in g_t . Consequently, the price-

dividend ratio is time-varying and also conditional on g_t . This means that in the absence of breaks (as in GT, 2003), the CI stock price is simply,

$$S_t^{CI, \pi=0} = \frac{(1+g_t)^{1-\alpha}}{1+\rho-(1-\pi)(1+g_t)^{1-\alpha}} D_t = \Psi^{CI} D_t,$$

where Ψ^{CI} is the price-dividend ratio. Note that in the logarithmic utility case (see Veronesi, 1999), it is well known that, when $\alpha = 1$, then:

$$S_t^{CI, \alpha=1} = \frac{1+\pi\rho}{1+\rho-(1-\pi)} D_t, \quad (8)$$

so that the price-dividend ratio is a constant that also depends on π .

Similarly, the one period zero-coupon bond changes over time due to shifts in g_t . The one period zero-coupon bond price is given by the expected pricing kernel in the absence of breaks (when $\pi = 0$), $(1 + g_t)^{-\alpha}/(1 + \rho)$, multiplied by the probability of no breaks $(1 - \pi)$ plus the expected pricing kernel in the case of breaks, $\int_{g_d}^{g_u} (1 + g_t)^{-\alpha} dG(g_t)/(1 + \rho)$, multiplied by the probability of breaks, π . Additionally, the current expected forward price of a one period zero-coupon bond in the very long term is equal to the expected value of the pricing kernel in the scenario of a break, since the probability of having no shifts in the mean in the distant future is practically zero (i.e., $\lim_{s \rightarrow \infty} E_t[B_s^{CI}] = \int_{g_d}^{g_u} (1 + g_t)^{-\alpha} dG(g_t)/(1 + \rho)$)).

Furthermore, pricing European call option contracts is straightforward under complete information. We assume that there are no arbitrage opportunities, and that the agent makes portfolio choices considering asset menus that include stocks and bonds only. This derives from our earlier assumption that markets are complete, so that European options are redundant by construction. In the case of an economy without breaks (i.e., $\pi = 0$), no-arbitrage option prices can be computed as BS prices deriving from equilibrium models in which the dividend fundamental process is stationary [see GT (2003), and references and proofs therein].⁹ However, the BS formula fails to hold when breaks in g_t are possible. Breaks make dividend yields and interest rates time-varying, thereby introducing non-

⁹ Technically, this result obtains only in the continuous time limit. However, here we refer to a discretized BS, fundamental-based formula that in fact goes back to the seminal paper by Rubinstein (1976).

stationarities in the dynamics process followed by the primitive assets that underlie the no-arbitrage of European options. Therefore, the correct discount factors to be applied—even under the risk-neutral measure that characterizes BS pricing—are path-dependent. Nevertheless, option contracts can be priced by a change of measure to the state-price density. Proposition 2 presents an expression for European call option prices based on this change of measure; Section 3 shows how the resulting expression may be solved using numerical methods.

PROPOSITION 2 (Complete Information): *Under complete information, the no-arbitrage price of a European call contract written on the stock, with strike price K and time-to-maturity τ , can be obtained by assuming that the number of breaks, z , between t and $t + \tau$ is a random variable drawn from a binomial distribution $\varphi(z|\tau, \pi)$ with parameters τ and π , $\{h_i\}_{i=0}^z$ are the time intervals between breaks drawn from geometric distributions $\eta(h_i|\pi)$ in which $\tau = \sum_{i=0}^z h_i$, and the new post-break growth rates $\{g_{t+h_i}\}_{i=1}^z$ are drawn from a univariate density $g_{t+h_{i-1}} \sim G(\cdot)$ with pdf $\varrho(\cdot)$ defined on the support $[g_d, g_u]$, where $g_{t+h_0} = g_t$ and $g_{t+\tau} = g_{t+h_z}$. Thus, the price European call contract is:*

$$Call_t^{CI}(K, \tau) = \int_0^{\infty} \max\{S_{t+\tau}^{CI} - K, 0\} \tilde{p}_t(S_{t+\tau}^{CI}) dS_{t+\tau}^{CI} \quad (9)$$

in which $S_{t+\tau}^{CI} = D_{t+\tau} \Psi(g_{t+\tau})$, $D_{t+\tau} = D_t \exp(\sqrt{\tau} \sigma \varepsilon_{t+\tau} - \tau \sigma^2 / 2) \prod_{i=0}^z (1 + g_{t+h_i})^{h_i}$, and $\varepsilon_{t+\tau}$ is the innovation term of the dividend geometric random walk process in (1) distributed as a normal density $\phi(\varepsilon_{t+\tau}|0, \sigma)$ with mean zero and variance σ . Finally, the state price density is characterized as:

$$\tilde{p}_t(S_{t+\tau}^{CI}) = \beta^\tau \left(\frac{D_{t+\tau}}{D_t}\right)^{-\alpha} \phi(\varepsilon_{t+\tau}|0, \sigma) \varphi(z|\tau, \pi) \eta(h_0|\pi) (\eta(h_1|\pi) \varrho(g_{t+h_1}) \cdot \dots \cdot \eta(h_z|\pi) \varrho(g_{t+h_z})).$$

Proof: See Appendix A

Proposition 2 shows that an option contract should now be priced by taking into account that there exists a probability $\pi > 0$ that g_t may be affected by a structural, and possibly permanent change in any period before the option's expiration date. Therefore, given that after a break the new value of g_t randomly changes (i.e., after a break g_t can take any value from the density $G(\cdot)$), this induces additional instability in the model and modifies the risk-

neutral probability distribution. Thus, breaks (even without learning) make wider the risk-neutral probability distribution than in the BS case with constant g_t values. This affects option variables including option prices, IVs, and deltas. For instance, we show in Section 3 (see Table 1) that, when there are breaks and complete information, option prices and IVs increase in magnitude, they are more volatile, and have more skewness and kurtosis.

2.2 *Option pricing under breaks and incomplete information with Bayesian learning*

In Section 2.1, the equations for pricing the bond, the stock (tree), and any cross-section of option contracts introduced in Propositions 1 and 2 were derived assuming that the agent knows the true mean dividend growth rate at each point in time. However, these expressions are not valid when there is incomplete information in the economy and an investor needs to learn the unknown parameters driving the process of fundamentals. Suppose that g_t is unknown and the representative agent efficiently uses all available information to price the assets following a Bayesian updating procedure. Within each regime, as defined by the last occurrence of a break, the agent receives new, independent signals about the mean dividend growth on a daily basis, $\{D_i/D_{i-1}\}_{i=t-n+1}^t$, which are random and follow a lognormal distribution where n is the number of periods since the last break (see equation (1)). However, breaks are still assumed to consist of rarely occurring and rather visible events so that they are observable. The investor's learning therefore concerns only the actual value of g_t following the most recent break. Although it would be possible to extend our set up to model the effects induced by any learning/estimation of breakpoint dates, the cost of this extension in terms of analytical complexity is remarkable, with a consequent loss of intuition for the results derived below. Moreover, assuming knowledge of the breakpoint dates does not appear to be completely unrealistic, as a number of recent econometric advances have shown that it is possible to perform real time tests that monitor for breaks in the mean function, attaining a considerable degree of accuracy (Chu, Stinchcombe, and White, 1996; Leisch, Hornik, and Kuan, 2000).¹⁰

¹⁰ Additionally, Lettau and Van Nieuwerburgh (2008) show that the uncertainty generated by the detection of breakpoint dates in the process of economic fundamentals is not critical to explaining stock return anomalies.

The agent uses her prior beliefs while she recursively learns, incorporating all the new information received. Therefore, as in Timmermann (2001), the expected value of any asset, $\lambda_t(g_t|\pi, \rho, \alpha, \sigma)$, given an agent's prior beliefs $q(g_t)$, can be obtained from the following updating Bayesian rule:

$$E_{t,n}^{BL}[\lambda_t(g_t)|\mathbf{q}_t] = \frac{\int_{g_d}^{g_u} \lambda_t(g_t) L(g_t|\mathbf{q}_t)_n q(g_t) dg_t}{\int_{g_d}^{g_u} L(g_t|\mathbf{q}_t)_n q(g_t) dg_t}, \quad (9)$$

where $L(g_t|\mathbf{q}_t)_n$ is the sample likelihood function, and the vector of signals is represented by $\mathbf{q}_t = [(D_t/D_{t-1}) \dots (D_{t-n+1}/D_{t-n})]$. The intuition behind equation (9) is simple. The agent does not know the new value of g_t after a break but she knows the distribution followed by g_t ($q(g_t)$) and the distribution followed by the signal vector \mathbf{q}_t . Given this knowledge, the agent recursively updates her expectations about g_t and on the value of all assets that depend on g_t , $\lambda_t(g_t)$, as new signals are observed using Bayes' rule. After a break affecting the true but unknown value of g_t , the new value will only be gradually learned because of the contemporaneous presence of the random innovation in the random walk process followed by dividends, for a given growth time-varying parameter g_t .

Therefore, in addition to the non-stationarities induced by the presence of breaks in the process followed by the price of all assets, incomplete information and learning generate incremental randomness in the value of stocks, bonds, and option prices. Immediately after a break, historical information is scarce concerning reliable values for the asset prices. Thus, there is an initial period of intense learning that generates important changes in the agent's beliefs, which induces an additional uncertainty in the valuation process of all assets. The incremental randomness will affect option prices and their implicit volatilities as options are non-linear securities (see Figure 1 in Section 3). Therefore, since there is incomplete information, breaks modify the quantity and reliability of the information the agent can access about the mean dividend growth rate g_t ; hence, breaks impact the speed and intensity by which the investor updates her knowledge of the economic conditions.

They also show that the main source of uncertainty is caused by the estimation of the magnitude of the new parameters in the aftermath of the break dates, similarly to our modeling approach.

Nevertheless, these large adjustments in beliefs are reduced recursively over time, as more information is learned.

In equation (9), instead of dealing with a complex sample likelihood function concerning log-normally distributed data, it is convenient to re-write Bayes' rule taking μ_t as the unknown parameter without loss of generality, because $1 + g_t = \exp(\mu_t + \sigma^2/2)$. The main advantage of parameterizing (9) as a function of μ_t is that the observable signals, $\{\ln(D_i/D_{i-1})\}_{i=t-n+1}^t$, will now follow a normal density, so that equation (9) can be written as:

$$E_{t,n}^{BL}[\lambda_t(\mu_t)|\xi_t] = \frac{\int_{\mu_d}^{\mu_u} \lambda_t(\mu_t) L(\mu_t|\xi_t)_n f(\mu_t) d\mu_t}{\int_{\mu_d}^{\mu_u} L(\mu_t|\xi_t)_n f(\mu_t) d\mu_t} \quad (10)$$

with

$$L(\mu_t|\xi_t)_n = \frac{1}{\sqrt{2\pi\sigma^2/n}} \exp\left[-\frac{(\bar{\xi}_t - \mu_t)^2}{2\sigma^2/n}\right] \quad (11)$$

in which $\xi_t = [\ln(D_t/D_{t-1}) \dots \ln(D_{t-n+1}/D_{t-n})]$, and $\bar{\xi}_t = (1/n) \sum_{i=t-n+1}^t \xi_i$ is the sample mean. At this point, building on equations (10) and (11), it is possible to derive results for the price of assets in the presence of infrequent breaks and under incomplete information with learning, collected in the Propositions 3 and 4 that follow. Interestingly, these use the closed-form expressions under complete information already derived in Propositions 1 and 2.

PROPOSITION 3 (Bayesian Learning): *Assuming incomplete information and learning, the stock and bond prices are given by:*

$$S_t^{BL} = \frac{\int_{\mu_d}^{\mu_u} S_t^{CI} L(\mu_t|\xi_t)_n f(\mu_t) d\mu_t}{\int_{\mu_d}^{\mu_u} L(\mu_t|\xi_t)_n f(\mu_t) d\mu_t} \quad (12)$$

and

$$B_t^{BL} = \frac{\int_{\mu_d}^{\mu_u} B_t^{CI} L(\mu_t|\xi_t)_n f(\mu_t) d\mu_t}{\int_{\mu_d}^{\mu_u} L(\mu_t|\xi_t)_n f(\mu_t) d\mu_t} \quad (13)$$

where S_t^{CI} and B_t^{CI} are the stock and bond price expressions under breaks and complete information defined in Proposition 1.

PROPOSITION 4 (Bayesian Learning): *Under incomplete information and learning, the prices of European call options written on the stock, with strike price K , and time-to-maturity $\tau = T - t$ are:*

$$Call_t^{BL}(K, \tau) = \frac{\int_{\mu_d}^{\mu_u} \int_0^{\infty} \max\{S_{t+\tau}^{CI} - K, 0\} \tilde{p}_t(S_{t+\tau}^{CI}) dS_{t+\tau}^{CI} L(\mu_{t+\tau} | \xi_{t+\tau}) n_{t+\tau} f(\mu_{t+\tau}) d\mu_{t+\tau}}{\int_{\mu_d}^{\mu_u} L(\mu_{t+\tau} | \xi_{t+\tau}) n_{t+\tau} f(\mu_{t+\tau}) d\mu_{t+\tau}}, \quad (14)$$

where $S_{t+\tau}^{CI} = D_{t+\tau} \Psi(g_{t+\tau})$, $g_{t+\tau} = \exp(\mu_{t+\tau} + \sigma^2/2) - 1$, dividends on expiration date follow $D_{t+\tau} = D_t \exp(\sqrt{\tau} \sigma \varepsilon_{t+\tau} - \tau \sigma^2/2) \prod_{i=0}^z (1 + g_{t+h_i})^{h_i}$, $\varepsilon_{t+\tau}$, $z \leq \tau$, $\{h_i\}_{i=0}^z$, $\{g_{t+h_i}\}_{i=1}^z$, and $\tilde{p}_t(S_{t+\tau}^{CI})$ are as in Proposition 1. In addition, $n_{t+\tau}$ is the number of dividend signals since the last break.¹¹

The simplicity of the Bayesian updating procedure underlying Propositions 3 and 4 is useful to our understanding of the effects of learning on asset prices. Using this perspective, we start by analyzing two special cases that are illustrative of the mechanics of the effects of rational learning. Firstly, suppose that the probability of a break is very large, $\pi \rightarrow 1$. This implies that the agent faces very frequent breaks, at the same frequency as calendar time (say, daily). In this case, learning has no effect because g_t changes in all periods, so that “there is no time for the investor to learn.” In this case, the expressions (12)-(14) simplify to (6)-(9) under the restriction that $\pi = 1$. Secondly, when $\pi = 0$ learning vanishes as $t \rightarrow \infty$, as in GT (2003). In this case the agent should have sufficient information after a long period to calculate accurate estimates for g_t and asset prices; and thus the effects of learning will disappear asymptotically. In this case, the expressions (12)-(14) converge to (6)-(9) under the restriction that $\pi = 0$.

Propositions 3 and 4 show that after a break, substantial revisions in the agent’s expectations may occur, which strongly affect asset prices. Immediately after breaks, as mentioned previously, the agent does not have enough historical information to obtain reliable estimates, which induce large adjustments in beliefs. Beliefs’ revisions generate important variations in the prices of all assets. The effect of incomplete information and

¹¹ Note that $n_{t+\tau} \neq n + \tau$ since there are chances of breaks occurring between t and $t + \tau$; therefore $n_{t+\tau}$ is also a random variable, where $n_{t+\tau} \leq n + \tau$.

learning will modify asset prices depending on the level of risk aversion of the agents, as shown in Section 3. However, initial revisions of large beliefs (after breaks) progressively abate over time as more information is received and learned. This updating process of agents' beliefs, caused by recursive information acquisition, will induce rich patterns of predictable dynamics in option prices and IVs.

3 Simulation results: making sense of the econometric evidence

Section 3.1 explains the structure of the simulation/calibration work that follows. Section 3.2 calibrates our model. Section 3.3 shows the general, qualitative pattern of our key results. Section 3.4 takes our calibration and simulation seriously and tries to match the key stylized features of the data.

3.1 Research design

The main goal of our research is to provide an understanding of whether (and how) a Bayesian learning scheme applied to processes subject to infrequent breaks may be used to explain a number of stylized facts concerning the pricing of European index options. We perform such an investigation, also assessing the implications of alternative assumptions concerning learning dynamics (i.e., the features of the way investors update their expectations over time). Following the same arguments as Timmermann (1993, 1996, 2001), Veronesi (1999, 2000), Guidolin (2006), and David and Veronesi (2013), who all evaluate how the learning process affects the properties of stock returns by performing extensive sets of simulations, we use a quantitative Monte Carlo approach. The aforementioned authors argue that learning influences the pricing function of all assets in a highly nonlinear way, which would be poorly approximated by any attempts at log-linearization, so that simulations are necessary to understanding the wider scope of outcomes that learning may induce.¹² Moreover, the use of simulations allows us to modify parameter configurations and observe the impact of learning in multiple environments.

¹² Kleidon (1986) shows that the use of standard tests to evaluate an equilibrium model using a single economy represented by market data may lead to inaccurate conclusions. He emphasizes that asset prices in

As a first step in our Monte Carlo experiments, we reproduce the Chicago Board Options Exchange (CBOE) rules in terms of strike price intervals, expiration dates, and listing and delisting policies. The main objective of the replication of the CBOE rules in the generation of our simulations is to provide realism to Monte Carlo experiments; and thus to make our outcomes “more” comparable to the results obtained from real market data. A detailed explanation about how the CBOE rules are implemented in our model simulations is provided in Appendix B.

Throughout this paper, we calculate IVs by numerically inverting the BS model, which is consistent with both previous academic studies and investor practice, where IVs are estimated using this model even though it is well-understood that its assumptions are violated by market data. Obviously, as stated in Section 2, our model departs from the BS model because of the richer dynamics of the process of fundamentals, as well as because the index options written on the market are priced off fundamentals in complete markets, as in GT (2003).¹³ Moreover, the well-known predictability patterns in both IVs as well as in the shapes of the IVS, which we explain through our learning model, have been always reported and discussed with reference to implicit volatilities calculated under the BS model (e.g., Harvey and Whaley, 1992; Gonçalves and Guidolin, 2006; Konstantinidi, Skiadopoulos, and Tzagkaraki, 2008; Chalamandaris and Tsekrekos, 2010). In this respect, one may see the (probably misspecified) use of BS to compute implicit IVs as a “wash out”: in the same way in which market data that are not generated from BS assumptions are transformed into BS IVs, simulated prices computed under alternative assumptions on the mechanism of expectation formation are transformed into IVs using the same, commonly used device, Black-Scholes, to make any comparisons possible.

equilibrium are calculated based on agents’ expectations about future events across multiple and different economies. Instead, Kleidon proposes the use of multiple realizations by simulation techniques.

¹³ Note that this claim is already true under complete information provided that $\pi > 0$. Of course, this is all the more correct under incomplete information, because of the effects of Bayesian learning and independently of whether $\pi > 0$ or not. However, when $\pi = 0$, note that the no-arbitrage option prices asymptotically converge to BS/Rubinstein prices as $t \rightarrow \infty$.

3.2 Calibration

We assume the following parameter values to be held constant in the simulations. In a few cases, especially where preferences are concerned, we produce, tabulate, and discuss results across a range of parameters to emphasize that these are hardly relevant—unless otherwise noted—to the general tone of our qualitative findings. The subjective rate of impatience, ρ , is set to equal either 0.713% or 0.767% (on a monthly basis). Using methods similar to GT (2003) and Guidolin (2006), we verify that on average, over our 1996-2007 sample path and using the parameters that follow, under incomplete information and Bayesian learning these parametric choices imply annualized equilibrium short-term rates that appear to be realistic with reference to the long-run properties of the U.S. financial market. We also assume that when $\rho = 0.713\%$ ($\rho = 0.767\%$) the new mean dividend growth rate after breaks is extracted from a uniform distribution with upper and lower boundaries of $g_u = 8.8\%$ ($g_u = 9.5\%$) and $g_d = -1.5\%$ ($g_d = -5.0\%$), expressed in annualized terms. As a result, one can verify that $1 + \rho > (1 + g_u)^{1-\alpha}$ for all the values of CRRA employed in this paper (see below).¹⁴ Proposition 1 shows that such a condition guarantees the existence of an equilibrium pricing function; Proposition 3 relies on the same assumptions made in Proposition 1 under the case of complete information, so this inequality is also sufficient for existence when learning occurs. The dividend process volatility, σ , is also set at two alternative values, 5% and 30% (on an annual basis), to span a range of possibilities. Of course, 5% is consistent with the typical parameterizations for the process followed by real dividends in the U.S. (e.g., Timmermann, 2001; GT, 2003); 30% represents instead a high-volatility case in which investors may be learning directly from past stock market returns, more than from the process of fundamentals itself.¹⁵ Finally, for the CRRA coefficient α , we use $\alpha = 0.2$, $\alpha = 0.5$, and $\alpha = 5$. Levels of α below 1 are both consistent with the evidence in the data of a relatively high (certainly in excess of 2) intertemporal elasticity of substitution in consumption; under power utility, such an

¹⁴ Therefore, and given that the new mean dividend growth rate after breaks is extracted from a uniform distribution with probability density function $f(g_t) = 1/(g_u - g_d)$, in corollary 1 and corollary 2 the dividend drift has as probability density function: $f(\mu_t) = \exp(\mu_t + \sigma^2/2)/(g_u - g_d)$, where $\mu_d = \ln(1 + g_d) - \sigma^2/2$ and $\mu_u = \ln(1 + g_u) - \sigma^2/2$.

¹⁵ In this paper, we also examine the case of $\sigma = 30\%$ per year as in this case learning may only occur very slowly between structural breaks, as any signals concerning the drift of fundamentals is confounded by the high variability of diffusive shocks that hit them.

intertemporal elasticity of substitution is simply the inverse of α , as $\alpha < 1$ appears sensible. Moreover, evidence in Timmermann (2001) and Guidolin (2006) has shown that under rational learning, provided an equilibrium exists, for $\alpha < 1$ the equity premium appears to be increasing in α as α declines towards zero (see also David, 2008), while the riskless short-term rate declines (as is customary in Lucas tree models). However, given that this level of the CRRA is also commonly perceived as “acceptable” (even though it is inconsistent with commonly estimated intertemporal elasticity of substitution coefficients), we also entertain the case of $\alpha = 5$.

We use the recursive, real time monitoring breakpoint test introduced by Chu, Stinchcombe, and White (1996) to estimate a probability of breaks, π , affecting the mean real dividend growth rate. In Appendix C, we describe the model introduced by Chu, Stinchcombe, and White (1996) and the breaks detected by the application of their method to a series of S&P 500 stock dividends over a sample of daily data from the 1996 to 2007 period, which are also de-seasonalized and adjusted by the Consumer Price Index to obtain real dividends. We find eight breaks in the 3,012 days of the 12-year sample we analyze. Therefore we set π at 0.0030 per day (0.667 on an annual basis). In essence, real dividend data confirm that breaks are indeed possible on a daily basis, but only with a negligible probability of less than 1% per day; equivalently, the absence of breaks is expected to last on average for 333 days in a row.

As mentioned above, we simulate multiple case scenarios depending on three assumptions about the representative agent’s expectations: (A) an economy without breaks and complete information;¹⁶ (B) an economy with breaks and complete information; and (C) an economy with breaks and incomplete information, under Bayesian learning. For case (A) (when g is constant), we calculate stock and bond prices assuming that $\pi = 0$ using equations (6) and (7), respectively. In this case, European option prices are obtained from the BS model in which the dividend yield is $\delta^{BS} = (1 + \rho - (1 + g_t)^{1-\alpha}) / (1 + g_t)^{1-\alpha}$, see GT (2003) for a proof. In case (B) of an economy with breaks and complete information, stock

¹⁶ In this case, it is irrelevant to specify whether information is complete or must be learned. Under Bayesian learning, if we were to simulate such an economy for a period of 12 years, we would find that by the end of the exercise, such an economy would behave in the same way as a complete information one. Therefore, we simply assume that information is complete.

and bond prices are calculated from (6) and (7) assuming $\pi = 0.0030$. In addition, European call prices are calculated using equation (8) where the main integral is solved by Monte Carlo methods, on the basis of 20,000 independent paths from the stochastic process described in Proposition 2. In case (C), the case of breaks and incomplete information with learning, stock and bond prices are obtained from (12) and (13) with $\pi = 0.0030$, while option prices are taken from equation (14), again using Monte Carlo methods.¹⁷

Under each of the three alternative case scenarios and for each possible combination of parameters—12 in total, obtained by combining two values for ρ , two for σ , and three for the CRRA coefficient α —we generate 2,000 simulations. On each simulated path, we produce 12 years of daily real dividends (3,018 days), which represent the observable signals received by the investor and used to learn about g_t . The simulations are generated by the two-step subordinated stochastic processes described in equation (1). This means that in the absence of breaks (for a constant g), we simulate time series of 12 years of daily dividends using a geometric random walk process. Additionally, we induce stochastic breaks in g_t on each time step of each simulated path (hence, we generate breaks in μ_t) according to the assumed geometric process parameterized by π . For instance, in the case in which a break occurs at time $t = m$, we obtain a new value for g_m drawn from a uniform distribution $g_m \sim U(\cdot)$ defined on the support $[g_d, g_u]$ and keep this value constant until the next break is generated.

Even though we price the stock index and the bond on a daily basis, we calculate option prices across strikes and maturities on a weekly basis, with the objective of saving computational time. The use of weekly data has been common in the empirical option pricing literature (e.g., Dumas, Fleming, and Whaley, 1998). In particular, we calculate option prices on the Wednesday of each week (which corresponds to steps of five simulated days), without any loss in our qualitative insights.¹⁸

¹⁷ In addition, using Monte Carlo methods we simultaneously estimate the expected dividend yield and expected zero curves over the residual “life” of each option contract with the objective of obtaining the necessary inputs for IV computation.

¹⁸ In all simulations, we use the same trading dates that effectively happened over the 12-year sample of 1996 - 2007, thus accounting for holidays and unexpected events in which the market was closed (see Appendix B for additional explanations of our simulations). The idea of using “real” trading dates is to increase the realism, as well as the reliability of our results. Note that most of the literature has completely ignored the

3.3 Qualitative results

The existence of breaks in the mean dividend growth rate and the need of investors to learn about such an unstable, time-varying parameter causes non-stationarities in stock, bond, and option prices. These are at the core of the ability of a model with incomplete information and rational learning to explain key stylized facts concerning IVs. To get some intuition for the nature of the instabilities captured by our framework, Figure 1 displays one complete simulation path in terms of simulated mean dividend growth rates (g), equilibrium stock prices (S), and at-the-money ($K/S = 1$) short-term (30 calendar days to maturity) IVs (IVATM, Short-T) under our three cases (A)-(C) listed above. Equilibrium stock and option prices are computed for the case of $\alpha=0.2$.^{19,20} In the upper panel, we plot three time series: one is trivially the constant level of g in the absence of breaks; the second, step-like function, corresponds to the time series of g_t when the mean dividend growth rate is affected by infrequent breaks (seven breaks over the simulation periods, which is realistic in the light of the evidence in Appendix C); the third is the recursive inference on mean dividend growth rate obtained by a rational investor using Bayes' rule based on the empirical likelihood of the data. Looking at this third series, one can notice that learning may occasionally take a long time. Estimates of g_t progressively adjust towards the true values after each break. However, there are also cases in which the investor's estimate of g_t since the most recent break actually drifts away from the fixed but unknown value (see the upper panel around the simulated observation 500). On the one hand, the observable dividend signals received by the investor are noisy because of the presence of an innovation term in the geometric random walk process. Consequently, the agent needs time to learn

effects of such rare occurrences on econometric tests. In the case of weekly option valuations, we select Wednesdays since this minimizes the incidence of the number of holidays because, as mentioned above, we simulate according to the actual authentic CBOE trading calendar between 1996 and 2007.

¹⁹ As simulations replicate option prices in a realistic way that tracks CBOE rules, so that 30-day at-the-money option contracts are not always offered and traded, we calculate IVATM, Short-T IVs by simple linear interpolation using the four contracts around a 30-day time-to-maturity mark and with closest strike price to S .

²⁰ The lower panel of Figure 1 depicts IVs instead of option prices because the former are easier to interpret and analyze than the underlying option prices, where IVs are extracted from prices using the BS model. The direct use of option prices is not advisable in comparative analyses due to the fact that option prices differ in their 'level' depending on option contract features.

and thus to obtain accurate values for the unknown g_t . On the other hand, in this figure a new break often appears when learning has improved the accuracy of the agent's estimations and hence her cognitive process strengthens once again. Importantly, the simulated real dividends underlying both the complete versus the incomplete information scenarios in Figure 1 are identical, and the differences are purely due to the need by the investor to learn in the second case.

[Insert Figure 1 here]

The middle panel of Figure 1 shows a particular path for stock prices. Visibly, at least in this particular realized path, the simulated time series of prices in the presence of breaks—both for the complete and incomplete information case scenarios—are substantially lower than in the case of no breaks. This result is simply due to the lower g_t values in the case scenarios with breaks than in the stationary scenario through the whole simulated period (see Figure 1, upper panel). However, this effect is not structural, in the sense that alternative simulations might have generated different effects (i.e., break-induced stock prices that are higher than no-break equilibrium ones). Finally, the third panel of the figure shows that the occasionally intense revisions of agents' expectations about the (new, post-break) value of g_t induce an increase in IVs, especially in the immediate aftermath of breaks, when the learning speed accelerates and revisions are stronger. The difference recorded between the times series for the case of breaks but complete information and the time series under breaks and learning shows that it is mostly learning and not breaks that are responsible for the elevated IVs and the spikes visible in the third panel. This lower panel also points to the possibility of serial correlation and volatility clustering in option IVs, which is one of the features we focus on in the following section. All in all, Figure 1 helps emphasize that the interaction between learning and breaks may strongly affect both the level and the dynamics of IVs, an indication that option pricing is potentially affected by the induced dynamic premia.

On average, the compounding of the infrequent, limited non-stationarities in fundamentals generated by rational learning affect the deep properties of the security market economy. This is emphasized for two alternative configurations of the calibration parameters, but always with reference to the case of $\alpha = 0.2$ in Table 1. In Table 1, we report summary

statistics across simulations for real dividends, for the mean dividend growth rate (observable and constant in one case, and subject to breaks, observable or not observable in the rest of the table), the short-term interest rate, the stock price, the short-term at-the-money call price, and its IV (subject to the same approximations that we have described above). The top panel of Table 1 concerns the case of low volatility of fundamentals, the lower panel concerns the high volatility case.²¹

[Insert Table 1 here]

Table 1 shows that dividends are exactly the same for all scenarios, as it should be. Breaks only affect the nature of the subordinate process of fundamentals, not its average or median levels. However, breaks inflate the standard deviation and the tail thickness of the dividend distribution. Additionally, it is particularly interesting to observe that the standard deviation of the g_t is higher in the scenario with breaks and complete information than in the case of breaks and learning (for estimated mean growth), which is easily understood observing the upper panel in Figure 1: under complete information, when mean growth is observable, the plot shows large changes in g_t on breaking dates, where by definition shifts in g_t are immediately recognized by the agent. Conversely, in the initial periods after a break, in an economy with Bayesian learning, the agent recursively incorporates new information giving some weight to her prior beliefs, and thus producing only gradual movements and smoother adjustments. As already observed, there are no structural differences in the means (of approximately 813-814 index points) and medians (approximately 761-770 index points) of stock prices across alternative scenarios. However, as one would expect—both because dividends are more volatile and because the price-dividend ratio also becomes time-varying—stock prices become more volatile, and slightly more skewed to the right. Correspondingly, as one would expect [see also the proof in GT (2003) for the case of no breaks], average option prices are higher in the presence of breaks and Bayesian learning. For instance, the average at-the-money, one-month call price is 4.4 points in the absence of breaks, 5.4 points when observable breaks are introduced, and 12.8 points when breaks support a sustained learning process.

²¹ We report additional summary statistics for fundamentals and asset prices in an Online Appendix, in which we report an experiment where α is set at 0.5 and 5.0 using the same combinations of parameters and scenarios as in Table I.

A comparison of cases (A) and (B) in Table 1 shows that breaks by themselves induce an increase in IVs (e.g., in the top panel from an average of 5% that exactly matches the assumed level of σ to 5.75%). However, this effect is smaller than the impact of the information incompleteness and learning. For instance, the top (lower) panel shows that IVs increase from 5.75% (31.63%) in under breaks and full information (case (B)) to 19.24% (43.28%) in the case under breaks and incomplete information with learning (case (C)). In addition, learning also induces skewness and kurtosis in option prices and IVs that are otherwise absent.

However, the power of learning to induce these realistic features in option prices and IVs is strongly affected by the assumed curvature of the representative investor's utility function. Figure 2 shows the results of a sensitivity analysis using a range of relative risk-aversion levels α applied to the IVS shape features, in an economy under breaks and incomplete information with learning. Figure 2 reports the average behavior of the IVs of multiple option contracts one month after a break in g_t . In Figure 2, the average values of IVs are presented across both the moneyness dimension using short-term option contracts (the first row of plots) and the maturity dimension by the use of at-the-money option contracts (lower windows). Panels to the right refer to the case of $\alpha < 1$, while panels to the left the case of $\alpha > 1$.²²

[Insert Figure 2 here]

The two upper plots in Figure 2 show that rational learning produces the typical skews of IV/asymmetric “smiles” that have often been reported in the literature: IVs are higher for deep in-the-money calls and deep out-of-the-money puts. The intuition behind the results is that because option prices depend on expectations of future fundamentals in a highly nonlinear way, the effects of Bayesian learning across alternative moneyness levels is asymmetric and—even when option prices have been filtered through the BS' formula—they create highly asymmetric IVS shapes. Additionally, the two upper panels of Figure 2 imply a negative relation between α and IV levels when $\alpha < 1$, while the relation turns positive when $\alpha > 1$. The intuition behind these results is simple: Learning has the lowest

²² In the case of $\alpha = 1$, the analysis that follows Proposition 1 shows that despite incomplete information and learning, the price-dividend ratio becomes a constant and learning has no effect [see Veronesi (1999) and David (2008) for similar remarks]. Therefore our analysis abstracts from such a limited case.

(zero) impact on stock and option prices when $\alpha = 1$ since the components of the valuation formulas that are affected by any unknown, time-varying parameter g_t disappear (see equations (6) and (14)). In fact, in this situation, the BS case of a completely flat IVS obtains [unreported in the figure, see GT (2003)]. Learning has its strongest effects as $\alpha \rightarrow 0^+$, where it has to be taken into account that the existence of the equilibrium requires that the condition $1 + \rho > (1 + g_u)^{1-\alpha}$ always has to be satisfied, which prevents α from being set to zero if $g_u > \rho$, as the absence of arbitrage requires. As $\alpha > 1$ grows, the effects of learning progressively weaken, but the risk premium associated with market variance grows, which explains why in the rightmost upper panel, average IV resumes an increasing pattern as $\alpha > 1$ gets larger.

Moreover, the two upper panels of Figure 2 show that when $\alpha < 1$, slopes and curvatures of the IV skews increase with α , which means that IV skews become steeper, while when $\alpha > 1$, slopes and curvatures of IV decrease as α grows (i.e., the IVS flattens in the moneyness dimension). These results are due to the fact that learning has its strongest effects over in-(out-)of-money call (put) contracts, as mentioned above. Moreover, as $\alpha \rightarrow 1$, the impact of learning on option prices tends to disappear faster for at-the-money and out- (in-)of-money call (put) contracts, at least in relative terms. This causes the IV shapes to display steeper slopes and higher curvatures when $\alpha \rightarrow 1$ than in other cases.

The two lower panels of Figure 2 show that learning also induces downward sloping shapes in the IVS as a function of time-to-maturity, as IVs strongly decrease as time-to-maturity increases. Immediately after (Figure 2 takes a picture of BS IVs one month after) a break, there are intense revisions of agents' expectations concerning the new, unknown value of g_t . However, a Bayesian agent expects that she will learn progressively because she will receive further information to make her perception of the mean growth rate of fundamentals increasingly precise. These expectations of future learning reduce the price and hence the IVs of long-term option contracts in relation to short-term option contracts. Furthermore, the lower panels in Figure 2 show that the IV levels are the lowest when CRRA is close to 1. This is explained by the same arguments used above: learning effects on stock and option prices are nil when $\alpha = 1$. In conclusion, the evidence in Figure 2 shows that our Bayesian learning model results are consistent with the large literature on IV

variations across moneyness and time-to-maturity (e.g., Rubinstein, 1985; Dumas, Fleming, and Whaley, 1998; Das and Sundaram, 1999).

Figure 3 reports sensitivity analysis results concerning the effects of the CRRA parameter α similar to Figure 2; however, Figure 3 concerns the average behavior of IVs as a function of moneyness and time-to-maturity *one year* after a break in g_t . The choice of one year corresponds to the fact that, as shown in Section 3.1, our estimates indicate that the average duration of a given regime as defined by the current value of g_t should be of approximately one year. The comparative analysis results in Figures 2 and 3 allows us to make comments on the varying learning effects on the IVS over time. They show that the average level of IVs decreases as more information is received since the last observed break. The speed of learning, and consequently its effects, weaken when an investor receives a growing amount of dividend signals that ought to allow her to form relatively precise inferences concerning g_t . In contrast to earlier papers, such as GT (2003), the effects of learning never disappear altogether, even one year after the most recent break.

Moreover, the upper rightmost panel of Figure 3 shows that, when plotted as a function of moneyness, IVs describe convex functions for α values close to 1 when $\alpha > 1$; however, as α increases, the IV shapes describe concave curves. As mentioned, additional uncertainty (induced by the information incompleteness in the context of Bayesian learning) induces an asymmetric effect on option contracts, since different option contracts have diverse levels of nonlinearities across alternative moneyness values; which generates asymmetric IVS shapes. In particular, the reason for concave shapes is that when $\alpha > 1$, agent endowment-based asset pricing models in general display a counterintuitive feature by which stock prices are lower when g_t increases (Abel, 1988; Cecchetti, Lam, and Mark, 1990). This counterintuitive feature, that in any event takes place only when learning is weak, induces these concave forms in the moneyness dimension.²³

[Insert Figure 3 here]

²³ Despite this counterintuitive feature of dynamic equilibrium models when $\alpha > 1$, we include them in our analyses to be consistent with the larger literature in which $\alpha > 1$ has been estimated or used to explain properties of asset prices that do not directly concern our paper.

3.4 Quantitative analysis

Besides using the qualitative methods and intuition in Figures 2 and 3 and Table 1, we also employ statistical methods to quantify the effects of breaks and learning under incomplete information on the dynamics of the IVS. In Table 2 (Panels A and B) we find evidence on the dynamic features for IVs as well as the IVS shape movements from simulations concerning an economy characterized by infrequent breaks and incomplete information. Such features are represented by means of simulation-specific moments (e.g., means, standard deviations, serial correlations, and ARCH coefficients). In the case of serial correlation and ARCH(q) tests, besides the average of the corresponding test statistics across simulations, the percentages in parentheses in Panels A and B are the fraction of the total number of simulation trials that imply Ljung-Box and ARCH Lagrange multiplier tests with p -values of 1% or lower. The results in Panels A and B are obtained from simulations using two different parameter setups, as in Table 1. The Online Appendix provides simulation results for additional parameter combinations as a robustness check. We define $Slope_{Mon}$ ($Slope_{Mat}$) as the average across simulation trials of numerical first derivatives with respect to moneyness (time-to-maturity) computed from all the pairs of priced options with neighboring moneyness levels and 30 days to maturity (neighboring maturity levels and closest at-the-money). In addition, $Curv_{Mon}$ ($Curve_{Mat}$) is the average across simulation trials of numerical second derivatives with respect to moneyness (time-to-maturity) computed from all triplets of priced options with neighboring moneyness levels and 30 days to maturity (neighboring maturity levels and closest at-the-money).^{24,25}

[Insert Table 2 here]

²⁴ A numerical first derivative is simply defined as $f'(x_1) \equiv (f(x_1) - f(x_0))/(x_1 - x_0)$; a numerical second derivative is instead $f''(x_1) \equiv (f(x_2) - 2f(x_1) + f(x_0))/(0.5(x_2 - x_0))^2$.

²⁵ An alternative way to characterize the IVS shape and its dynamics is through deterministic IVS models, which describe the implied volatilities as a function of an option strike price and time-to-maturity (Dumas, Fleming, and Whaley, 1998). Moreover, these polynomial functional forms have been successfully used to capture the presence of predictability in the shape of the IVS itself (e.g., Gonçalves and Guidolin, 2006). However, deterministic IVS models impose cross-sectional relations among different factors that could add noise to the analysis of our theoretical equilibrium model. Instead, we prefer the simplicity of numerical derivatives, which are calculated independently in each of the two IVS dimensions (i.e., moneyness and maturity). However, as a robustness check, in Appendix E we also assess whether a rational learning model may produce deterministic IVS estimates comparable to those commonly found in the literature.

Panels A and B of Table 2 show that, on average, learning induces negative slopes and convex IV shapes in both the moneyness and maturity dimensions. Moreover, rational learning generates kurtosis, serial correlation, and volatility clustering in both the IVs themselves and in the slope and curvature indices computed from the simulated IVS. The levels of IV along with all of the shape features that describe the IVS imply a large and significant serial correlation coefficient in more than 50% of the simulations using a first-order Box-Pierce test statistic. This means that if we observe today: i) a higher IV level; ii) a steeper negative IVS slope; or/and iii) a more convex IVS shape, they forecast: 1) high IV levels; 2) negative slopes; and/or 3) and convex IVS shapes in the future, respectively. Furthermore, the IV level, the slopes and curvatures of the IVS under learning imply on average widespread volatility clustering as measured by the percentage of significant ARCH LM tests (both with one and three lags). This means that when IV levels, slope or convexity of the IVS become variable over time, this instability tends to persist over time. However, ARCH effects are weaker in the case of the slope and curvature indices measured with respect to moneyness, $Slope_{Mon}$ and $Curve_{Mon}$, although 10% statistical significance is preserved for at least 25% of the simulations.

The simulation results presented in Panels A and B of Table 2 are consistent with the evidence reported in the literature (e.g., Harvey and Whaley, 1992; Gonçalves and Guidolin, 2006; Fengler, Härdle, and Mammen, 2007; Konstantinidi, Skiadopoulos, and Tzagkaraki, 2008). To provide evidence on the model with breaks and Bayesian learning, Panels A* and B* in Table 2 have the same structure as Panels A and B in the same table, but are no longer based on simulated option prices. Instead, Panels A* and B* concern a large set of traded, non-zero volume stock options sampled between 1997 and 2007. Panel A* concerns IV levels and IVS shape and predictability patterns measured on S&P 500 index options; Panel B* covers a set of 150 individual equity options in which the underlying stocks pay dividends.²⁶ The options data used to compute the statistics reported in Panel A* and B* are described in Appendix D. The predictability patterns in the IVS, shown in Tables 2 and 3,

²⁶ Although equity options are American-style, there is empirical evidence that they follow similar IVS dynamics as European contracts such as S&P 500 index options (Dennis and Mayhew, 2002; Goyal and Saretto, 2009; Bernales and Guidolin, 2014). In addition, possible small biases and heterogeneities across Panels A* and B* probably carry modest importance when compared to the enormous benefits we may obtain from observing the rich cross-sectional dynamic behaviors by the use of 150 different equity options.

show the strong similarities between the properties of the data and the features that emerge from the rational learning model with infrequent breaks introduced in Section 2. For instance, S&P 500 data deliver an average at-the-money, one month to maturity IV of 16.7% versus 19.2% from our simulations, under the first set of (low volatility, $\sigma = 5\%$) parameters; the slope (curvature) index in the moneyness dimension is -0.64 (13.8) in the data and -0.35 (30.9) in our simulations. Therefore the signs are always correct, although a model with learning yields IVS shapes versus moneyness that are considerably more convex than what can be detected in the S&P 500 data. Results are less accurate in the case of IVS shapes versus time-to-maturity, because the empirical slope (curvature) index in the time-to-maturity dimension is 0.03 (-0.12) in the data and -0.31 (3.28) in our simulations, and the last sign appears to be incorrect, while it is realistic that IVS be approximately flat when plotted against maturity.²⁷

It is important to notice the ability of our simple dynamic equilibrium model, with breaks and Bayesian learning, to re-produce the shape of the IVS of individual equity options, as shown by a comparison between simulated and market data. For instance, in Panels B and B* in Table 2, at least qualitatively, all the key properties of the (average) IVS from U.S. options markets are matched by our simulations. The data show an average at-the-money, one month to maturity IV of 40.3% vs. 43.3% in our simulations; the empirical, numerically computed slope (curvature) index in the moneyness dimension is -0.21 (2.62) in the data and -0.19 (2.87) in our simulations; the empirical slope (curvature) index in the time to maturity dimension is -0.04 (0.08) in the data and -0.31 (3.53) in our simulations.

The rightmost columns of all the panels in Table 2 are instead devoted to the predictability and instability of the IVS shape and level. Empirically, for both S&P 500 and individual equity options data, the average IV level tends to be highly (and positively) persistent; for instance, for 98% of the short-term ATM individual option series, the null of no serial correlation can be rejected. This is fully mimicked by our calibrated results, where for both sets of parameters in Table 2, 98.8% and 93.3% of the simulations reveal statistically significant autocorrelations in IV levels, an indication that a high positioning of the IVS

²⁷ However, below we show that when there are high levels of uncertainty (when learning speed is high in the aftermath of a structural break) in S&P 500 index options, maturity slopes tend to be negatively sloped, which is consistent with our calibration results.

today forecasts the same for the following weeks. A similar finding holds with reference to both slope and convexity indicators in the case of S&P 500 options, which is fully captured by the properties of simulated IV series from our model. Interestingly, in the case of individual equity options, there is strong evidence of serial correlation in slope and curvature of the IVS in the time-to-maturity dimension, while there is little evidence in the data to show similar phenomena compared to moneyness. The second set of calibration parameters can then reproduce such a persistence in slope and curvature indices in the maturity dimension, although it also tends to create excessive persistence in the moneyness one. Similar results appear when looking at any volatility clustering (i.e., the persistence of instability in the very shape of the IVS when rapidly changing would tend to remain so for a few consecutive weeks): the calibration in Panel A of Table 2 provides a good match for most of the results concerning S&P 500 IVS dynamics (Panel A*); interestingly, the calibration results in Panel B of Table 2 show that ARCH effects would weaken in terms of slope and convexity indices in the maturity dimension.

An agent's learning process also affects how the level of IV and the IVS shape characteristics are related to each other in a cross-sectional sense. For example, whether slope and convexity in the moneyness dimension tends to lower (i.e., the IVS is flatter) when the entire level of the IVS shifts upwards, which is empirically the case. For instance, Mixon (2007) found that the slope of at-the-money IV over different maturities has predictive ability for the level of future short-dated IV (although not to the extent predicted by a simple expectations hypothesis). To examine these interesting and delicate effects, Table 3 shows the matrices of cross-indicator simultaneous correlations in our calibrations (Panels A and B) and in market data (Panels A* and B*). Also in this case, two different sets of parameters appear in Panels A and B, while Panels A* and B* report estimated correlation coefficients for the S&P 500 and the average across 150 distinct individual stock options, respectively. Nevertheless in the Online Appendix, we report further correlation analyses using alternative parameterizations, as a robustness check.

[Insert Table 3 here]

Once more, our model with infrequent breaks and incomplete information provides an impressive fit to the properties of the IVS. Table 3 reveals a number of non-zero and

statistically significant cross-correlations. For instance, the IVS becomes flatter (the slope less negative and the smile weaker) in the moneyness dimension as well as more negatively sloped by becoming more convex in the maturity dimension, when the general level of the IVS shifts up. When the IVS becomes steeper compared to moneyness, it also becomes steeper (flatter) in the maturity dimension. A steeper negative slope as a function of maturity tends to be accompanied by less convexity (i.e., the term structure of the IVS tends to approximate a negative sloped line, that however approaches smoothly and asymptotically the zero axis for very high maturities). An inspection of the estimated correlations among these properties of the IVS in Table 3 reveals that most of these features are well captured by our model, especially in the high-volatility calibration reported in Panel B.

4 Model extension by allowing the dividend volatility to vary

In this section, we extend the model presented in Section 2 by allowing the dividend volatility to vary. We assume that dividend returns follow a GARCH(1,1) given that this type of process could reflect additional learning mechanisms followed by agents. In fact, Engle (2001, p. 160) states in relation to GARCH type models that: “Such an updating rule is a simple description of adaptive or learning behavior and can be thought of as Bayesian updating.” Thus, the dividends evolve according to the following stochastic process:

$$\ln\left(\frac{D_{t+1}}{D_t}\right) = \mu_{t+1} + \zeta_{t+1}$$

$$\zeta_{t+1} = \sigma_{t+1}z_{t+1} \tag{15}$$

$$\sigma_{t+1}^2 = \kappa + a\zeta_t^2 + b\sigma_t^2$$

where z_{t+1} are i.i.d. innovations with zero mean and unit variance. However, and similar to the model in Section 2, the fundamental mean dividend growth rate g_{t+1} (and hence μ_{t+1} given that $g_{t+1} = \exp(\mu_{t+1} + \sigma^2/2) - 1$) presents breaks and consequently changes over time, although the value of g_{t+1} is constant between break events. Time periods between breaks

follow a geometric process with parameter π ; therefore, the number of breaks in a given time window is characterized by a binomial distribution.

Under full information, stock and bond prices have the same expressions when dividends follow equation (15) as in the case when the dividend volatility is constant, since the dividend volatility does not affect stock and bond prices in our model. Therefore, stock and bond prices can be calculated using equations (6) and (7), respectively. In the case of European call option contracts, when there is full information, option prices are calculated using equation (9) by Monte Carlo methods on the basis of 20,000 independent paths. Nevertheless, Monte Carlo methods are based on a modified stochastic process like the one described in Proposition 2, in which we include the dividend volatility process characterized in equation (15).

We also relax the full information assumption to observe the effect of learning on option pricing. Consequently, we assume that g_t is unknown; however, the agent observes dividends received from the underlying asset on a daily basis, which can be used to obtain an estimated value of g_t . Thus, the agent receives signals about the mean dividend growth rate, $\{D_i/D_{i-1}\}_{i=t-n+1}^t$, which are random and follow equation (15), where n is the number of periods since the last break. We assume that the representative agent uses the information available efficiently to price all assets by following a Bayesian updating procedure. Hence, the expected value under Bayesian learning at time t , $E_{t,n}^{BL}[\cdot]$, of any asset or variable that depends on μ_t , $\lambda_t(\mu_t)$, is:

$$E_{t,n}^{BL}[\lambda_t(\mu_t)|\xi_t] = \frac{\int_{\mu_d}^{\mu_u} \lambda_t(\mu_t) L(\mu_t|\xi_t)_n f(\mu_t) d\mu_t}{\int_{\mu_d}^{\mu_u} L(\mu_t|\xi_t)_n f(\mu_t) d\mu_t} \quad (16)$$

where $f(\cdot)$ is the pdf of μ_t . However, in this case the sample likelihood function, $L(\cdot)_n$, is:

$$L(\mu_t|\xi_t)_n = \prod_{t=1}^n \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left[-\frac{(\xi_t - \mu_t)^2}{2\sigma_t^2}\right] \quad (17)$$

in which $\xi_t = [\ln(D_t/D_{t-1}) \dots \ln(D_{t-n+1}/D_{t-n})]$. Therefore, stock and bond prices are calculated using equation (16) through Monte Carlo estimations and by making $\lambda_t(\mu_t) = S_t^{CI}$

as in equation (6) and $\lambda_t(\mu_t) = B_t^{CI}$ as in equation (7), respectively. In the case of European call option contracts, option prices are obtained using Monte Carlo estimations through equation (16) and by assuming that $\lambda_t(\mu_t) = Call_t^{CI}(K, \tau)$ as in equation (9). However, as explained above, Monte Carlo methods are based on the stochastic process in which the dividend volatility follows: $\sigma_{t+1}^2 = \kappa + a\zeta_t^2 + b\sigma_t^2$.

In this model extension, we also perform an extensive set of simulations to analyze learning effects on option prices and IVs. Thus, in each combination of parameters in the model, we generate 2,000 simulations. For each of these simulations, we produce 12 years (3,024 trading days) of daily dividends, which are the signals observed by agents and thus to learn about g_t (which represents 6,048,000 simulated trading days).

We assume one of the same plausible sets of parameters described in Section 3. The assumed parameters are: $\pi=0.0030$, $\rho=8.9\%$, $g_u=8.8\%$, and $g_d=-1.5\%$, and we use different levels for the coefficient of relative risk aversion. Nevertheless, as explained above, we assume that the dividend volatility is not constant and follows a GARCH(1,1) process. We estimate the parameters of the GARCH(1,1) using daily dividend time series from the S&P 500 Index for the 1996 and 2007 period [which are deseasonalized and adjusted by the consumer price index to obtain real dividends as in Shiller (2000)]. Thus, we set $\kappa = 6 \cdot 10^{-6}$, $a = 0.31$ and $b = 0.09$. This gives an unconditional volatility of the GARCH(1,1) equal to 5%, which is consistent with the case of the constant volatility parameter described in subsection 3.2, where we also set the dividend volatility at 5%.

Table 4 presents the results of simulations from an economy with breaks and incomplete information under Bayesian learning; however, different from the results in Table 2 Panel A, in this case the dividend volatility follows a GARCH(1,1) process. The table reports time series statistics concerning the level, slope, and curvature of the IVS in both moneyness and maturity dimensions. Table 2 presents the results with a coefficient of relative risk aversion, α , at 0.2; however the analysis using other α values are available from the authors upon request. Table 4 shows that learning, when the dividend volatility varies, induces an increase in the IV (20.71%) of at-the-money short-term option contracts in relation to the case of learning and constant volatility (19.24%) reported in Table 2 Panel A, which is also

higher than the unconditional volatility of the GARCH(1,1) that is equal to 5%. The changes in the dividend volatility with learning produce a steeper negative slope (-0.39) and more curvature (35.31) on the moneyness dimension of the IVS than in the scenario of learning and constant volatility (see Table 2 Panel A). The effect of changes in the dividend volatility also generate higher skewness and kurtosis in the IV $Slope_{Mon}$ and $Curv_{Mon}$ than in the case of learning and invariable dividend volatility.

Given the fact that the main purpose of our study is to understand the predictable dynamics in the IV surface, the most important result of this model extension is that it induces stronger predictability patterns in IVS shape characteristics. Table 4 shows that serial correlation coefficients and ARCH effects are stronger for the GARCH specification than the model with constant dividend volatility described in Table 2 (Panels A and B). The improvement in the predictability patterns are observable for all IVS features (i.e., the IVS level, slopes, and curvatures). For instance, 100% and 94.50% (91.30% and 91.10%) of the simulations in Table 4 (Table 2 Panel A) have significant ARCH(1) and ARCH(3) effects. In addition, the GARCH specification helps with the fitting of market data. For example, Table 2 Panel A* shows that while S&P 500 options imply significant results for ARCH(1) and ARCH(3) tests, also 74.00% and 69.33% of the 150 equity options are characterized by ARCH(1) and ARCH(3) effects.

[Insert Table 4 here]

Interestingly, the results in Table 4 show that learning, combined with the GARCH(1,1) process on the dividend volatility generates lower slopes in magnitude and an inferior curvature in the maturity dimension of the IVS than in the case of learning and constant volatility presented in Table 2 Panel A. The intuition behind this result is simple. On the one hand, in the case of learning and constant volatility, we show in Figures 2 and 3 that in general the IVs strongly decrease as time-to-maturity increases. Immediately after a break there are intense revisions concerning the new value of g_t ; hence there is an increase in the IVs of short-term option contracts. However, the Bayesian agent expects that she will learn progressively in the future, because she will receive further information in the following periods; hence IVs are reduced as the maturity increases. On the other hand,

when the dividend volatility follows GARCH(1,1), from time to time the dividend volatility can evolve according to increasing or decreasing paths due to the cyclical behavior of the GARCH(1,1). This cyclical pattern is also anticipated by the agent and changes the shape of the IV term structure, which is reflected in the values presented in Table 4. To understand graphically the effect of the GARCH(1,1) on the IV term structure, Figure 4 shows the average behavior of IVs as a function of time-to-maturity when there is learning, and the dividend volatility is *increasing* (upper windows) and *decreasing* (lower windows) in the GARCH(1,1) process. This figure also presents the average IV term-structure *one month* (left-hand windows) and *one year* (right-hand windows) after a break in g_t .

[Insert Figure 4 here]

Figure 4 also shows that the dividend GARCH(1,1) modifies the IV term structure in relation to the results obtained by assuming the volatility of dividends to be constant. The bottom row of graphs in Figure 4 show that learning induces a more negative slope and stronger curvature on the IV term structure, when the dividend volatility follows a decreasing pattern in the GARCH(1,1) process than in the case of constant dividend volatility (see Figures 2 and 3). Nevertheless, an increasing tendency of the dividend volatility generates a particular outcome. For instance, the upper left-hand window in Figure 4 shows that after a month since a break, the IV term structure has a smile shape. Immediately after a break there is an increase in the IVs of short-term contracts; however IVs are reduced as the time-to-maturity increases to reach a minimum in the middle-term (90 days) because the agent “expects” to learn when she will receive more information in the future. Afterwards, the increasing pattern of the GARCH(1,1) process followed by the dividend volatility generates a growing path on IVs obtained from long-term option contracts.

The impact of the increasing pattern of the dividend volatility on the IV term structure is more evident when learning effects are reduced. For example, the upper right-hand window in Figure 4 shows a positive slope on the IV term structure, when learning has reached a high accuracy regarding the unknown value of g_t (because a year has elapsed since the last break). This is an important improvement in terms of flexibility in relation to the model with constant dividend volatility, in which we normally find IVs that are decreasing

functions of time-to-maturity. For instance, we show in Table 2 Panel A* (Panel B*) that the IV term structure has on average a positive (negative) slope for index options (equity options).

Furthermore, the agent's learning process also affects the cross-sectional relations between the level of IV and the IVS shape characteristics when the dividend volatility follows a GARCH(1,1). Table 5 presents the results of a correlation analysis for the level, slope, and curvature of the IVS in both the moneyness and maturity dimensions, when there is learning and the dividend volatility evolves according to equation (15). Similar to Table 3, Table 5 shows that learning induces strong relations between the level of the IV, the slope, and the curvature on the moneyness dimension. However, the number of simulations in Table 5 with significant correlations is lower than in the case of constant dividend volatility (see Panels A and B in Table 3) for the slope and curvature on the maturity dimension. The intuition behind the reduction in the number of simulations — with significant correlations in Table 5 between the IV term structure features and other IVS variables — is related to our explanations for Figure 4. Figure 4 shows that after a break and when the dividend volatility increases in the GARCH(1,1) process, learning induces an augmentation in the IV of short-term contracts, which is reduced as the maturity increases, since the agent expects to receive more information and thus to have more accuracy regarding the fundamental dividend growth rate. These opposite effects, when there is learning and the dividend volatility is rising, indicate that the relations between the IV term structure and other implied variables are slightly reduced.

[Insert Table 5 here]

We conclude that although far from perfect, even a simple equilibrium asset pricing model such as ours (with and without constant dividend volatility) has explanatory power for the predictable dynamics in the IV surface, which is our main contribution. Our focus is providing a rational explanation for the predictability patterns in the IVS, rather than calibrating the shape of the IVS itself. However, our model is still able to characterize key properties of option prices and IVs [as shown to some extent already in static analysis by GT (2003) and David and Veronesi (2002)].

5 Conclusions

The fact that option prices and especially (BS) IVs are predictable has been identified by academics and exploited by practitioners in a number of valuable applications, from hedging to trading. Nevertheless, there is a gap in the literature regarding possible explanations for such puzzling predictability patterns. In fact, under the simplest option pricing benchmark, the ineffable Black-Scholes pricing models and their simple extensions, the IVS would be flat and not moving over time. Also more complex pricing frameworks that go well beyond the simplistic assumptions underlying BS are usually silent about the specific features of the dynamics in the IVS. In this paper, we contribute to this body of literature by providing evidence to support the hypothesis that the investors' learning impacts the dynamic predictability patterns characterizing the IVS.

We present an equilibrium model in which the fundamental mean dividend growth rate (the drift of the corresponding stochastic process) is subject to infrequent but observable breaks, which therefore only occur with a small probability. Under incomplete information, which represents the realistic description of the world in which deep parameters and mean growth rates may at best be estimated, an agent receives independent but noisy daily signals about the unknown fundamental value that are used to update her beliefs using a Bayesian updating algorithm.

We show that an investor's learning process may cause the typical predictability patterns for option prices and the IVS reported in the literature. Rational learning makes agents' beliefs time-varying and in equilibrium this leads to sizable dynamic risk premia that affect both option prices and the movements of the (BS) volatilities implicit in such prices (Mixon, 2007). Moreover, our modeling approach shows that learning generates heterogeneous dynamic properties for option contracts depending of their moneyness and residual maturity; these heterogeneous effects, due to the complex shape of the perceived, time-varying pricing kernel under rational learning, are responsible for the (non-flat) and predictable shape of the IVS.

One issue we want to stress is the role played by our study. It would, of course, be naive to think that a simple model may describe all salient aspects of the option market. However, our concern is not to capture every single feature of option markets as closely as possible. Instead, our interest is to understand whether our learning model can "explain the predictable dynamics in the implied volatility surface," Of course, we want to characterize all option market features in a reasonable way, but our model seems to be a satisfactory starting point in understanding the predictable dynamics of the IVS.

One aspect of our model may also teach a more general lesson to scholars in the field. Option markets have been widely used to capture forward-looking information since they reflect agents' expectations about future scenarios, in which forecasting horizons match the expiration dates of options contracts (Fleming, 1998). The information captured from option prices has been used by investors in a range of markets and with applications to a broad spectrum of financial issues including risk management, asset allocation, and capital budgeting.²⁸ However, through our model we also show that option markets may display not only the classical, forward-looking features, but they may at the same time be affected by backward-looking characteristics since option traders also need to learn recursively as new information arrives. Participants in option markets face a sequential process of information acquisition in which signals are received and processed with reference to historical information and prior beliefs. Therefore, the forward-looking information obtained from option markets is generated by a backward-looking learning process. Such tight intertwining between backward- and forward-looking information processing imposes useful restrictions that ought to be carefully considered and tested when models of option pricing are quantitatively assessed.

Finally, our model is simple and intuitive. Yet, it may be extended in a variety of directions while at the same time preserving its key intuition that any dynamics in the IVS may be consistent with no arbitrage pricing restrictions, and derive from time-varying risk premia. These premia compensate investors for the additional risk deriving from belief revisions at

²⁸ For instance, option prices have been used to forecast underlying returns (Xing, Zhang, and Zhao, 2010; Cremers and Weinbaum, 2010; Bakshi, Panayotov, and Skoulakis, 2011), realized volatilities (e.g., Christensen and Prabhala, 1998; Busch et al., 2011), betas (e.g., Siegel, 1995; Chang *et al.*, 2009), correlations (e.g., Driessen, Maenhout, and Vilkov, 2009), and to estimate the moments required in standard asset allocation problems (e.g., Kostakis, Panigirtzoglou, and Skiadopoulos, 2011; DeMiguel, Plyakha, Uppal, and Vilkov, 2012).

a varying speed due to the infrequent occurrence of breaks. For instance, the model could be used to isolate the portion of variation in the IVS that is due to rational pricing factors from those that are simply irrational (e.g., Kim and Lee, 2013). Moreover, researchers have investigated the properties of option returns (Broadie, Chernenkov, and Johannes, 2009) and detected a number of difficulties in solving anomalies with reference to standard asset pricing frameworks. It would be interesting to further investigate how rational learning would affect such pricing frameworks and whether this may teach us something about the nature of option returns.

References

- Abel, A., 1988. Stock prices under time-varying dividend risk: An exact solution in an infinite-horizon general equilibrium model. *J. Monet. Econ.* 22, 375-393.
- Bai, J., R. Lumsdaine, and J. Stock, 1998. Testing for and dating common breaks in multivariate time series. *Rev. Econ. Stud.* 63, 395-432.
- Bakshi, G., G. Panayotov, and G. Skoulakis, 2011. Improving the predictability of real economic activity and asset returns with forward variances inferred from option portfolios. *J. Financ. Econ.* 100, 475-495.
- Beber, A., and M. W. Brandt, 2006. The effect of macroeconomic news on beliefs and preferences: Evidence from the options market. *J. Monet. Econ.* 53, 1997-2039.
- Beber, A., and M. W. Brandt, 2009. Resolving macroeconomic uncertainty in stock and bond markets. *Rev. Financ.* 13, 1-45.
- Bernales, A., and M. Guidolin, 2014. Can we forecast the implied volatility surface dynamics of equity options? Predictability and economic value tests. *J. Bank. Finance* 46, 326-342.
- Black, F., and M., Scholes, 1973. The pricing of options and corporate liabilities. *J. Polit. Econ.* 81, 637-654.
- Brennan, M., and H. Cao, 1996. Information, trade, and derivative securities. *Rev. Financ. Stud.* 9, 163-208.
- Broadie, M., M., Chernenkov, and M. Johannes, 2009. Understanding index option returns. *Rev. Financ. Stud.* 22 4493-4529.
- Cecchetti, S. G., P. Lam, and N. C. Mark, 1990. Mean reversion in equilibrium asset prices. *Am. Econ. Rev.* 80, 221-242.
- Chalamandaris, G., and A.E. Tsekrekos, 2010. Predictable dynamics in implied volatility surfaces from OTC currency options. *J. Bank. Finance* 34, 1175-1188.
- Chang, B. Y., P. Christoffersen, K. Jacobs, and G. Vain Berg, 2009. Option-implied measures of equity risk. Working paper, McGill University.

- Christoffersen, P.F., S. Heston, and K. Jacobs, 2009. The shape and term structure of the index option smirk: Why multifactor stochastic volatility models work so well. *Manage. Sci.* 55, 1914-1932.
- Chu, C. S. J. , M. Stinchcombe, and H. White, 1996. Monitoring structural change. *Econometrica* 64, 1045-1065.
- Das, S., and R. Sundaram, 1999. Of smiles and smirks: A term structure perspective. *J. Financ. Quant. Anal.* 34, 211-239.
- David, A., 2008. Heterogeneous beliefs, speculation, and the equity premium. *J. Finance* 63, 41-83.
- David, A., and P. Veronesi, 2002. Option prices with uncertain fundamentals. Working paper, University of Chicago.
- David, A., and P. Veronesi, 2013. What ties return volatilities to price valuations and fundamentals?. *J. Polit. Econ.* 121, 682 - 746.
- DeMiguel, V., Y., Plyakha, R., Uppal, and G. Vilkov, 2012. Improving portfolio selection using option-implied volatility and skewness. *J. Financ. Quant. Anal.*, forthcoming.
- Donders, M., R. Kouwenberg, and T. Vorst, 2000. Options and earnings announcements: An empirical study of volatility, trading volume, open interest, and liquidity. *Eur. Financ. Manag.* 6, 149-172.
- Driessen, J., P. J. Maenhout, and G. Vilkov, 2009. The price of correlation risk: Evidence from equity options. *J. Finance* 64, 1377-1406.
- Duan, J. C., and J. G. Simonato, 2001. American option pricing under GARCH by a Markov chain approximation. *J. Econ. Dyn. Control* 25, 1689-1718.
- Dubinsky, A., and M. Johannes, 2006. Earnings announcements and equity options. Working paper, Columbia University.
- Dumas, B., J. Fleming, and R. Whaley, 1998. Implied volatility functions: Empirical tests. *J. Finance* 53, 2059-2106.
- Ederington, L., and J. H. Lee, 1996. The creation and resolution of market uncertainty: The impact of information releases on implied volatility. *J. Financ. Quant. Anal.* 31, 513-539.
- Engle, R.F., 2001. GARCH 101: The use of ARCH/GARCH models in applied econometrics. *J. Econ. Perspect.* 15, 157-168.
- Fengler, M.R., W.K. Härdle, and E. Mammen, 2007. A semiparametric factor model for implied volatility surface dynamics. *J. Financ. Economet.* 5, 189-218.
- Fleming J., 1998. The quality of market volatility forecasts implied by S&P 100 index option prices. *J. Empir. Financ.* 5, 317-345.
- Gonçalves, S., and M. Guidolin, 2006. Predictable dynamics in the S&P 500 index options implied volatility surface. *J. Bus.* 79, 1591-1635.
- Granger, C. and N. Hyung, 2004. Occasional structural breaks and long memory with an application to the S&P 500 absolute stock returns. *J. Empir. Financ.* 3, 399-421.
- Guidolin, M., 2006. High equity premia and crash fears. Rational foundations. *Econ. Theor.* 28, 693-708.

- Guidolin, M., and A. Timmermann, 2003. Option prices under Bayesian learning: Implied volatility dynamics and predictive densities. *J. Econ. Dyn. Control* 27, 717-769.
- Guidolin, M., and A. Timmermann, 2007. Properties of equilibrium asset prices under alternative learning schemes. *J. Econ. Dyn. Control* 31, 161-217.
- Harvey, C. R., and R. E. Whaley, 1992. Market volatility prediction and the efficiency of the S&P 100 index option market. *J. Financ. Econ.* 31, 43-73.
- Heston, S., and S. Nandi, 2000. A closed-form GARCH option valuation model. *Rev. Financ. Stud.* 13, 585-625.
- Huang, C. F., and R. H. Litzenberger, 1988. *Foundations for Financial Economics*. first ed. North-Holland, The Netherlands.
- Kim N., and J. Lee, 2013. No-arbitrage implied volatility functions: Empirical evidence from KOSPI 200 index options. *J. Empir. Financ.* 21, 36-53.
- Konstantinidi, E., G. Skiadopoulos, and E. Tzagkaraki, 2008. Can the evolution of implied volatility be forecasted? Evidence from European and US implied volatility indices. *J. Bank. Finance* 32, 2401-2411.
- Kostakis, A., N. Panigirtzoglou, G. Skiadopoulos, 2011. Market timing with option-implied distributions: A forward-looking approach. *Manage. Sci.* 57, 1231-1249..
- Leisch, F., K. Hornik, and C. M. Kuan, 2000. Monitoring structural changes with the generalized fluctuation test. *Economet. Theor.* 16, 835-854.
- Lettau, M., and S. Van Nieuwerburgh, 2008. Reconciling the return predictability evidence. *Rev. Financ. Stud.* 21, 1607-1652.
- Lucas, R., 1978, Asset prices in an exchange economy, *Econometrica* 46, 1429-1445.
- Mixon, S., 2007. The implied volatility term structure of stock index options. *J. Empir. Financ.* 14, 333-354.
- Ni, S. X., J. Pan, and A. M. Poteshman, 2008. Volatility information trading in the option market. *J. Finance* 63, 1059-1091.
- Pastor, L., and R. F. Stambaugh, 2001. The equity premium and structural breaks. *J. Finance* 56, 1207-1239.
- Pliska, S. R., 1997. *Introduction to Mathematical Finance*. First ed. Blackwell, Oxford.
- Rubinstein, M., 1985. Nonparametric tests of alternative option pricing models using all reported trades and quotes on the 30 most active CBOE option classes from August 23, 1976 through August 31, 1978. *J. Finance* 40, 455-480.
- Rubinstein, M., 1994. Implied binomial trees. *J. Finance* 49, 771-818.
- Shaliastovich, I., 2015. Learning, confidence, and option prices. *J. Econometrics* 187, 18-42.
- Shiller, R., 2000. *Irrational Exuberance*. first ed. Princeton University Press, Princeton.
- Timmermann, A., 1993. How learning in financial markets generates excess volatility and predictability of excess returns. *Q. J. Econ.* 108, 1135-1145.
- Timmermann, A., 1996. Excess volatility and predictability of stock prices in autoregressive dividend models with learning. *Rev. Econ. Stud.*, 63, 523-557.
- Timmermann, A., 2001. Structural breaks, incomplete information, and stock prices. *J. Bus.*

Econ. Stat., 19, 299-314.

Veronesi, P., 1999. Stock market overreaction to bad news in good times: A rational expectations equilibrium model. *Rev. Financ. Stud.* 12, 975-1007.

Veronesi, P., 2000. How does information quality affect stock returns?. *J. Finance*, 55, 807-837.

Xing, Y., X. Zhang, and R. Zhao, 2010. What does individual option volatility smirk tell us about future equity returns. *J. Financ. Quant. Anal.* 45, 641-662.

Yan, S., 2011. Jump risk, stock returns, and slope of implied volatility smile. *J. Financ. Econ.* 99, 216-233.

Appendix A: Proofs

Proof of Proposition 1: Assuming that the expression that describes S_t^{CI} can be written as $S_t^{CI} = D_t \Psi(g_t)$ for some function $\Psi_t^{CI}(\cdot)$, we define a “break indicator,” s_t , that signals the occurrence of breakpoints in the mean dividend growth rate. In the case in which there is no break in g_{t+1} $s_{t+1} = s_t$; if instead $s_{t+1} = s_t + 1$, then a break has taken place at $t + 1$. Additionally, $\Pr(s_{t+1} = s_t) = (1 - \pi)$ is the probability of no break, and the probability of a break out of the state prevailing at time t is $\Pr(s_{t+1} = s_t + 1) = \pi$. Therefore, from equation (4):

$$\begin{aligned}
 (1 + \rho)\Psi(g_t)D_t &= \sum_{i=0}^1 E_t [(\Psi_t^{CI}(g_t)D_{t+1} + D_{t+1}) \left(\frac{D_{t+1}}{D_t}\right)^{-\alpha} | s_{t+1} = s_t + i] \Pr(s_{t+1} = s_t + i) \\
 &= (1 - \pi)D_t \int_{-\infty}^{\infty} (1 + \Psi_t^{CI}(g_t)) (1 + g_t)^{1-\alpha} \exp\left((1 - \alpha)\left(\sigma\varepsilon_{t+1} - \frac{\sigma^2}{2}\right)\right) \\
 &\quad \cdot \phi(\varepsilon_{t+1}|\sigma^2) d\varepsilon_{t+1} + (1 - e^{-\pi})D_t \int_{g_d}^{g_u} \int_{-\infty}^{\infty} (1 + \Psi_t^{CI}(g_{t+1})) (1 + g_{t+1})^{1-\alpha} \\
 &\quad \cdot \exp\left((1 - \alpha)\left(\sigma\varepsilon_{t+1} - \frac{\sigma^2}{2}\right)\right) \phi(\varepsilon_{t+1}|\sigma^2) d\varepsilon_{t+1} dG(g_{t+1}), \tag{A1}
 \end{aligned}$$

where $G(\cdot)$ is the c.d.f. of g_{t+1} defined on $[g_d, g_u]$, ε_{t+1} is the innovation term of the geometric random walk process for dividends, and $\phi(\cdot|\sigma)$ is a normal density function with mean zero and variance σ . Therefore, given that ε_{t+1} and g_{t+1} are independent, we can rewrite equation (A1) as:

$$\begin{aligned}
 (1 + \rho)\Psi(g_t)D_t &= (1 - \pi)D_t(1 + g_t)^{1-\alpha}(1 + \Psi_t^{CI}(g_t)) \\
 &\quad + \pi D_t \int_{g_d}^{g_u} (1 + g_{t+1})^{1-\alpha} dG(g_{t+1}) + \pi D_t \int_{g_d}^{g_u} \Psi_t^{CI}(g_{t+1})(1 + g_{t+1})^{1-\alpha} dG(g_{t+1}) \tag{A2}
 \end{aligned}$$

-or equivalently,

$$\begin{aligned}
D_t \Psi(g_t) &= (1 - \pi) \frac{D_t}{1 + \rho - (1 - \pi)(1 + g_t)^{1-\alpha}} (1 + g_t)^{1-\alpha} \\
&+ \pi \frac{D_t}{1 + \rho - (1 - \pi)(1 + g_t)^{1-\alpha}} \int_{g_d}^{g_u} (1 + g_{t+1})^{1-\alpha} dG(g_{t+1}) \\
&+ \pi \frac{D_t}{1 + \rho - (1 - \pi)(1 + g_t)^{1-\alpha}} \int_{g_d}^{g_u} \Psi_t^{CI}(g_t) (1 + g_{t+1})^{1-\alpha} dG(g_{t+1}).
\end{aligned} \tag{A3}$$

Because $G(\cdot)$ does not vary over time in equation (A3), we multiply both sides by $(1 + g_{t+1})^{1-\alpha} dG(g_{t+1})/D_t$ and integrate over $[g_d, g_u]$, to obtain:

$$\begin{aligned}
\int_{g_d}^{g_u} \Psi(g_t) (1 + g_t)^{1-\alpha} dG(g_t) &= \int_{g_d}^{g_u} (1 - \pi) \frac{(1 + g_t)^{2-2\alpha}}{1 + \rho - (1 - \pi)(1 + g_t)^{1-\alpha}} dG(g_t) \\
&+ \int_{g_d}^{g_u} \pi \frac{(1 + g_{t+1})^{1-\alpha} \int_{g_d}^{g_u} (1 + g_{t+1})^{1-\alpha} dG(g_{t+1})}{1 + \rho - (1 - \pi)(1 + g_{t+1})^{1-\alpha}} dG(g_{t+1}) \\
&+ \int_{g_d}^{g_u} \pi \frac{(1 + g_{t+1})^{1-\alpha}}{1 + \rho - (1 - \pi)(1 + g_{t+1})^{1-\alpha}} dG(g_{t+1}) \\
&\cdot \int_{g_d}^{g_u} \Psi(g_{t+1}) (1 + g_{t+1})^{1-\alpha} dG(g_{t+1}).
\end{aligned} \tag{A4}$$

The term on the left-hand side is equal to the second part of last term on the right-hand side, and consequently:

$$\begin{aligned}
&\int_{g_d}^{g_u} \Psi(g_{t+1}) (1 + g_{t+1})^{1-\alpha} dG(g_{t+1}) \\
&= \left(\int_{g_d}^{g_u} (1 - \pi) \frac{(1 + g_{t+1})^{2-2\alpha}}{1 + \rho - (1 - \pi)(1 + g_{t+1})^{1-\alpha}} dG(g_{t+1}) \right. \\
&+ \left. \int_{g_d}^{g_u} \pi \frac{(1 + g_{t+1})^{1-\alpha} \int_{g_d}^{g_u} (1 + g_{t+1})^{1-\alpha} dG(g_{t+1})}{1 + \rho - (1 - \pi)(1 + g_{t+1})^{1-\alpha}} dG(g_{t+1}) \right) \\
&/ \left(1 - \int_{g_d}^{g_u} \pi \frac{(1 + g_{t+1})^{1-\alpha}}{1 + \rho - (1 - \pi)(1 + g_{t+1})^{1-\alpha}} dG(g_{t+1}) \right).
\end{aligned} \tag{A5}$$

Finally, inserting equation (A5) into equation (A3), we have:

$$\begin{aligned}
S_t^{CI} &= \frac{D_t}{1 + \rho - (1 - \pi)(1 + g_t)^{1-\alpha}} \{(1 - \pi)(1 + g_t)^{1-\alpha} \\
&+ \pi \int_{g_d}^{g_u} (1 + g_{t+1})^{1-\alpha} dG(g_{t+1}) + \pi \int_{g_d}^{g_u} (1 - \pi) \frac{(1 + g_{t+1})^{2-2\alpha}}{1 + \rho - (1 - \pi)(1 + g_{t+1})^{1-\alpha}} dG(g_{t+1}) \\
&+ \int_{g_d}^{g_u} \pi \frac{(1 + g_{t+1})^{1-\alpha} \int_{g_d}^{g_u} (1 + g_{t+1})^{1-\alpha} dG(g_{t+1})}{1 + \rho - (1 - \pi)(1 + g_{t+1})^{1-\alpha}} dG(g_{t+1})) \\
&/ (1 - \int_{g_d}^{g_u} \pi \frac{(1 + g_{t+1})^{1-\alpha}}{1 + \rho - (1 - \pi)(1 + g_{t+1})^{1-\alpha}} dG(g_{t+1}))\}.
\end{aligned} \tag{A6}$$

The integrals in equation (A6) are constant over time and can be labeled as:

$$I_1 = \int_{g_d}^{g_u} (1 + g_{t+1})^{1-\alpha} dG(g_{t+1}) \tag{A7}$$

$$I_2 = \int_{g_d}^{g_u} \frac{(1 + g_{t+1})^{2-2\alpha}}{1 + \rho - (1 - \pi)(1 + g_{t+1})^{1-\alpha}} dG(g_{t+1}) \tag{A8}$$

$$I_3 = \int_{g_d}^{g_u} \frac{(1 + g_{t+1})^{1-\alpha}}{1 + \rho - (1 - \pi)(1 + g_{t+1})^{1-\alpha}} dG(g_{t+1}) \tag{A9}$$

and as a result one can see that:

$$\begin{aligned}
S_t^{CI} &= \frac{D_t}{1 + \rho - (1 - \pi)(1 + g_t)^{1-\alpha}} \{(1 - \pi)(1 + g_t)^{1-\alpha} + \pi I_1 \\
&+ \pi \left(\frac{(1 - \pi)I_2 + \pi I_1 I_3}{1 - \pi I_3} \right)\} \\
&= \frac{D_t}{1 + \rho - (1 - \pi)(1 + g_t)^{1-\alpha}} \{(1 - \pi)(1 + g_t)^{1-\alpha} + \pi \left(\frac{I_1 + (1 - \pi)I_2}{1 - \pi I_3} \right)\}.
\end{aligned} \tag{A10}$$

Therefore, equation (A10) shows that $S_t^{CI} = D_t \Psi_t^{CI}(g_t)$, as we state in equation (A1).

In the case of the bond, we use the second Euler equation (5) to obtain:

$$\begin{aligned}
B_t^{CI} &= \frac{1}{(1 + \rho)} \sum_{i=0}^1 E_t \left[\left(\frac{D_{t+1}}{D_t} \right)^{-\alpha} |s_{t+1} = s_t + i\right] \Pr(s_{t+1} = s_t + i) \\
&= \frac{1}{(1 + \rho)} \{(1 - \pi) \int_{-\infty}^{\infty} (1 + g_t)^{-\alpha} \exp(-\alpha(\sigma \varepsilon_{t+1} - \frac{\sigma^2}{2})) \phi(\varepsilon_{t+1} | \sigma^2) d\varepsilon_{t+1} + \pi
\end{aligned}$$

$$\begin{aligned}
& \int_{g_d}^{g_u} \int_{-\infty}^{\infty} (1 + g_{t+1})^{-\alpha} \exp\left(-\alpha\left(\sigma\varepsilon_{t+1} - \frac{\sigma^2}{2}\right)\right) \phi(\varepsilon_{t+1}|\sigma^2) d\varepsilon_{t+1} dG(g_{t+1}) \\
& = \frac{1}{(1 + \rho)} \left\{ (1 - \pi)(1 + g_t)^{-\alpha} + \pi \int_{g_d}^{g_u} (1 + g_{t+1})^{-\alpha} dG(g_{t+1}) \right\}.
\end{aligned} \tag{A11}$$

The last equality derives from the fact that $G(\cdot)$ does not vary over time and by the independence of ε_{t+1} and g_{t+1} . \square

Proof of Proposition 2: The result can be obtained from no-arbitrage arguments applied to a European contingent claim with terminal value given by $\max\{S_{t+\tau}^{CI} - K\}$, when the mean dividend growth rate underlying the pricing of $S_{t+\tau}^{CI}$ is subject to breaks. Therefore, it is necessary to proceed to the risk neutralization of the probabilities that enter the state price density. Following Huang and Litzenberger (1988, p. 229), we take the Euler equation (4) and divide it by the price of a one-period zero-coupon bond:

$$\begin{aligned}
& \frac{(1 + \rho)S_{t+k}^{CI}}{(1 - \pi)(1 + g_{t+k})^{-\alpha} + \pi \int_{g_d}^{g_u} (1 + g_{t+k})^{-\alpha} dG(g_{t+k})} = E_{t+k} \left[\beta \left(\frac{D_{t+k+1}}{D_{t+k}} \right)^{-\alpha} \right. \\
& \left. \cdot (S_{t+k+1}^{CI} + D_{t+k+1}) \frac{(1 + \rho)}{(1 - \pi)(1 + g_{t+k})^{-\alpha} + \pi \int_{g_d}^{g_u} (1 + g_{t+k})^{-\alpha} dG(g_{t+k})} \right].
\end{aligned} \tag{A12}$$

It turns out that the forward price and the forward cumulative dividend process are:

$$S_{t+k}^{CI*} = \frac{(1 + \rho)S_{t+k}^{CI}}{(1 - \pi)(1 + g_{t+k})^{-\alpha} + \pi \int_{g_d}^{g_u} (1 + g_{t+k})^{-\alpha} dG(g_{t+k})} \tag{A13}$$

and

$$D_{t+k}^* = \sum_{s=0}^k D_{t+s} \frac{(1 + \rho)}{(1 - \pi)(1 + g_{t+s})^{-\alpha} + \pi \int_{g_d}^{g_u} (1 + g_{t+s})^{-\alpha} dG(g_{t+s})}. \tag{A14}$$

In addition, we know from the Euler condition that the pricing kernel must be such that:

$$E_t \left[\beta \left(\frac{D_{t+1}}{D_t} \right)^{-\alpha} \frac{(1 + \rho)}{(1 - \pi)(1 + g_{t+k})^{-\alpha} + \pi \int_{g_d}^{g_u} (1 + g_{t+k})^{-\alpha} dG(g_{t+k})} \right] = 1. \tag{A15}$$

Using equation (A15) and adding D_{t+k}^{CI*} to both sides of equation (A12), we obtain:

$$\begin{aligned}
& S_{t+k}^{CI*} + D_{t+k}^* \\
&= E_{t+k} \left[\beta \left(\frac{D_{t+k+1}}{D_{t+k}} \right)^{-\alpha} \frac{(1+\rho)}{(1-\pi)(1+g_{t+k})^{-\alpha} + \pi \int_{g_d}^{g_u} (1+g_{t+k})^{-\alpha} dG(g_{t+k})} (S_{t+k+1}^{CI*} \right. \\
&\quad \left. + D_{t+k+1}^*) \right].
\end{aligned} \tag{A16}$$

This shows that $(S_{t+k}^{CI*} + D_{t+k}^*)$ follows a martingale under this conditional probability measure so that the risk-neutral density is:

$$\hat{p}_t(S_{t+k}^{CI}) = \beta \left(\frac{D_{t+1}}{D_t} \right)^{-\alpha} \frac{(1+\rho)}{(1-\pi)(1+g_{t+k})^{-\alpha} + \pi \int_{g_d}^{g_u} (1+g_{t+k})^{-\alpha} dG(g_{t+k})} p_t(D_{t+k}). \tag{A17}$$

Consequently, the one-period state-price density can be written as:

$$\begin{aligned}
\tilde{p}_t(S_{t+k}^{CI}) &= \beta \left(\frac{D_{t+k}}{D_t} \right)^{-\alpha} \frac{1}{1+r_{t+k}^{CI}} \cdot \frac{(1+\rho)}{(1-\pi)(1+g_{t+k})^{-\alpha} + \pi \int_{g_d}^{g_u} (1+g_{t+k})^{-\alpha} dG(g_{t+k})} p_t(D_{t+k}) = \\
&\beta \left(\frac{D_{t+1}}{D_t} \right)^{-\alpha} p_t(D_{t+k}) \quad ,
\end{aligned} \tag{A18}$$

where r_{t+k}^{CI} is the one-period risk-free interest rate. Additionally, Pliska (1997) shows that if the risk-neutral measure on a single period model is unique and exists, this is a sufficient condition to have a unique risk-neutral measure on an infinite period model obtained as a repetition of many static, single-period models. In our case, the infinite period model risk-neutral measure can be obtained using the independence of breaks on the mean dividend growth rates and by taking all paths that could guide to a particular state in $t + \tau$ periods. In this context, $p_t^{CI}(S_{t+\tau}^{CI})$ is the state price density of all paths that lead to the state in which the dividend is $D_{t+\tau}$, while the expected value of $D_{t+\tau}$ is:

$$E_t[D_{t+\tau}] = D_t E_t \left[\frac{D_{t+1}}{D_t} E_{t+1} \left[\left(\frac{D_{t+2}}{D_{t+1}} \right) \dots E_{t+\tau-1} \left[\left(\frac{D_{t+\tau}}{D_{t+\tau-1}} \right) \right] \right] \right]. \tag{A19}$$

Furthermore, using the independence of $\{\varepsilon_{t+i}\}_{i=1}^{\tau}$ and $\{g_{t+i-1}\}_{i=1}^{\tau}$ we have:

$$E_t[D_{t+\tau}] = D_t E_t \left[\exp(\sqrt{\tau} \sigma \varepsilon_{t+\tau} - \tau \sigma^2 / 2) \prod_{i=1}^{\tau} (1 + g_{t+i-1}) \right]. \tag{A20}$$

At this point, let z be the number of breaks between t and $t + \tau$; this is a random variable drawn from a binomial distribution, $\varphi(z|\tau, \pi)$, with parameters τ and π ; $\{h_i\}_{i=0}^z$ are the time

periods between breaks, which are also random variables that follow geometric distributions with parameter π , $\eta(h_i|\pi)$, where $\tau = \sum_{i=0}^z h_i$. Then, on each path:

$$D_{t+\tau}^{CI} = D_t \exp(\sqrt{\tau}\sigma\varepsilon_{t+\tau} - \tau\sigma^2/2) \cdot \prod_{i=1}^{z+1} (1 + g_{t+v_{i-1}})^{h_i}, \quad (\text{A21})$$

where $\{g_{t+h_i}\}_{i=1}^z$ are drawn from a univariate density $g_{t+h_{i-1}} \sim G(\cdot)$ and pdf $\varrho(g_{t+h_i})$ defined on the support $[g_d, g_u]$, while $g_{t+h_0} = g_t$ and $g_{t+\tau} = g_{t+h_z}$. Consequently,

$$p_t(D_{t+k}) = \phi(\varepsilon_{t+\tau}|0, \sigma)\varphi(z|\tau, \pi)\eta(h_0|\pi) (\eta(h_1|\pi)\varrho(g_{t+h_1}) \cdot \dots \cdot \eta(h_z|\pi)\varrho(g_{t+h_z})) \quad (\text{A22})$$

and thus from equation (A18), we have:

$$\tilde{p}_t(S_{t+k}^{CI}) = \beta^\tau \left(\frac{D_{t+\tau}}{D_t}\right)^{-\alpha} \phi(\varepsilon_{t+\tau}|0, \sigma)\varphi(z|\tau, \pi)\eta(h_0|\pi) (\eta(h_1|\pi)\varrho(g_{t+h_1}) \cdot \eta(h_z|\pi)\varrho(g_{t+h_z})). \quad (\text{A23}) \quad \square$$

Appendix B: Replication of CBOE rules for data generated by model simulations

In previous studies (e.g., Duan and Simonato, 2001; Yan, 2011), options data have been simulated assuming constant moneyness and time-to-maturity for a specific option contract (e.g., exactly 30-day to expiry contracts with a moneyness exactly equal to one).²⁹ Clearly, on the one hand, the moneyness ratio changes constantly because strike prices are fixed by option exchanges while stock prices vary over time. On the other hand, the time-to-maturity decreases gradually since expiration dates are also fixed. Therefore, the assumption of regular and invariable features of the traded option contracts in a simulation exercise is not consistent with actual options data. Our goal is to investigate whether rational learning may reproduce a range of small-sample results generated from standard econometric tests applied to actual data. Therefore to generate a cross-section (across strikes and maturities) of time series of option prices in a realistic way plays a crucial role, in case the null hypothesis of learning not being fundamentally responsible for the reported stylized facts

²⁹ We define moneyness as $Mon \equiv K/S$, where K and S are the strike and the underlying stock prices, respectively.

were not to be rejected.³⁰ Instead, to increase the realism as well as the reliability of our results, we follow the detailed rules of the CBOE.

In particular, firstly, we use the same trading dates that were effectively listed over the 1996 – 2007 sample, thus accounting for holidays and unexpected events in which the market was closed.³¹ Secondly, we fix expiration dates for option contracts in the same way as the CBOE was doing between 1996 and 2007. Therefore, for contracts to be offered in a given month, the expiration dates are set to coincide with the three subsequent months followed by three additional long-term maturities aligned on the March quarterly cycle (i.e., March, June, September, and December). In addition, expiration dates fall in line with the Saturday after the third Friday of each expiration month. Thirdly, strike price intervals are set around the underlying asset price; contracts with expiration dates in the three near-term months are spaced at five-point intervals around the underlying index price as of the day in which contracts are offered, while contracts with long-term expirations are spaced at 25-point intervals.

Appendix C: Chu, Stinchcombe, and White’s real time breakpoint test

We use the test introduced by Chu, Stinchcombe, and White (1996) to estimate the probability of breaks in the mean dividend growth rate. The authors present a dynamic test for structural breaks through which market participants can identify a real time shift in the (conditional) mean function. Consider the dividend random walk process in equation (1). Let v be the minimum number of periods over which the drift is assumed to be constant, given that n is the number of periods from the most recent break (i.e., $\mu_{t-n+1} = \mu_{t-n+2} = \dots = \mu_{t-n+v}$). Assuming that the representative agent starts detecting the presence of breaks after a period v , the authors propose the use of the following *fluctuation detector* in the case of a univariate location (mean function) model:

$$\widehat{Z}_t = n\widehat{s}_0^{-1}(\widehat{\mu}_t - \widehat{\mu}_v), \quad (\text{B1})$$

³⁰ This means that we want to minimize the chances of Bayesian learning explaining option pricing stylized features and puzzles, but this hypothesis is rejected because prices are simulated following over-simplistic rules that make simulated results not perfectly comparable to the ones obtained from the data, which are instead generated following CBOE rules.

³¹ However, for simplicity such infrequent and unexpected events (e.g., September 11, 2001) are not simulated and are held fixed throughout all simulation trials.

where $\hat{\mu}_t$ and $\hat{\mu}_v$ are the parameter estimates at time t and v . We defined ξ_t as the vector of signals about μ_t in equation (10); therefore $\hat{\mu}_t = \bar{\xi}_t = (1/n) \sum_{i=t-n+1}^t \xi_i$ and $\hat{\mu}_v = \bar{\xi}_v = (1/v) \sum_{i=t-n+1}^{t-n+v} \xi_i$, while $\hat{s}_0 = (v^{-1} \sum_{i=t-n+1}^{t-n+v} (\xi_i - \hat{\mu}_v)^2)^{0.5}$. Under the null hypothesis of no breaks, Chu, Stinchcombe, and White report asymptotic bounds for the statistic $|\hat{Z}_t|$:

$$\lim_{v \rightarrow \infty} P \{ |\hat{Z}_t| \geq \sqrt{v} \left(\frac{n-v}{v} \right) \left[\left(\frac{n}{n-v} \right) [a^2 + \ln \left(\frac{n}{n-v} \right)] \right]^{\frac{1}{2}} \} \cong 2(1 - \Phi(a) + a\phi(a)). \quad (\text{B2})$$

Here $\Phi(\cdot)$ and $\phi(\cdot)$ are the cdf and pdf of a standard normal random variable, respectively, while a is a constant related to the chosen significance level of the test. The intuition behind this test is that given a significance level, an agent could start the calculation of \hat{Z}_t recursively and in real-time after v signals received from the previous break to detect a new one. The testing process starts again after the detection of the new break. In this paper we assume that dividends are paid out daily, which is true for many market indexes. For that reason and with the objective of detecting breaks, we use daily dividend time series for the S&P 500 index between 1996 and 2007 that are de-seasonalized and adjusted by the Consumer Price Index. We set $v=125$, which represents six months of trading, and we use a 5% level of significance. We detect eight breaks in the period between 1996 and 2007. Figure B1 shows the breaks detected in the sample period.

[Insert Figure B1 here]

Appendix D: Options data

We use data from the U.S. option market for the 1996 - 2007 period to estimate a few typical indicators concerning the shape and dynamics of the IVS to be compared to the results obtained from the simulations of calibrated versions of our incomplete information, Bayesian learning framework. We include individual call equity options and call S&P 500 index options that are American and European style, respectively. We obtain the data from the OptionMetrics database, which reports daily closing bid and ask quotes, BS IVs, maturities, strike prices, synchronous (after appropriate adjustments that employ the put-call parity) closing underlying stock (index) prices, and the risk-free term-structure of interest rates. Option prices correspond to closing bid-ask midpoints. In relation to single-name stock options, we select only options in which the underlying stocks pay dividends, to

ensure the realism of our model. We choose the 150 names with the highest volume that have been continuously traded over our sample period.³² We sample option market data only on Wednesdays, as with our simulations. We apply four exclusionary criteria to filter out observations that represent noisy data, possibly recording errors, and that can hardly be thought to be expressions of well-functioning markets. Firstly, we eliminate all observations that violate basic no-arbitrage bounds, such as European put-call parity, American put-call boundaries, the lower bound etc. [see Bernales and Guidolin (2014) for a complete list of restrictions that are applied as filters]. Secondly, we delete all contracts with less than six trading days and more than one year to expiration as their prices are usually noisy. Thirdly and similar to Gonçalves and Guidolin (2006), we exclude contracts with prices lower than \$0.30 for equity options and \$3/8 for S&P 500 index options to avoid the effects of price discreteness on IVs (note that in the case of equity options, the minimum tick is \$0.05 for trading prices lower than \$3, while for index options the smallest tick is \$1/16). Finally, following Dumas, Fleming, and Whaley (1998), we exclude options contracts for which the moneyness is either less than 0.90 or in excess of 1.10, because deep in- and out-of-the money option contracts could cause additional noise in the analyses, and option series beyond these thresholds are normally illiquid and infrequently traded.

Appendix E: Calibrating the Fit of Deterministic IVS Models

In this appendix, we report the estimated fit of a simple and yet popular deterministic IVS model proposed in Dumas, Fleming, and Whaley (1998), using IVs simulated from a calibrated economy under breaks and incomplete information with learning. The implied volatility polynomial function that has been estimated is:

$$IV(Mon, \tau) = b_0 + b_1 Mon + b_2 Mon^2 + b_3 \left(\frac{\tau}{365}\right) + b_4 \left(\frac{\tau}{365}\right)^2 + b_5 Mon \left(\frac{\tau}{365}\right) + \epsilon \quad , \quad (D1)$$

where $IV(Mon, \tau)$ is the IV of a call option contract with moneyness Mon and time-to-maturity τ . Table E1 presents the coefficient estimates obtained by OLS and overall measures of fit for two alternative calibrations (high and low σ) and three alternative values for α . The coefficients are to be compared to the empirical ones estimated from data

³² The full list of 150 option series is available from the authors upon request.

on S&P 500 index calls or the average across 150 deterministic IVS regressions estimated for each of the stock options detailed in Appendix D. Table E1 shows that the IVS generated by the Bayesian learning model can be characterized by an IV polynomial function as in Dumas, Fleming, and Whaley (1998) in a similar way to the IVs reported in the option trading data.

[Insert Table E1 here]

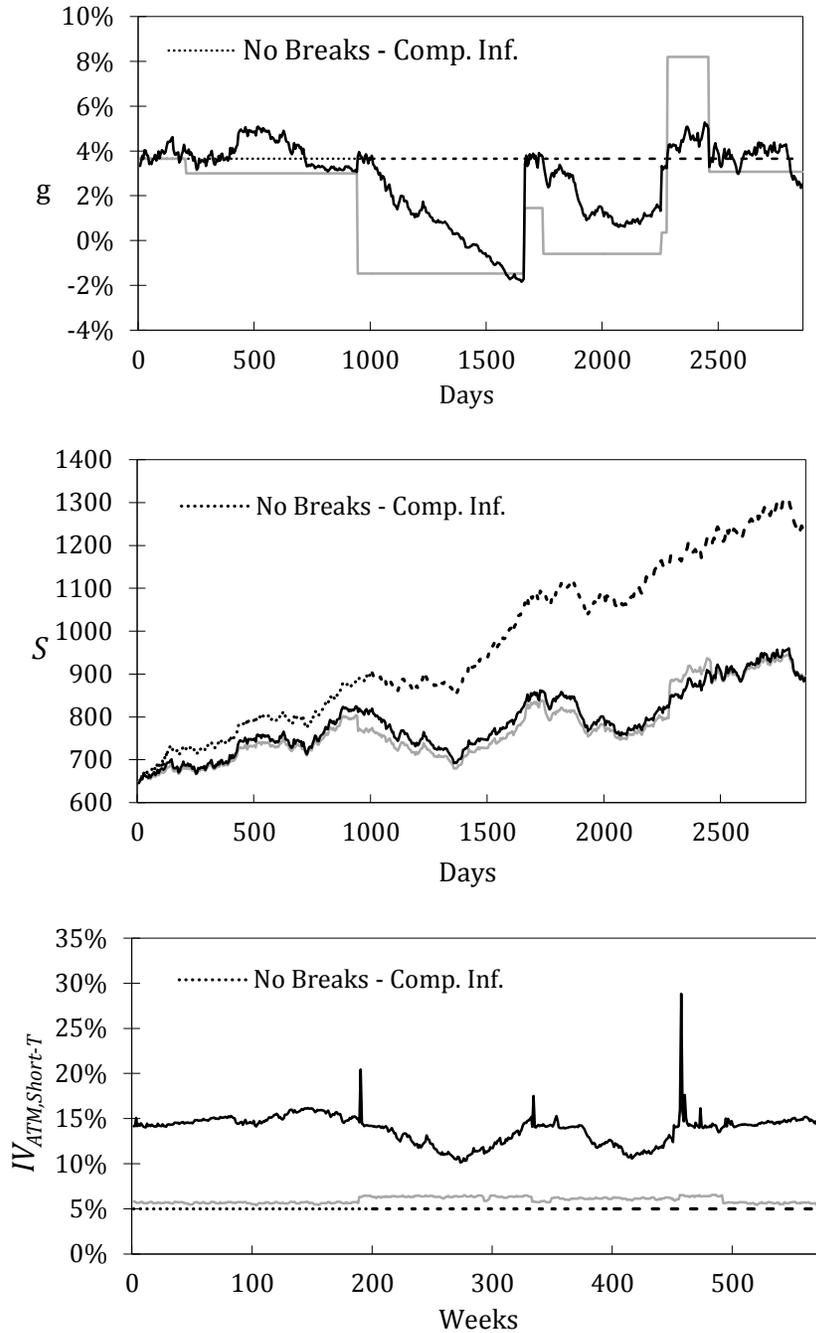


Figure 1. Evolutions of mean dividend growth rates, stock prices, and at-the-money short-term IVs under three scenarios. The figure shows the outcome for one simulated path concerning the dynamics over a 12-year sample for the mean dividend growth rate (g), stock prices (S), and at-the-money short-term IVs ($IV_{ATM,Short-T}$) under three case scenarios: no breaks; breaks and complete information; and breaks and incomplete information with rational learning. $IV_{ATM,Short-T}$ is the implied volatility corresponding to a call option contract with 30-days to the expiration date (calendar days) and at-the-money. Given that the simulations replicate option prices according to CBOE rules, we calculate $Call_{ATM,Short-T}$ by linear interpolation using the four contracts around the 30-day time-to-maturity and with closest strike price to S . The assumed parameters are: $\alpha=0.2$, $\pi=0.00301$, $\rho=8.9\%$, $\sigma=5.0\%$, $g_u=8.8\%$, and $g_d=-1.5\%$.

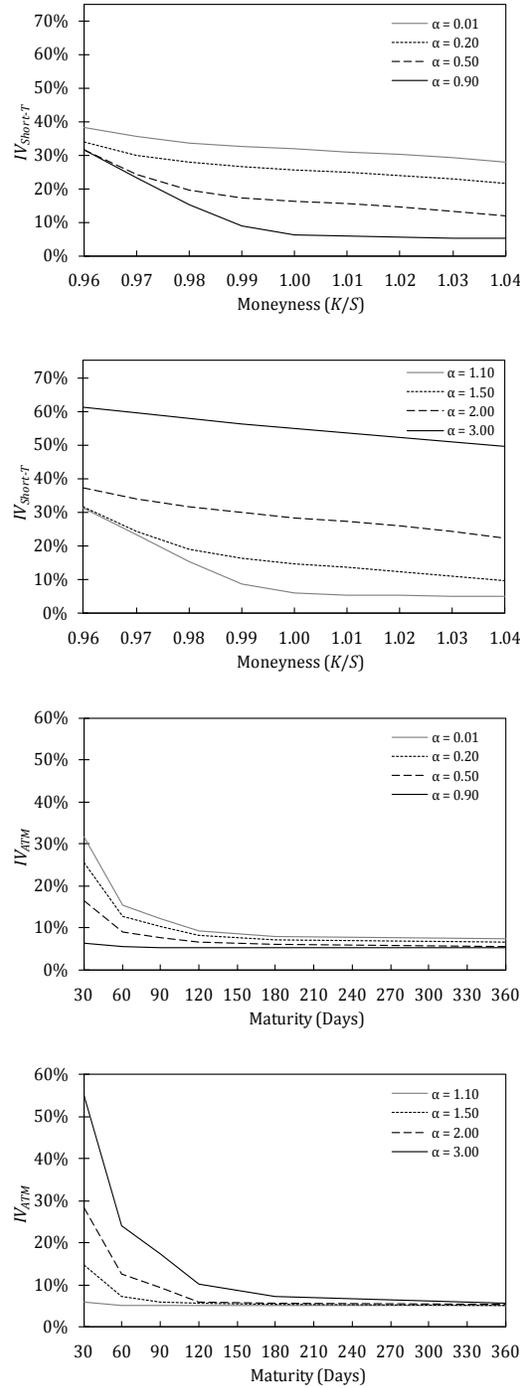


Figure 2. Sensitivity analysis on the average behavior of the implied volatility surface one month after a break in an economy with learning. The figure presents the average behavior, one month after a break in g_t , of implied volatilities in an economy under breaks and incomplete information with learning. This figure shows implied volatilities as a function of moneyness using short-term option contracts (upper windows) and implied volatilities as a function of time-to-maturity using at-the-money option contracts (lower windows). $IV_{Short-T}$ (IV_{ATM}) represents the implied volatilities corresponding to call option contracts with 30 days to the expiration date (strike prices equal to S). Given that the simulations replicate option prices according to CBOE rules, we calculate $Call_{ATM,Short-T}$ by simple linear interpolation using the four contracts around the 30-day time-to-maturity and with closest strike price to S . The assumed parameters are: $\pi=0.00301$, $\rho=8.9\%$, $\sigma=5.0\%$, $g_u=8.8\%$, and $g_d=-1.5\%$.

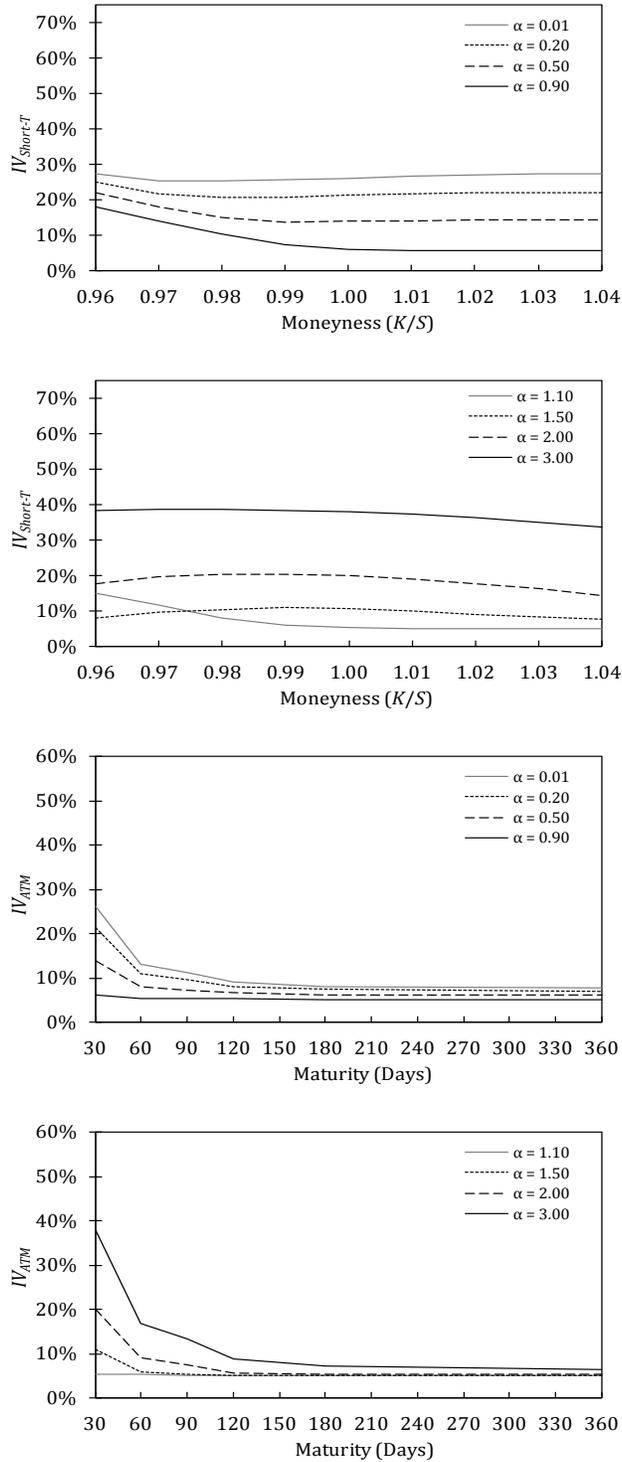


Figure 3. Sensitivity analysis on the average behavior of the implied volatility surface one year after a break in an economy with learning. The figure presents the average behavior, one year after a break in g_t , of implied volatilities in an economy under breaks and incomplete information with learning. This figure shows implied volatilities as a function of moneyness using short-term option contracts (upper windows) and implied volatilities as a function of time-to-maturity using at-the-money option contracts (lower windows). The assumed parameters are: $\pi=0.00301$, $\rho=8.9\%$, $\sigma=5.0\%$, $g_u=8.8\%$, and $g_d=-1.5\%$.

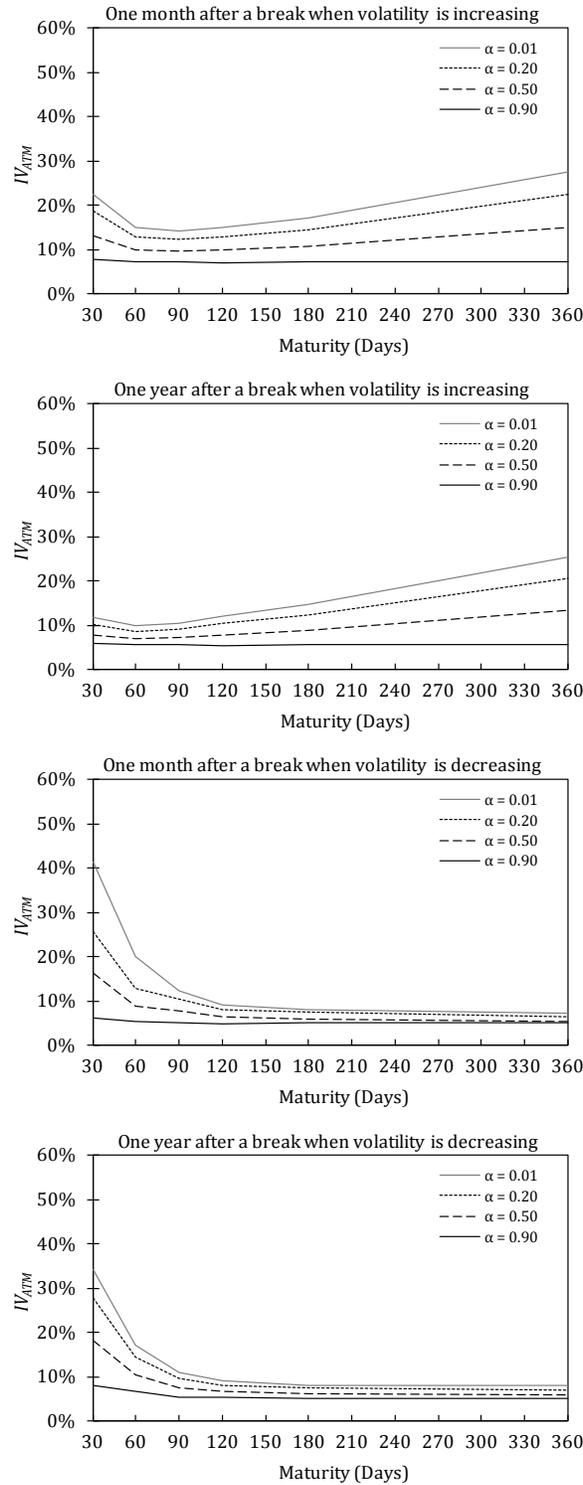


Figure 4. Average behavior of the implied volatility term-structure in an economy with learning as a function of dividend volatility. The figure presents the average behavior of implied volatility term-structure in an economy under breaks and incomplete information with learning when the dividend volatility is *increasing* (upper windows) and *decreasing* (lower windows) in the GARCH (1,1) process. The figure reports the average behavior of implied volatilities as a function of time-to-maturity using at-the-money option contracts one month (left-hand windows) and one year (right-hand windows) after a break in g_t . The assumed parameters are: $\pi=0.00301$, $\rho=8.9\%$, $g_u=8.8\%$, and $g_u=-1.5\%$.

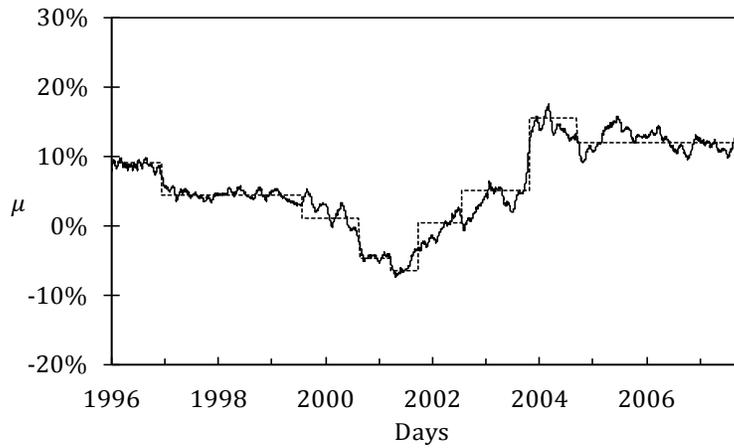


Figure B1. Structural breaks affecting the drift of the random walk process in equation (1). The solid line represents the recursive mean dividend growth rate (the drift) estimated with a rolling window of 125 trading days using continuously compounded dividends on the S&P 500 index, de-seasonalized and adjusted by the Consumer Price Index to obtain real dividends, between 1996 and 2007. The dotted line shows structural breaks estimated for the conditional mean function of dividends. Breaks were detected in December 1996, August 1999, September 2000, April 2001, October 2001, August 2002, November 2003, and October 2004.

Table 1

Simulation summary under three alternative case scenarios for investor's expectations

The variables g , S , and $IV_{ATM,Short-T}$ are defined in the note to Figure 1. Div is the daily dividend simulated while $r_{f,1\ day}$ is the one-day risk-free interest rate. $Call_{ATM,Short-T}$ is the price of a call option contract with 30 days to the expiration date (calendar days) and at-the-money. Given that the simulations replicate option prices according to CBOE rules, we calculate $Call_{ATM,Short-T}$ by simple linear interpolation using the four contracts around the 30-day time-to-maturity and with closest strike price to S .

Scenario	Variable	Mean	Median	Std. Dev.	Skewness	Excess Kurtosis	Min.	Max.
$\alpha=0.2$								
$\pi=0.00301, \rho=8.9\%, \sigma=5.0\%, g_u=8.8\%, \text{ and } g_d=-1.5\%$								
No Breaks - Comp. Inf.	Div	0,17	0,17	0,03	1,12	1,74	0,10	0,37
	g	3,65%	3,65%	0,00%	NA	NA	3,65%	3,65%
	$r_{f,1\ day}$	9,68%	9,68%	0,00%	NA	NA	9,68%	9,68%
	S	813,22	761,05	153,29	1,12	1,74	109,43	1301,20
	$Call_{ATM,Short-T}$	4,39	4,30	1,57	0,38	-0,07	1,01	11,95
	$IV_{ATM,Short-T}$	5,00%	5,00%	0,00%	NA	NA	5,00%	5,00%
Breaks - Comp. Inf.	Div	0,18	0,17	0,04	1,28	1,92	0,11	0,38
	g	3,66%	3,65%	2,96%	0,00	-1,21	-1,49%	8,78%
	$r_{f,1\ day}$	9,67%	9,75%	0,53%	-0,05	-0,90	9,29%	10,01%
	S	814,27	768,34	174,55	1,23	1,69	446,67	1678,97
	$Call_{ATM,Short-T}$	5,42	5,21	2,08	0,66	0,97	1,02	16,81
	$IV_{ATM,Short-T}$	5,75%	5,82%	0,31%	0,76	0,01	5,10%	7,44%
Breaks - Inc. Inf. (Learning)	Div	0,18	0,17	0,04	1,28	1,92	0,11	0,38
	g	3,64%	3,60%	1,62%	0,08	-0,16	-0,52%	8,33%
	$r_{f,1\ day}$	9,67%	9,56%	0,29%	0,04	-0,15	9,31%	10,03%
	S	814,10	769,82	173,25	1,25	1,84	449,06	1730,23
	$Call_{ATM,Short-T}$	12,75	12,34	3,12	1,24	9,53	2,93	91,95
	$IV_{ATM,Short-T}$	19,24%	19,62%	3,60%	0,93	3,23	7,13%	74,11%
$\pi=0.00301, \rho=9.6\%, \sigma=30.0\%, g_u=9.5\%, \text{ and } g_d=-5.0\%$								
No Breaks - Comp. Inf.	Div	0,18	0,13	0,16	2,46	8,97	0,00	1,64
	g	2,25%	2,25%	0,00%	NA	NA	2,25%	2,25%
	$r_{f,1\ day}$	10,09%	10,09%	0,00%	NA	NA	10,09%	10,09%
	S	698,66	501,26	626,41	2,46	9,07	18,47	6023,90
	$Call_{ATM,Short-T}$	18,32	13,93	12,83	0,80	-0,17	0,13	210,41
	$IV_{ATM,Short-T}$	30,00%	30,00%	0,00%	NA	NA	30,00%	30,00%
Breaks - Comp. Inf.	Div	0,18	0,13	0,17	2,56	9,07	0,00	1,66
	g	2,27%	2,25%	4,17%	0,01	-1,19	-4,94%	9,48%
	$r_{f,1\ day}$	10,08%	10,11%	0,79%	-0,09	-0,95	9,52%	10,52%
	S	700,93	504,80	676,20	2,62	9,67	18,91	6797,87
	$Call_{ATM,Short-T}$	26,13	23,30	18,44	1,28	2,35	1,60	241,42
	$IV_{ATM,Short-T}$	31,63%	31,28%	0,83%	0,83	0,77	30,06%	33,20%
Breaks - Inc. Inf. (Learning)	Div	0,18	0,13	0,17	2,56	9,07	0,00	1,66
	g	2,24%	2,22%	2,68%	0,25	-0,35	-1,73%	9,23%
	$r_{f,1\ day}$	10,04%	10,04%	0,13%	0,21	-0,31	9,79%	10,55%
	S	699,91	503,95	673,46	2,60	9,34	19,96	6990,88
	$Call_{ATM,Short-T}$	35,13	34,71	22,62	3,36	15,78	1,74	424,78
	$IV_{ATM,Short-T}$	43,28%	42,19%	5,40%	4,23	9,15	30,79%	221,23%

Table 2
Simulated properties of the implied volatility surface under rational learning

The table contains time series statistics concerning the level, slope, and curvature of the IVS in both the moneyness and maturity dimensions. The table presents the results of analyses for: i) an economy under breaks and incomplete information with learning on the left-hand side (Panels A and B); and ii) option market data in the U.S. over the period 1997 - 2007 (Panel A* and Panel B*). Panels A and B show average simulation outcomes using two parameter set ups; while Panels A* and B* report statistics for IVS shape features using S&P 500 index options and 150 individual equity options on dividend-paying stocks, respectively. The option market dataset is described in Appendix C. $IV_{ATM,Short-T}$ is defined in the note to Figure 1. $Slope_{Mon}$ ($Slope_{Mat}$) is the average across simulation trials of numerical first derivatives with respect to moneyness (time-to-maturity) computed from all the pairs of priced options with neighboring moneyness levels and 30 days to maturity (neighboring maturity levels and closest at-the-money). In addition, $Curv_{Mon}$ ($Curv_{Mat}$) is the average across simulation trials of numerical second derivatives with respect to moneyness (time to maturity) computed from all triplets of priced options with neighboring moneyness levels and 30 days to maturity (neighboring maturity levels and closest at-the-money). Serial Correlation refers to a Box-Pierce test applied to the first-order Ljung-Box statistic. The ARCH(1) and ARCH(3) statistics are the values of the LM test for ARCH effects. The percentage of simulations with significant statistics at a 10% level for the associated test statistics are reported in parentheses. Because in Panel A only the S&P 500 index option series is used in the tests, in this case the percentage in parentheses is simply 0 or 100.

Variable	Mean	Std. Dev.	Skew	Excess Kurt.	Serial Corr.	ARCH(1)	ARCH(3)	Variable	Mean	Std. Dev.	Skew	Excess Kurt.	Serial Corr.	ARCH(1)	ARCH(3)
Simulated data ($\alpha=0.2$)								Market Data							
Panel A: $\pi=0.00301$, $\rho=8.9\%$, $\sigma=5.0\%$, $g_u=8.8\%$, and $g_d=-1.5\%$								Panel A*: S&P 500 Options							
$IV_{ATM,Short-T}$	19,24%	3,60%	0,93	3,23	55,83	41,05	42,57	$IV_{ATM,Short-T}$	16,65%	5,90%	0,71	0,42	422,82	266,05	267,56
					(98.80)	(91.30)	(91.10)						(100.00)	(100.00)	(100.00)
$Slope_{Mon}$	-0,35	0,14	-0,31	2,35	18,04	5,66	10,10	$Slope_{Mon}$	-0,64	0,21	-0,48	0,69	82,70	15,21	18,21
					(78.50)	(39.30)	(42.70)						(100.00)	(100.00)	(100.00)
$Curv_{Mon}$	30,90	20,99	0,13	11,62	12,80	4,04	7,60	$Curv_{Mon}$	13,84	14,40	0,08	7,38	14,02	0,35	15,42
					(62.20)	(28.20)	(29.40)						(100.00)	(0.00)	(100.00)
$Slope_{Mat}$	-0,31	0,08	-0,02	1,81	53,16	37,03	38,76	$Slope_{Mat}$	0,03	0,07	-0,60	2,48	205,68	126,97	141,14
					(98.80)	(87.50)	(86.00)						(100.00)	(100.00)	(100.00)
$Curv_{Mat}$	3,28	1,12	0,21	4,47	47,56	29,66	31,61	$Curv_{Mat}$	-0,12	0,60	0,24	5,06	23,02	0,18	36,69
					(98.80)	(84.80)	(81.90)						(100.00)	(0.00)	(100.00)
Panel B: $\pi=0.00301$, $\rho=9.6\%$, $\sigma=30.0\%$, $g_u=9.5\%$, and $g_d=-5.0\%$								Panel B*: Equity Options							
$IV_{ATM,Short-T}$	43,28%	5,40%	4,23	9,15	50,46	39,74	38,31	$IV_{ATM,Short-T}$	40,34%	5,63%	0,67	0,86	59,33	29,06	31,39
					(93.30)	(88.70)	(90.30)						(98.00)	(74.00)	(69.33)
$Slope_{Mon}$	-0,19	0,15	-0,57	7,54	16,07	4,41	8,89	$Slope_{Mon}$	-0,21	0,70	1,03	26,31	4,72	2,18	7,02
					(75.80)	(31.90)	(35.30)						(32.67)	(13.33)	(17.33)
$Curv_{Mon}$	2,87	2,40	0,44	17,60	9,37	2,86	7,23	$Curv_{Mon}$	2,62	10,61	-0,35	11,50	1,85	2,64	4,98
					(53.40)	(27.80)	(26.20)						(18.67)	(15.33)	(18.67)
$Slope_{Mat}$	-0,31	0,09	-0,03	3,14	49,85	33,87	36,65	$Slope_{Mat}$	-0,04	0,09	-0,90	2,82	36,23	12,44	14,94
					(92.20)	(78.90)	(79.20)						(96.00)	(60.00)	(57.33)
$Curv_{Mat}$	3,53	1,72	0,62	6,38	44,38	27,61	26,48	$Curv_{Mat}$	0,08	0,55	0,59	3,92	17,39	7,38	10,40
					(92.20)	(77.90)	(78.30)						(90.00)	(48.67)	(44.67)

Table 3

Cross-sectional relations of IVS features under rational learning

The table contains correlations for the level, slope, and curvature of the IVS in both the moneyness and maturity. The table presents the results of analyses for: i) an economy under breaks and incomplete information with learning on the left hand side (Panels A and B); and ii) option market data in the U.S. over the period 1997 - 2007 (Panels A* and B*). Panels A and B show average simulation outcomes using two parameter set ups; while Panels A* and B* report statistics for IVS shape features using S&P 500 index options and 150 individual equity options on dividend-paying stocks, respectively. $IV_{ATM,Short-T}$ is defined in the notes to Figure 1, while $Slope_{Mon}$, $Curv_{Mon}$, $Slope_{Mat}$, and $Curv_{Mat}$ are defined in the notes in Table 2. The percentage of simulations with significant statistics at a 10% level for the associated test statistics are reported in parentheses. Because in Panel A only the S&P 500 index option series is used in the tests, in this case the percentage in parentheses is simply 0 or 100.

Variable	$IV_{ATM,Short-T}$	$Slope_{Mon}$	$Curv_{Mon}$	$Slope_{Mat}$	$Curv_{Mat}$	Variable	$IV_{ATM,Short-T}$	$Slope_{Mon}$	$Curv_{Mon}$	$Slope_{Mat}$	$Curv_{Mat}$
Simulated data ($\alpha=0.2$)						Market Data					
Panel A: $\pi=0.00301$, $\rho=8.9\%$, $\sigma=5.0\%$, $g_u=8.8\%$, and $g_d=-1.5\%$						Panel A*: S&P 500 Options					
$IV_{ATM,Short-T}$	1,00 (100.00)					$IV_{ATM,Short-T}$	1,00 (100.00)				
$Slope_{Mon}$	-0,38 (80.10)	1,00 (100.00)				$Slope_{Mon}$	-0,24 (100.00)	1,00 (100.00)			
$Curv_{Mon}$	-0,14 (67.90)	-0,47 (81.30)	1,00 (100.00)			$Curv_{Mon}$	-0,35 (100.00)	-0,09 (100.00)	1,00 (100.00)		
$Slope_{Mat}$	-0,96 (99.20)	0,32 (76.40)	0,17 (64.50)	1,00 (100.00)		$Slope_{Mat}$	-0,40 (100.00)	0,00 (0.00)	0,27 (100.00)	1,00 (100.00)	
$Curv_{Mat}$	0,89 (98.90)	-0,28 (74.80)	-0,15 (70.30)	-0,91 (98.50)	1,00 (100.00)	$Curv_{Mat}$	0,02 (0.00)	0,07 (0.00)	-0,05 (0.00)	-0,29 (100.00)	1,00 (100.00)
Panel B: $\pi=0.00301$, $\rho=9.6\%$, $\sigma=30.0\%$, $g_u=9.5\%$, and $g_d=-5.0\%$						Panel B*: Equity Options					
$IV_{ATM,Short-T}$	1,00 (100.00)					$IV_{ATM,Short-T}$	1,00 (100.00)				
$Slope_{Mon}$	-0,35 (74.60)	1,00 (100.00)				$Slope_{Mon}$	-0,11 (46.67)	1,00 (100.00)			
$Curv_{Mon}$	-0,12 (66.90)	-0,46 (79.70)	1,00 (100.00)			$Curv_{Mon}$	-0,29 (86.00)	-0,11 (64.67)	1,00 (100.00)		
$Slope_{Mat}$	-0,42 (89.90)	0,31 (72.40)	0,00 (58.90)	1,00 (100.00)		$Slope_{Mat}$	-0,57 (97.33)	0,28 (80.00)	0,06 (46.67)	1,00 (100.00)	
$Curv_{Mat}$	0,38 (79.70)	-0,11 (54.60)	-0,09 (52.80)	-0,72 (81.20)	1,00 (100.00)	$Curv_{Mat}$	0,24 (76.67)	-0,25 (80.00)	0,03 (41.33)	-0,62 (99.33)	1,00 (100.00)

Table 4**Simulated properties of the implied volatility surface under rational learning when the dividend volatility varies**

The table contains time series statistics concerning the level, slope, and curvature of the IVS in both the moneyness and maturity dimensions in an economy under breaks and incomplete information with learning when the dividend volatility follows a GARCH(1,1) process. The table shows average simulation outcomes using a one-parameter set up. $IV_{ATM,Short-T}$ is defined in the notes to Figure 1, while $Slope_{Mon}$, $Slope_{Mat}$, $Curv_{Mon}$ and $Curv_{Mat}$ are defined in the notes to Table 2. Serial Correlation refers to a Box-Pierce test applied to the first-order Ljung-Box statistic. The ARCH(1) and ARCH(3) statistics are the values of the LM test for ARCH effects using one and three lags, respectively. The percentage of simulations with significant statistics at a 10% level for the associated test statistics are reported in parentheses.

Variable	Mean	Std. Dev.	Skewness	Excess Kurtosis	Serial Correlation	ARCH(1)	ARCH(3)
<i>Dividend volatility follows a GARCH(1,1) where $\alpha=0.2$</i>							
<i>$\pi=0.00301$, $\rho=8.9\%$, $g_u=8.8\%$, and $g_d=-1.5\%$</i>							
$IV_{ATM,Short-T}$	20,71%	3,88%	1,02	3,51	56,74 (100.00)	45,43 (100.00)	51,27 (94.50)
$Slope_{Mon}$	-0,39	0,18	-0,36	2,84	21,63 (83.60)	7,07 (55.70)	11,43 (59.20)
$Curv_{Mon}$	35,31	21,61	0,16	12,90	16,30 (71.30)	4,60 (55.30)	9,62 (41.60)
$Slope_{Mat}$	-0,12	0,16	-0,02	2,34	67,49 (100.00)	38,89 (100.00)	47,39 (94.60)
$Curv_{Mat}$	2,09	2,39	0,25	5,39	60,80 (100.00)	35,32 (98.30)	40,05 (91.40)

Table 5**Cross-sectional relations of IVS features under rational learning when the dividend volatility varies**

The table contains correlations for the level, slope, and curvature of the IVS in both the moneyness and maturity dimensions in an economy under breaks and incomplete information with learning when the dividend volatility follows a GARCH(1,1) process. The table shows average simulation outcomes using one parameter set up. $IV_{ATM,Short-T}$ is defined in the notes to Figure 1, while $Slope_{Mon}$, $Curv_{Mon}$, $Slope_{Mat}$, and $Curv_{Mon}$ are defined in the notes to Table 2. The percentage of simulations with significant statistics at a 10% level for the associated test statistics are reported in parentheses.

Variable	$IV_{ATM,Short-T}$	$Slope_{Mon}$	$Curv_{Mon}$	$Slope_{Mat}$	$Curv_{Mat}$
<i>Dividend volatility follows a GARCH(1,1) where $\alpha=0.2$</i>					
$\pi=0.00301, \rho=8.9\%, g_u=8.8\%, \text{ and } g_d=-1.5\%$					
$IV_{ATM,Short-T}$	1,00 (100.00)				
$Slope_{Mon}$	-0,45 (85.30)	1,00 (100.00)			
$Curv_{Mon}$	-0,15 (73.90)	-0,58 (86.90)	1,00 (100.00)		
$Slope_{Mat}$	-0,88 (88.40)	0,25 (72.10)	0,13 (59.30)	1,00 (100.00)	
$Curv_{Mat}$	0,73 (87.20)	-0,24 (70.30)	-0,13 (64.30)	-0,87 (95.20)	1,00 (100.00)

Table E1**Comparing Fitted Deterministic Volatility Functions on Market Data vs. from Simulations from an Economy under Incomplete Information with Rational Learning**

The table contains OLS coefficient estimates and fit measures obtained from estimating equation (E1) on average implied volatilities from the U.S. option market compared to simulated option IVs from a Bayesian learning model under alternative calibrations.

	b_0	b_1	b_2	b_3	b_4	b_5	R^2	F Statistic	p-value
Panel A: $\pi=66.7\%$, $\rho=8.9\%$, $\sigma=5.0\%$, $g_u=8.8\%$, and $g_d=-1.5\%$									
$\alpha=0.2$	2,70	-4,76	2,21	-0,74	0,20	0,47	0,89	53,84	0,00
$\alpha=0.5$	4,12	-7,53	3,52	-0,88	0,13	0,71	0,91	37,70	0,00
$\alpha=5.0$	-7,04	14,44	-6,94	-0,32	0,71	-0,68	0,94	71,55	0,00
Panel B: $\pi=66.7\%$, $\rho=9.6\%$, $\sigma=30.0\%$, $g_u=9.5\%$, and $g_d=-5.0\%$									
$\alpha=0.2$	2,94	-4,90	2,40	-0,73	0,21	0,48	0,88	49,12	0,00
$\alpha=0.5$	4,11	-7,16	3,41	-0,94	0,13	0,72	0,90	36,30	0,00
$\alpha=5.0$	-5,54	13,99	-7,03	-0,33	0,72	-0,62	0,93	67,68	0,00
Market Data									
S&P 500 Options	7,31	-13,55	6,39	0,50	-0,03	0,54	0,85	31,75	0,00
Equity Options	4,60	-8,11	3,92	-0,22	0,05	0,12	0,81	43,61	0,00