How Bad Will the Potential Economic Disasters Be? Evidences From S&P 500 Index Options Data

June, 2008

Abstract

This paper proposes a new methodology of using the S&P 500 index option data to gauge the magnitudes of the potential economic disasters in the U.S. with a setup incorporating a small-probability consumption jump and habit formation. The estimated economic disasters strike once every 36-64 years in the form of 13.5-17.6% consumption contractions, which induce 36-56% stock market crashes. These results are much more in line with the empirical observations than those estimated with Peso problem models in which habit formation is ignored. The setup also explains a wide variety of the observed pricing features of both options and stocks.

JEL code: G10, G12

Key words: economic disasters, habit formation, volatility smirk, equity premium

Introduction

Understanding the relation between financial markets and the real economy is one of the key programs in finance, since asset prices are ultimately driven by the real, macroeconomic risks[1]. Perhaps the most extreme version of macroeconomic risks strikes in the form of disastrous economic events, such as the Great Depression, during which the levels of aggregate consumption and GDP decrease dramatically. While it is apparent that potential economic disasters can have profound effects on the financial markets[2], empirically we lack direct observations to calibrate their magnitudes since disasters are rare and usually yield no more than a few observation points during a long period of time. Whereas disasters that happened decades ago can shed some evidences about what will happen in the future, they could be very different from the potential ones that investors are concerned about today, which have not yet been realized in sample.

In the literature, potential economic disasters are originally used in peso problem models[3] to resolve the equity premium puzzle (e.g., Rietz, 1988), hence a natural methodology is to infer their magnitudes from the observed equity premium. In Rietz's model, as in other peso problem models in which disasters strike once every a few decades, generating an average equity premium comparable to what is observed in the data under reasonable risk aversion degree[4] requires a potential 40-60% consumption or GDP contraction which looks too dramatic. In a recent empirical paper, however, Barro (2006) reports evidences of very large contractions in consumption and GDP up to 64% from the data of international economic disasters that occurred during the twentieth century. While both Rietz and Barro focus on the stock market, potential disasters may also have profound effects on the derivative markets. As widely acknowledged in the option pricing literature, poten-

[1]See Cochrane (2006) for a comment about the relation between macroeconomics and finance.

[2]The phenomenon that infrequent but disastrous events (not necessarily economic events) can have profound effects on empirical observations is sometimes refered to as "Peso problem", often attributed to Milton Friedman for his comments about the Mexican peso market in the early 1970s. Peso problem situation provides a good explaination for a number of asset pricing anomalies including the equity premium puzzle (e.g., Rietz (1988)), the failure of the expectation hypothesis (e.g., Cochrane, 2001), and the observed pricing differentials among options across moneyness (e.g., Rubinstein, 1994).

[3]In this paper, peso problem models refer to those modeling occasional extremely bad states in the distribution of consumption or GDP in the otherwise standardard consumption-based asset pricing framework.

[4]As commented in Cochrane (2001), a high degree of risk aversion in the standard consumption-based model induces counterfactually high volatility of the risk-free rate which is empirically quite stable across time.

tial stock market crashes, which are induced by the potential economic disasters in peso problem models, play an important role in explaining the observed pricing differentials among options across moneyness, paticularly the prevalent "smirk" pattern documented in the index option market. It is thus natural to ask whether economic disasters of the size required to justify the observed equity premium are consistent with the observed option pricing differentials.

In the upper Panel of Figure 1, I plot the average prices of S&P 500 index options with 30 days to expirations, quoted in terms of the implied Black-Scholes volatility (B/S-vol)[5], against the options' moneyness[6], where the data are collected in the period from April 4, 1988 to June 30, 2005. As well documented in the literature (e.g., Rubinstein, 1994), out-of-the-money (OTM) put options are more valuable than at-the-money (ATM) options, generating a pronounced "smirk" pattern in the cross-sectional plot of the option-implied volatility against moneyness, normally referred to as the "volatility smirk". I plot in the same panel the volatility smirk implied implied from Barro (2006) in which he assumes a potential 37% consumption contraction that strikes once every 60 years. While the smirk premium, measured as the price difference between 8% OTM put options (i.e., put options with moneyness equal to .92) and ATM options, is about 10.4% from the data, it is more than 18% in Barro (2006). It seems that the dramatic economic disaster required to rationalize the high U.S. equity premium is too severe for the observed option pricing differentials. Indeed, cutting the jump size in half while keeping all the other parameter values unchanged reduces the smirk premium implied from Barro (2006) to about 13%, which is also plotted in the upper Panel of Figure 1.

[Figure 1 goes approximately here]

The generated counterfactually high smirk premium casts doubt on the methodology of calibrating the potential economic disasters to the observed equity premium with peso problem models. It is likely that peso problem models are misspecified and have ignored other important driving forces behind the observed asset pricing phenomena, such as habit formation. Indeed, habit formation models were originaly raised to resolve the equity premium puzzle as well (e.g., Constantinides, 1990). Later researches (e.g., Compbell and

[5] Option prices in this paper are all quoted in terms of the implied Black-Scholes volatility (B/S-vol).

[6] Moneyness is defined as the ratio of the strike price of the option contract to the underlying stock price.

Cochrane, 1999, henceforth CC; Menzly, Santos and Veronesi, 2005, henceforth MSV) find that the idea of habit formation actually explains a wide variety of the observed stock pricing phenomena. Like peso problem models, however, habit formation models have not yet been applied to price options. In the lower Panel of Figure 1, I plot the volatility smirk implied from the same data together with its two counterparts implied from CC and MSV, where the model prices are computed at the parameter values reported in these two papers. We observe an "inverted" smirk implied from CC[7]. Whereas MSV generates the usual smirk pattern, the implied smirk premium is only 3.1%, which is clearly off compared to the observed 10.4% premium. In exercises that are not reported, I tried different parameter values for both CC and MSV while keeping their matches of the observed equity premium, but the generated smirk premia never exceed 6%. It seems that while peso problem situation (i.e., the presence of potential economic disasters) and habit formation are both important devices to boost the equity premium, the former is much more effective than the latter to generate the pricing differentials among options across moneyness.[8]

The key observation in this paper is that a model incorporating both peso problem situation and habit formation paves a new way to gauge the magnitude of the potential economic disasters. With habit formation pushing up the equity premium without affecting the smirk premium much, I can use index option data to back out the implied economic disasters priced in the form of the observed option pricing differentials across moneyness. Disasters thus estimated, together with the estimated habit formation, are likely to replicate simultaneously the observed equity premium and the observed smirk premium in option pricing.

The proposed methodology of calibrating the potential economic disasters to option data is linked to the literature that examines the asset pricing implications of reduced-form option pricing models in which the stock prices are subject to potential jumps (e.g., Bates, 2000; Pan, 2002; Eraker, 2004). While these papers also use option data to back

[7] The implied volatilities from CC are cut for small moneyness, because they do not exist for moneyness less than 0.97. The option implied volatility is monotonically increasing in the option price, and theoretically it can take negative values for a too low option price. We say that the implied volatility does not exist if its theoretical value is negative.

[8] To give it an intuitive explanation, notice that the innovations of habit dynamics are assumed to be perfectly negatively correlated with the consumption diffusions in both CC and MSV, hence risks associated habit formation still belong to diffusion risks. As discussed in Liu, Pan and Wang (2005), OTM put options, in particular the deep OTM put options, are much more sensitive to jumps than to diffusions.

out the potential disasters, they differ from my estimations in two important dimensions. First and foremost, their estimations are based on reduced-form models starting with exogenous specifications of the stock return processes. As a result, their methodology can not be applied to gauge the magnitudes of the potential economic disasters in the form of negative jumps in consumption or GDP. Second, option data through their models only back out jumps under the risk neutral measure, which are not empirically relevant until transformed into the physically-measured jumps with some exogenous specifications about the risk premia. In contrast, my model establishes the direct links between the implied asset prices and the potential disasters under the physical measure, hence the direct estimation of the empirically relevant jumps from the option data.

Following Barro (2006), the growth of log consumption is modeled as a random walk subject to a small-probability negative consumption jump with constant arrival intensity and jump size. Analogously to both CC and MSV, I adopt the external habit formation specification in which the representative agent's current utility depends not only on his own current consumption, but also on the history of the aggregate consumption. For the mathematical tractability, I use MSV habit formation specification (in a slightly extended version), which is the time-continuous analog of that in CC. Three models are considered in this paper under the same specified potential economic disasters and habit formation but with different dividend processes. Closed-form stock prices are obtainable in all the three models, which greatly facilitate the simulation of the model-implied option prices.

Both potential disasters and habit formation are priced in the financial markets through the pricing kernel, hence I adopt an integrated approach of using the joint data of options and stocks to estimate the jump parameters simultaneously with the habit formation parameters. S&P 500 index option data are collected from two datasets: CBOE from 04/04/1988 to 12/29/1995, and Ivy DB from 01/04/1996 to 06/30/2005; the other data are from CRSP and St. Louis Fed. I find that the pricing differentials between deep OTM put options and ATM options provide particularly important information for the identification of the potential disasters. Due to the fact that the trading of OTM put options is low during the CBOE period, I use Ivy DB data only for the estimation. CBOE data together with the out-of-the-sample Ivy DB data are used later to test the models' option pricing implications.

I apply the method of simulated moments (MSM) to implement the estimation which selects parameter values to best match the simulated asset pricing moments to their data counterparts. Only unconditional moments are used because the underlying economic state, i.e., the surplus consumption ratio, is not observable. I choose to match the average prices of nine options, all with 30 days to expiration. For the non-option pricing moments, I choose to match the average values of dividend-price ratio, the equity premium and the risk-free rate. I combine Chebyshev interpolation with Monte Carlo simulation to efficiently compute the unconditional option pricing moments implied from the model, and I use the actual data to compute the optimal weighting matrices approximated by Newey-West. The minimization problem is solved using the global optimization toolbox provided by Tomlab which yields stable parameter estimates for different starting values.

Depending on the assumed dividend processes, potential economic disasters implied from the option market strike once every 36-64 years in the form of a 13.5%-17.6% consumption contraction, which induces an average 36-56% stock market crash. Empirically, Barro (2006) documents two major economic downturns in the U.S. during the past 100 years: the Great Depression (1929-1933) and the aftermaths of WWII (1944-1947), during which aggregate consumption contracts by as much as 9.9% per year. The estimated arrival intensity seems consistent with the historical observations so far. While my estimated consumption jump sizes still look dramatic relative to the annual consumption contractions during the historical economic disasters, they are much more in line with the observed data than those estimated with peso problem models in which a potential 40-60% consumption contraction is usually demanded to match the observed average equity premium. In summary, my estimation implies that the U.S. investors are expecting another major economic crisis that is more severe than the Great Depression within the next 50 years or so.

Economic disasters in my model induce the crashes in the stock market through two channels: i) the usual channel due to the jumps in the dividends that are of the same magnitudes as those of the consumption jumps, which induces jumps of the same sizes in the stock market; ii) the extra channel due to the jumps in the agent's instantaneous risk aversion degree at the presence of habit formation, which causes excessive depressions in the stock price relative to dividends. As a result, the induced stock market crashes are much more dramatic than the estimated consumption jumps. In contrast, consumption jumps induce stock market crashes only though the usual channel in peso problem models, which imply the same jump sizes in the consumption and in the stock price. Empirical evidences, however, seem to support the implication of excessive stock price adjustment. For example, Liu et al (2003) report two major crashes in the U.S. stock market during the past 100 years: one is from mid-October to mid-November in 1929, and the other is the black Monday of October 19, 1987, when the Dow index fell by 44% and 23%, respectively. While these falls are accumulated within only one month or even one day, they are already much larger than the maximumm 9.9% annual consumption contraction during the Great Depression.

The three estimated models, regardless of the different dividend processes, all match well a wide variety of the observed pricing phenomena of both options and stocks. While the parameters are only chosen to replicate the unconditional moments and only short term options data are used for the estimation, the models perform well in many other dimensions as well, such as the the volatility smirks conditional on particular economic states, the term structure in option pricing, stock return predictability, and the historical fluctuations in the smirk premia and the dividend-price ratios. To my best knowledge, the matches of the term structure in option pricing and the historical smirk premia data have never been attempted before. By comparsion, the other two consumption-based option pricing models that I know, Liu et al (2005) and Benzoni et al (2005), both focus on the match of the unconditional volatility smirk. In particular, neither paper can generate the observed variations of volatility smirks across time or across different economic states[9]. My models' asset pricing ability, besides being interesting of its own, suggests that the model misspecification is a relatively minor issue in a setup incorporating both potential economic disasters and habit formation intended to replicate the joint pricing behaviors of options and stocks, hence an important support to the proposed methodology.

The structure of the paper is as follows: Section I presents the setup and derives the theoretical results for the baseline model treating consumption and dividend as a single process; Section II describes the data; Section III discusses the details of the methodology; Section IV reports the estimation results, in particular, the estimated potential disasters, with the baseline model; Section V considers two alternative models with different dividend processes and report their implied estimations; Section VI discusses the models' out-of-the-sample asset pricing implications; and finally Section VII concludes. All mathematical proofs are in the Appendix.

[9] Liu et al (2005) do not model economic state, which implies that the prices of options with given maturity and moneyness are constants. Whereas Benzoni et al (2005) model the economic state, x, as a persistent stochastic growth variable in the consumption and dividend processes, they report that option prices computed at values of x up to ± 3 standard deviations away from its steady state value $\bar{x}$ still turn out to be very similar to those computed at $\bar{x}$ itself.

I Setup

A. The preference

Time is continuous and infinite and the representative agent in the economy maximizes

$$E\left[\int_0^\infty e^{-\rho t} u(C_t, X_t) dt\right] \tag{1.1}$$

where

$$u(C_t, X_t) = \left\{ \begin{array}{ll} \frac{(C_t - X_t)^{1-\gamma}}{1-\gamma} & \text{if } \gamma > 0, \gamma \neq 1 \\ \ln(C_t - X_t) & \text{if } \gamma = 1 \end{array} \right\}; \tag{1.2}$$

$X_t > 0$ and ρ denote respectively the habit level and the subjective time-discount factor[10]. MSV assume a log habit formation setup by setting $\gamma = 1$. Here I follow CC by allowing any $\gamma > 0$. As in both CC and MSV, I adopt the external habit formation specification in which the agent's current utility depends not only on his own current consumption, but also on the history of the aggregate consumption summarized by X_t.

To capture the relation between habit and the aggregate consumption, CC models the *surplus consumption ratio*,

$$S_t \equiv \frac{C_t - X_t}{C_t} < 1, \tag{1.3}$$

which serves as the fundamental state variable in the habit formation models. Since habit is external, the local curvature of the utility function, η_t, which measures the agent's instantaneous risk aversion degree, is related to the surplus consumption ratio by

$$\eta_t \equiv -\frac{C_t u_{cc}(C_t, X_t)}{u_c(C_t, X_t)} = \frac{\gamma}{S_t} > \gamma. \tag{1.4}$$

From (1.3) and (1.4), a decline of consumption toward its habit level drives up the agent's degree of risk aversion, which plays the key role in explaining the observed discrepancy between the stock market jump size and the consumption jump size.

For the purpose of mathematical tractability, I follow MSV by imposing the stochastic structure on a function of S_t, which is denoted by G_t,

[10] In this paper, all variables without the time subscripts are assumed to be constants.

$$G_t \equiv S_t^{-\gamma} = \left(\frac{C_t}{C_t - X_t}\right)^{\gamma} \tag{1.5}$$

Since $\gamma > 0$, G_t moves one for one with S_t and hence can be used as the equivalent state variable. Analogously to CC and MSV, I assume that G_t is mean-reverting with shocks that are conditionally perfectly correlated with innovations in consumption growth,

$$dG_t = k(\bar{G} - G_t)dt - \alpha(G_t - \beta)(d\ln(C_t) - E_t[\ln(C_t)])^{11}, \tag{1.6}$$

where $\bar{G}$ is the long run average of G_t; k is the speed of mean reversion; $\alpha > 0$ captures the impact of unexpected consumption innovations on the evolution of G_t: a negative consumption innovation, for example, drives consumption towards its habit level resulting in an increase in G_t, or equivalently, a decrease in S_t; β is the lower bound for G_t and hence ensures an upper bound for S_t, and I impose $\beta \geq 1$ so that $S_t \leq 1$.

B. Consumption and dividend

Analogously to Barro (2006), I assume that the growth of the log consumption evolves as a random walk subject to a small-probability negative jump,

$$d\ln(C_t) = \tilde{\mu}dt + \sigma dB_{1t} - bdN_t \tag{1.7}$$

where B_{1t} is a standard Brownian motion; N_t is a Poisson process with arrival intensity λ capturing the random arrival of a small-probability economic disaster;[12] $b > 0$ captures that the economic disaster strikes in the form of negative consumption jumps. By Ito's

[11] Under the usual assumption that the habit X_t is the weighted average of the past consumptions, the drift and the diffusion terms in the G_t process are both nonlinear functions of G_t. The assumed G_t process in (1.6) is mainly for the purpose of tractability. Under the assumption in (1.6), the habit X_t is no longer the weighted average of past consumptions, but a more complicated nonlinear function of past consumption shocks.

[12] Both Barro (2006) and in my model, disasters last for an instant of time. In the data, however, economic crises usually perisist for varying number of years. As pointed out by Barro (2006), an alternative procedure is to allow for two regimes, normal and crisis, with transition probabilities reflecting the relative length of the two regimes. While that procedure is more realistic, it is mathematically more difficult to handle since it admits a number of rich complications such as the distribution of jump sizes. To focus on the first order effects of potential disasters on asset prices and to keep the model tractable, I thus adopt the simpler way, as Barro (2006) does, by modeling the potential disaster as a time invariant Poisson arrival process.

lemma,

$$\frac{dC_t}{C_t} = \mu dt + \sigma dB_{1t} + (e^{-b} - 1)dN_t \tag{1.8}$$

where $\mu \equiv \tilde{\mu} + \frac{1}{2}\sigma^2$; μ and σ denote respectively the mean and the standard deviation of the consumption growth rate conditional on no economic disasters. Upon the occurrence of an economic disaster, consumption contracts to $e^{-b} < 1$ of its level right before the arrival of the crisis, hence the consumption jump size is measured by $e^{-b} - 1$.

The model is closed by assuming an aggregate dividend process. I first impose the simplest yet the widely used way (e.g., CC; MSV; Liu et al, 2005; Barro, 2006) by treating dividend and consumption as a single process, i.e., $D_t = C_t$, and I call the model thus closed the "baseline model", which is used to explain the details of the proposed methodology. In Section V, I will consider two alternative models with different and more realistic dividend processes .

C. Stock pricing

Under the assumed external habit formation specification, pricing kernel in the economy is given by

$$\Lambda_t = e^{-\rho t} u_c(C_t, X_t) = e^{-\rho t} C_t^{-\gamma} G_t \tag{1.9}$$

which incorporates the effects of both the potential economic disasters and habit formation. Denote by P_t the price of the aggregate stock at period t, and by definition,

$$P_t = E_t \int_t^{\infty} \frac{\Lambda_s}{\Lambda_t} C_s ds \tag{1.10}$$

Closed-form stock pricing formulas obtain under the assumed preference and the consumption processes, which are summarized in the following proposition:

Proposition 1 *(a) Under the assumed utility in (1.1)-(1.2), the assumed habit formation in (1.3)-(1.6), and the assumed consumption processes in (1.8), the price-dividend ratio (which is equal to the price-consumption ratio in the baseline model) is given by:*

$$\frac{P_t}{C_t} = a_1 + a_2 \frac{1}{G_t} \tag{1.11}$$

where the two constant coefficients a_1 and a_2 are defined in (A.14) in the Appendix.

(b) *The risk-free rate (rf_t), the equity premium (EP_t) and the return volatility ($volR_t$) in the economy are given by:*

$$rf_t = -\mu_{\Lambda t} - \lambda J_{\Lambda t} \tag{1.12}$$

$$EP_t = \sigma_{\Lambda t}\sigma_{Pt} - \lambda J_{\Lambda t} J_{Pt} \tag{1.13}$$

$$volR_t = \sqrt{\sigma_{Pt}^2 + \lambda J_{Pt}^2} \tag{1.14}$$

where $J_{Xt} \equiv \frac{X_t^+}{X_t} - 1$ ($X = \Lambda, P$) denote respectively the jump sizes of the pricing kernel and the stock price; $\mu_{\Lambda t}$, $-\sigma_{\Lambda t}$ are the drift and the diffusion terms in the process of $\frac{d\Lambda_t}{\Lambda_t}$; σ_{Pt} is the diffusion term in the process of $\frac{dP_t}{P_t}$. $J_{\Lambda t}$, J_{Pt}, $\mu_{\Lambda t}$, $\sigma_{\Lambda t}$, and σ_{Pt} are all closed-form functions of the state G_t, and their formulas are given by (A.20)-(A.24) in the Appendix.

In habit formation models, a decrease in the surplus consumption ratio, or equivalently an increase in G_t, drives up the curvature of the utility function according to (1.4) making the agent more risk averse. As a result, the price is depressed relative to dividend, a higher equity premium is demanded, and the return volatility becomes more volatile. These predictions are verified through an (unreported) numerical analysis of (1.11), (1.13), and (1.14) using the parameter estimates reported in Table II. On the other hand, negative consumption jumps according to (1.9) and (1.11) induce both positive jumps in Λ_t and negative jumps in P_t, i.e., $J_{\Lambda t} > 0$ and $J_{Pt} < 0$. Hence from (1.12)-(1.13), the risk-free rate goes down and the equity premium goes up. These are the findings in Rietz (1988) and Barro (2006) that peso problem situation helps simultaneously resolve the equity premium puzzle and the risk-free rate puzzle.

Unlike peso problem models which imply the same jump sizes in both consumption and the stock price, negative consumption jumps in my setup induce the crashes in the stock market through two channels: the usual channel due to the negative jumps in dividends, which are of the same magnitudes as the consumption jumps; and the extra channel due to the postive jumps in the agent's instantaneous risk aversion degree as implied by the positive jumps in G_t at the presence of habit formation, which causes excessive depressions in the stock price relative to dividends. As a result, the induced stock market crashes are on average much more severe than the consumption jumps, a phenomenon which is consistent with empirical observation. In addition, the extra jumps in Λ_t due to the risk-aversion jumps imply even lower risk-free rates and even higher equity premia than predicted by the peso problem models according to (1.12)-(1.13). Intuitively, the agent has stronger incentive to shift his investment from the stock to the risk-free saving when a stock market crash is expected to occur at the same time when his risk aversion degree jumps upward. These implications are due to the "magnification effects" of habit formation in the peso problem situations, which was not previously explored.

D. Option pricing

At period t the price of a put option written on the stock that matures at period $t + \tau$ with the strike price K is by definition,

$$Put_t = E_t \left[\frac{\Lambda_{t+\tau}}{\Lambda_t} \max(K - P_{t+\tau}, 0) \right] \tag{1.15}$$

From (1.15), potential economic disasters have direct impacts on the generated option prices through the induced jumps in both the stock price and the pricing kernel. To show explicitly the effects of habit formation, I normalize Put_t by the spot stock price as follows,

$$O_t \equiv \frac{Put_t}{P_t} = E \left[e^{-\rho\tau} \left(\frac{C_{t+\tau}}{C_t} \right)^{-\gamma} \frac{G_{t+\tau}}{G_t} \max \left(M_t - \frac{C_{t+\tau}}{C_t} \frac{a_1 + a_2 \frac{1}{G_{t+\tau}}}{a_1 + a_2 \frac{1}{G_t}}, 0 \right) | G_t \right] \tag{1.16}$$

where the second equality uses (1.9) and (1.11); $M_t \equiv \frac{K}{P_t}$ is the moneyness. Given the starting value of G_t, the evolution of habit formation summarized by the assumed dynamics of G, together with the assumed consumption process, drives the evolution of both the pricing kernel and the stock price between t and $t + \tau$, whose end-of-the-period values determine the discounted end-of-the-period payment from the option. Option price by definition is the expectation of the discounted end-of-the-period payments which can be conveniently simulated thanks to the closed-form stock price.

Following convention, option prices are quoted in terms of the implied Black-Scholes volatility (B/S-vol) computed as:

$$\text{B/S-vol}_t = BSC^{-1}(\tau, M_t, r_{t,t+\tau}, d_{t,t+\tau}) \tag{1.17}$$

where BSC^{-1} is the inverse of the Black-Scholes formula for the put option, inverted over the argument σ; $r_{t,t+\tau}$ and $d_{t,t+\tau}$ are the interest rate and the dividend-price ratio between

t and $t + \tau$. In the following, I fix $r_{t,t+\tau}$ and $d_{t,t+\tau}$ at 1.4% and 2.3%, i.e., their average values in my option sample period, for all t and τ.

To illustrate the combined effects of peso problem situation and habit formation on the generated option prices, I plot in the upper Panel of Figure 2 the volatility smirks from options with 30 days to expiration that are implied from MSV and my baseline model, where the non-jump parameters are from Table 1 of MSV[13]; the potential disaster is assumed to strike once every 100 years in the form of a 10% consumption contraction, i.e., $\lambda = .01$, $e^{-b} - 1 = -10\%$. Facing a small-probability economic disaster with only a moderate jump size (relative to the usually calibrated 40-60% consumption contractions in peso problem models), the generated smirk premium leaps from 3.1% in MSV to 4.3% in my (baseline) model. I plot in the lower Panel of the same figure the implied volatility smirks from the model under another two jump scenarios with the arrival intensity and the jump size doubled respectively. As expected, the generated smirk patterns become even more pronounced with the smirk premia increased to about 5.5% in both scenarios. Intuitively, deep OTM put options have the attractive feature to hedge against potential stock market crashes, hence they are priced with significant premia relative to ATM options, referred to as the smirk premia, after factoring the potential disasters.

Whereas peso problem situation is important for generating the observed smirk premium which determines the size of the volatility smirk, habit formation is indispensable for matching the observed prices of ATM options which determine the level of the volatility smirk. As documented in Barro (2006), potential consumption jumps, even the very dramatic ones, turn out to be quantitatively insignificant for matching the observed return volatility, a result which he calls "excess-volatility puzzle". Consequently, peso problem models severely under-evaluate ATM options whose prices are tightly linked to the return volatility, as illustrated in the upper Panel of Figure 1. In contrast, habit formation is effective at generating a high return volatility due to the implied high risk aversion degree measured by η_t as defined in (1.4).[14] For example, the average return volatility implied from CC and MSV are 9% and 23%, respectively,[15] which are much higher than the 2-3%

[13] In particular, $\gamma = 1$ in MSV due to the log habit formation assumption.

[14] Unlike the standard consumption-based model, however, a high risk aversion degree in habit formation models do not necessarily imply a high volatility of the risk-free rates. For example, both MSV and my model are capable of matching the observed return volatility of the risk-free rates.

[15] CC report an average 15% return volatility. As pointed out by Watcher (2005), their numerical solutions are not accurate and the average return volatility implied from their calibration is in fact around 9%, which I have confirmed. Watcher (2005) also reports that a different calibration of CC's model

volatilities implied from the corresponding degenerated models without habit formation. By incorporating MSV habit specification with the peso problem situation, the generated average ATM option prices stay above 16% under all the three jump scenarios, which is close to the observed 18.7% average price in the data. My model is thus likely to simultaneously match both the size and the level of the volatility smirk. The point of this paper is to take advantage of this setup's pricing potentiality by formally estimating the economic disasters implied from the observed option data, which is the focus of the next three sections.

[Figure 2 goes approximately here]

II Data

A. Option data

I collect S&P 500 index option data from two datasets: CBOE from April 4, 1988 to December 29, 1995 and Ivy DB from January 4, 1996 to June 30th, 2005, which combine to cover a period of nearly 18 years. I exclude the six months following the 1987 market crash and the pre-crash period because the pronounced volatility smirks emerged only after the crash and have persisted ever since.[16] CBOE data are no longer publicly available; I instead use those from Eraker who independently computes the relevant option prices (see Eraker, 2004, for the details of the computation). Ivy DB data are available from OptionMetrics LLC, and option prices during the Ivy DB period are computed as the closing bid-ask averages following the usual convention (e.g., Bakshi et al, 1997; Pan, 2002; Buraschi et al, 2006)[17]

Several exclusion filters are applied to the combined data. To focus on options which seem to attract the greatest attention in the literature, options with maturities longer than 180 days, and with moneyness less than .9 and larger than 1.05 are dropped. Options with maturities less than 15 days are also dropped because too short term options are subject to liquidity related biases. I delete from the remaining data observations permitting obvious

produces an average return volatility of around 16%.

[16]Benzoni et al (2005) builds a model in a Bayesian setting which generates the regime shift from a relatively flat smirk before the 1987 market crash to a pronounced smirk after the crash.

[17]Only closing bids and asks are available in the Ivy DB data which are collected at 4 pm EST. To ensure synchronization, S&P 500 index used in this paper are also closing quotes collected at 4 pm EST.

arbitrage, i.e., those whose prices violate $C_t \geq \max(0, S_t - K)$, where S_t and K are the spot stock price and the strike price. Observations with prices less than 3/8 dollars are also deleted because the discreteness of quotes has a large impact on these observations. Finally, I use only OTM option data, which are known to be more liquid than ITM (in-the-money) options. Similar filtering policies are used in, for example, Bakshi et al (1997, 2000), Ait-Sahalia and Lo (1998), David and Veronesi (2002), and Eraker (2004).

Table I reports the summary statistics of the option data collected from both CBOE and Ivy DB. Both the average B/S-vols and the number of options (in the parentheses) are reported for different option categories according to their maturities and moneyness. We observe a pronounced smirk pattern in both datasets and across all three maturity categories. Take the short term options with 15-45 days to expiration for example: deep OTM put options (with moneyness between .92-.94) are priced at the premia of 8-12% relative to ATM options; the premia decrease with the maturity but always stay above 4.4%. An important fact is that there is much lower option trading in the earlier CBOE period (with the total trading number of 26,465) than in the more recent Ivy DB period (with the total trading number of 87,134), although the two periods are of similar lengths. This is particularly true for deep OTM put options: the number of these options traded during the CBOE period is about one order of magnitude smaller than that during the Ivy DB period in two of the three maturity categories.

[Table I goes approximately here]

B. Other data

By convention, consumption and the risk-free rate are proxied by the sum of nondurables and services and the 30 day T-bill rate, respectively. Monthly consumption, CPI and population data are from the FRED database at the St. Louis Fed. Monthly 30 day T-bill rates, S&P 500 index values, total market values of the S&P 500 firms denoted by P_t^{SP500}, ex-dividend and cum-dividend S&P 500 returns, denoted respectively by R_t^d and R_t^x, and the daily S&P 500 index values are all from CRSP. The range of the monthly data is from January 1959 to December 2005, and the range of the daily S&P 500 index coincides with that of the combined option data. Equity premia are computed as the cum-dividend returns of the S&P 500 index minus the 30 day T-bill rates. Monthly dividend-price ratios for the S&P 500 firm, $\frac{D_t}{P_t}$, are computed out of the difference between R_t^d and R_t^x and smoothed using 12 months trailing averages as discussed in Hansen, Heaton and Li (2004). Monthly series of dividends are computed as the product of P_t^{SP500} and the smoothed $\frac{D_t}{P_t}$, which will be used for the estimation of the two alternative models in Section V. Finally I use CPI and population data to convert nominal quantities into real quantities and real quantities per capita, whenever necessary. Following the method used by Fama and French for constructing the daily risk-free rates for their dataset, I compute the daily CPI, risk-free rates and dividend-price ratios as their monthly counterparts divided by the number of trading days in that month, which will be used, together with the daily option data, for the model estimations.[18]

III Methodology

The model is estimated in two steps. In the first step, I use the monthly consumption data to estimate the non-jump parameters in the consumption process, i.e., μ and σ, which are jointly denoted by θ_1. Treating the θ_1 estimate as given, in the second step I use the joint daily data of options, stocks and risk-free rates to estimate the remaining parameters, i.e., those associated with the potential economic disasters and the assumed preference (including the assumed habit specification), which are jointly denoted by θ_2. Any loss of efficiency as a result of this approach is expected to be small because μ and σ play a minor role in pricing options and stocks. In both steps, I apply the method of simulated moments (MSM) to implement the estimation which selects parameter values to best match the simulated moments, computed using a long simulated series obtained from the assumed data generating mechanism, to their data counterparts. For the technical details about MSM, see, for example, Duffie and Singleton (1993) and Gourieroux and Monfort (1996).

A. Consumption

In the baseline model, parameters in θ_1 denote the mean and the standard deviation of the consumption growth rate conditional on no economic disasters. The two U.S. economic

[18]Daily risk-free rates and CPI are not available. Although unsmoothed daily dividend-price ratios can be computed out of the daily index returns, it is generally difficult to get it smoothed. For example, I can no longer apply the methodology in Hansen, Heaton and Li (2004) because of the days when none of the S&P 500 firms distribute dividends.

downturns documented in Barro (2006) during the past 100 years both occured beyond my consumption period, hence the estimation of θ_1 using the monthly consumption data. Applying the Euler approximation to (1.8) without the jump term yields

$$\frac{C_{s+\Delta} - C_t}{C_s} = \mu\Delta + \sigma\sqrt{\Delta}W_1 \tag{3.1}$$

where Δ is the annualized time interval; W_1 is a standard normal error. Since C_t in the model denotes the consumption rate, the model-implied monthly consumption is approximated by:

$$C_s^{\theta_1} \simeq \Delta \cdot \sum_{j=0}^{n-1} C_{s-j\Delta} \tag{3.2}$$

where $\Delta = \frac{1}{12n}$. I pick $n = 5$ and the increase of n has little effect on the estimation results.

The simulated moments are computed out of $\{C_s^{\theta_1}\}_{s=0}^{t}$, which are then matched to their data counterparts. The moments that I use are the unconditional mean of the consumption growth rate and its square. With two moments to estimate the two unknowns in θ_1, I achieve an exact match. When computing the standard errors, I need to account for the serial correlation induced by the aggregation in (3.2). Working (1960) shows that a special linear form of serial correlation can be derived in closed-form. Due to the nonlinearity in the levels evolution for consumption, I'm compelled to make this adjustment numerically using the Newey-West.

B. Preferences and the potential disasters

Treating the θ_1 estimates as given, in the second step I use the joint daily data of options, stocks and the risk-free rates to estimate the remaining parameters associated with the potential disasters and the assumed preference, which are jointly denoted by $\theta_2 \equiv [\gamma, \rho, \bar{G}, k, \beta, \alpha, \lambda, b]$.

B.1. Model implied option prices

Applying Euler discretization after substituting (1.7) into (1.6) yields

$$G_{s+\Delta} - G_s = k(\bar{G} - G_s)\Delta - \alpha(G_s - \beta)\left[\sigma\sqrt{\Delta}W_1 - b(J_{s+\Delta} - \lambda)\right] \tag{3.3}$$

$$\frac{C_{s+\Delta} - C_s}{C_s} = \mu\Delta + \sigma\sqrt{\Delta}W_1 + (e^{-b} - 1)J_{s+\Delta} \tag{3.4}$$

where Δ and W_1 are defined the same way as in (3.1); $J_{s+\Delta} = 1$ with probability $\lambda\Delta$ indicating that a jump occurs between s and $s + \Delta$, and $J_{s+\Delta} = 0$ otherwise. Given the maturity τ, the moneyness M_t, and the starting value of G at G_t, I obtain the discounted end-of-the-period payment from the option, whose formula is given by (1.16) (within the expectation sign), for a monthly sample of $\left\{\left(\frac{C_{t+s}}{C_t}, G_{t+s}\right)\right\}_{s=0}^{\tau}$ simulated according to the data generating mechanism in (3.3)-(3.4). I repeat to obtain 10,000 independent random draws of the discounted end-of-the-period payments, and I use their average as the model price of the option contract with maturity τ and moneyness M_t when the underlying economic state is G_t, which is denoted by $O(G_t; \theta_2)$, where I suppress its dependence on θ_1, τ, and M_t. The biases due to the simulation are found to be negligible because a futher decrease of Δ and/or a futher increase of the number of the random draws produce almost identical model-implied option prices.

Option prices thus computed are dependent on the underlying economic state G. Since G is not observable in the actual data, only unconditional average option prices, i.e., $E[O(G_t; \theta_2)]$, where the expectation is with respect to G_t, can be used in the estimation. The stationary distribution of G_t cannot be derived in closed-form at the presence of potential disasters, and I obtain its numerical approximation using N simulated realizations of G in its stationary distribution region, i.e., $\{G^i\}_{i=1}^N$. I choose $N = 4,000$, and larger N produces almost identical numerical distributions. $E[O(G_t; \theta_2)]$ is thus approximated by $\frac{1}{N}\sum_{i=1}^N O(G^i; \theta_2)$, where I need to compute $O(G^i; \theta_2)$ for each of the realized G^i ($i = 1, 2, ...N$). If all $O(G_t^i; \Theta_2)$s were computed using the Monte Carlo simulation as described in last paragraph, it will take roughly one hour to compute just one approximated $E[O(G_t; \theta_2)]$ for $N = 4 \times 10^3$ based on a 1.6 GHz CPU. As a result, the whole estimation time is likely to be on the order of months!

To make the estimation feasible, I instead combine interpolation with simulation to compute $O(G^i; \theta_2)$. I first compute $\{S^i\}_{i=1}^N$ implied out of the simulated $\{G^i\}_{i=1}^N$ according to (1.5). Unlike G_t, the surplus consumption ratio is theoretically bounded both from below and from above, and I use $[S_{\min,} S_{\max}]$ to denote the range of the implied $\{S^i\}_{i=1}^N$. I then choose K collocation nodes,[19] $S_{(k)}$ ($k = 1, 2, ...K$), within $[S_{\min,} S_{\max}]$, and I simu-

[19] The choice of collocation nodes depends on the functional family that is used for interpolation. Here I use Chebyshev interpolation which works the best for most smooth functions. For detailed discussions

late $O(S_{(k)};\theta_2)$ for each $S_{(k)}$. $O(G^i;\theta_2) = O(S^i;\theta_2)$ is interpolated out of the simulated $\{O(S_{(k)};\theta_2)\}_{k=1}^{K}$ for each $i = 1, 2, ...N$. The extra biases introduced by interpolation are small because option prices are smooth functions of the state S in the model. For example, the differences between the simulated and the interpolated prices divided by the simulated prices are generally below 0.2% with $K = 8$. With the introduction of interpolation, It only requires $K << N$ simulations plus N interpolations to compute one approximated $E[O(G_t;\theta_2)]$, which is much faster than running N simulations. As a result, the estimation time is reduced from several months to a mere several hours.

B.2. Moment choices and the optimization problem

From (1.9), both peso problem situation and habit formation are directly priced through the pricing kernel. I therefore adopt an integrated approach of using the joint data of options, stocks and the risk-free rates to simultaneously estimate all the parameters in θ_2. For non-option pricing moments, I choose to match the (unconditional) averages of the dividend-price ratio, the risk-free rate and the equity premium. Option pricing moments used in the estimation are only from the short term option contracts with 30 days to expiration, which are known to generate the most pronounced volatility smirk in the data (e.g., Liu et al, 2004; Benzoni et al, 2005). Both OTM put options and ATM options are used for the estimation, and I exclude OTM call options because their prices are close to those of ATM options. In particular, I choose to match the average prices of short term options in nine moneyness categories: $\bar{M} \equiv [0.92, 0.93, 0.94, 0.95, 0.96, 0.97, 0.98, 0.99, 1]$. In summary, I have a total of twelve asset pricing moments to estimate eight unknown parameters in θ_2.

Daily option data for the estimation are selected according to the following policy. Among all available options on the nth trading day, I first select those with maturities closest to 30 days. From the pool of options thus chosen, I then select the one with moneyness closest to each of the elements in $\bar{M}$. Due to the low trading of deep OTM put options (with moneyness between 0.92-0.94) during the CBOE period, most "deep OTM put options" thus selected from the CBOE data have the actual moneyness well above 0.94, which is likely to introduce large biases to the data prices. One possible solution is to exclude the data of deep OTM put options, which I tried in an unreported

about interpolation, see Miranda and Fackler (2002).

exercise. Estimation in this way, however, yields jump estimates with very large standard errors and an average smirk premium only half of its data counterpart, suggesting that the observed pricing differentials between deep OTM put options and ATM options might provide important information for the identification of the potential disasters. I thus use Ivy DB data only for the estimation. CBOE data together with the out-of-the-sample Ivy DB data will be used later to test the model's asset pricing implications.

Figure 3 plots the time series of the maturity and a subset of the moneyness for the options thus selected during the Ivy DB period. Both maturity and moneyness vary across time around their targeted values with only moderate derivations: maturities generally stay within 20-40 days, and moneyness varies mainly between $M - .005$ and $M + .005$, where M is the targeted moneyness. In addition, we observe roughly the same amount of positive and negative derivations, hence biases due to the varying contract variables are likely to be averaged out for the unconditional option pricing moments.

[Figure 3 goes approximately here]

Assume θ_2 live in a compact space, whose values are selected to minimize the distance between the sample moments and the simulated moments using the efficient GMM metrics:

$$\hat{\theta}_2 = \arg\min_{\theta_2} g_T(\theta_2)' W_T g_T(\theta_2) \tag{3.5}$$

where

$$g_T(\theta_2) \equiv \quad \tfrac{1}{T}\textstyle\sum_{t=1}^{T} P_t^{data} - \tfrac{1}{\Im(T)}\textstyle\sum_{t=1}^{\Im(T)} P_t^{\theta_2}; \tag{3.6}$$

T and $\Im(T)$ are the sample sizes of the actual data and the simulated data, respectively; W_T is the optimal weighting matrix; P_t^{data} and $P_t^{\theta_2}$ denote the twelve asset pricing moments computed from the actual data and the simulated data, respectively.

I set $\frac{\Im(T)}{T}$ to 5. As shown by Gourieroux and Monfort (1996), even a small $\frac{\Im(T)}{T}$ ratio can achieve practically sufficient level of efficiency. To save the computation time, W_T is estimated using the actual data instead of the simulated data. MSM thus applied is a special case of the general moment method (GMM) developed by Hansen (1982), but the estimated θ_2 are still consistent and asymptotically efficient (e.g., Duffie and Singleton, 1996). I use the global optimizer, OQNLP, provided by Tomlab[20] to solve the optimization

[20] Tomlab is a platform for solving applied optimization problems in Matlab. Among the global min-

problem in (3.5)-(3.6), which yields stable estimations for various starting values of θ_2.

IV Estimation results

The estimation results are reported in Table II. Panel A reports the estimated non-jump parameters; Panel B reports the derived values about habit formation; Panel C reports the estimated potential disasters; and Panel D reports the goodness-of-fit tests. I also list for certain variables in Panel B and C their corresponding values in Barro (2006), CC and MSV for the purpose of comparsion. Standard errors are reported in the parentheses.

A. non-jump parameters

Consistent with many previous findings, the average and the standard deviation of the consumption growth rate are both around 2%. The γ estimate implies that the estimated preference is not statistically different from the log habit formation setup assumed in MSV. The 1.21% annual subjective time-discount rate is well within the range of their usually assumed values in the macroeconomics literature. Of the four habit formation parameters, asset prices are much more sensitive to k and α than to $\bar{G}$ and β. As a result, k and α are estimated with much higher precisions.

To compare my estimated habit formation with those in the previous models, I report in Panel B $S_{\max}$ and mean(η_t), denoting respectively the maximum value of the surplus consumption ratio and the average local curvature of the utility function, that are implied from my model as well as from CC and MSV in which the representative agent faces no potential economic disasters. While my $S_{\max}$ lies between the values of its two counterparts, the implied average utility curvature from my model is much lower than those from both CC and MSV. Intuitively, habit formation pushes up equity premium by increasing the effective risk aversion degree measured by η_t. Due to the potential economic disasters which help boost the demanded equity premium as well, a relatively low degree of risk aversion in my model is sufficient to resolve the equity premium puzzle.

imizers provided by Tomlab, OQNLP yields the smallest objective value for my problem within a given period of time.

B. Economic downturns

The estimated potential disasters are reported in Panel C of Table II. The standard errors of the two jump estimates, λ and b, are relatively large, which is consistent with the previous estimations with the reduced-form option pricing models that jump parameters are relatively hard to be accurately identified (e.g., Bates, 2000; Pan, 2002; Eraker, 2004). I report in the last two columns the implied consumption jump size, $e^{-b}-1$, and the induced average jump size in the stock market, mean(J_P). For comparsion, I also list in the last row of Panel C the corresponding values from Barro (2006). S&P 500 index option data through my baseline model imply a potential economic disaster that strikes once every 50 years in the form of a 17.6% consumption contraction, which induces an average 56% stock market crash. Whereas the estimated arrival intensity is close to that in Barro (2006), the estimated consumption jump size is less than half of his estimation. In contrast, my model generates an average equity premium of 6%, close to what is observed in the data, which is much higher than the 3.5% equity premium implied from Barro (2006). Intuitively, with habit formation pushing up the equity premium, it is possible to resolve the equity premium puzzle with a potential economic downturn much less dramatic than the usually demanded 40-60% consumption jumps in peso problem models.

Of the economic disasters documented in Barro (2006), those with more than 40% contractions in consumption or GDP (in terms of real per capita, accumulated over the entire crisis) all occurred outside of the U.S. The two U.S. economic disasters that Barro (2006) documents during the past 100 years are the Great Depression (1929-1933) and the aftermaths of WWII (1944-1947), with accumulated contractions of 31% and 28% in GDP, respectively. In a model where all parameters are annualized, it seems reasonable (e.g., Longstaff and Piazzesi, 2006) to interprete the estimated consumption jump size as the average or the maximum annual consumption contraction.[21] The maximum annual consumption contraction during the Great Depression is 9.9%, and the average annual consumption contractions during the two documented U.S. economic downturns are both around 8%. My estimated consumption jump size (17.6%), while still look dramatic compared with the two annualized numbers, are much more line with those during the historically documented economic disasters than those implied from peso problem models.

[21] Instead, Barro (2006) advocates the calibration of consumption jump size using the accumulated falls during the entire crisis.

C. Stock market crashes

In my model, as in peso problem models, stock market crashes are induced by the economic disasters, hence they also strike once every 50 years[22]. Unlike peso problem models, economic disasters in my model induce the crashes in the stock market through two channels: the usual channel due to the negative jumps in the dividends with the same magnitudes as the consumption jumps, which induces jumps of the same size in the stock market; the extra channel due to the positive jumps in the agent's instantaneous risk aversion degree at the presence of habit formation, as discussed in Section I.C, which causes excessive depressions in the stock price relative to dividends. As a result, the stock price jump sizes are on average around -56%, which is much larger (in absolute values) than the estimated 17.6% consumption contractions. In contrast, without modeling habit formation, peso problem models imply the same jump sizes in both consumption and in the stock price. Empirical evidences, however, seem to support the prediction of excessive stock price adjustment. For example, Liu et al (2003) report that there are two major crashes in the U.S. stock market during the past 100 years: one is from mid-October to mid-November in 1929, and the other is the black Monday of October 19, 1987, when the Dow index fell by 44% and 23%, respectively. While these falls are accumulated within only one month or even one day, they are already much more dramatic than the maximum 9.9% annual consumption contractions during the Great Depression.

In the literature, stock price jumps are usually directy estimated out of the option pricing data with some reduced-form models (e.g., Bakshi et al, 1997; Bates, 2000; Pan, 2002). In those estimations, potential stock market crashes strike far more frequently from once every half of a year to once every four years, but with much smaller jumps sizes between -0.3% and -5.4%. The differences are not unexpected given the different setups with which the estimations are implemented. In addition, previous estimations are usually based on the data of ATM and near ATM options, partly because OTM put options are thinly traded during the earlier CBOE period. In contrast, OTM put options, in particular, the deep OTM put options, are found to play an important role for the identification of the potential jumps in my estimation.

[22] Following the convention of consumption-based asset pricing models, my model traces the dynamics in the stock market, including the potential stock market crashes, back to the dynamics of the underlying economic variables such as consumption. In the real world, however, consumption jumps and stock price jumps do not always happen at the same time, a phenomenon whose explanation is beyond the scope of this paper.

D. Goodness-of-fit tests

Panel D of Table II reports goodness-of-fit tests including four individual tests and the J_T test evaluating the overall fit of the model. For the individual tests, I report the differences between the values simulated from the model and their data counterparts for the average prices of 8% OTM put options ($O|_{M=.8}$), the average price of ATM options ($O|_{M=1}$), the smirk premium (SP) and the equity premium (EP), with standard errors reported in the parentheses. The model matches well the observed smirk premium and equity premium both economically and statistically. For example, the model-implied equity premium (6.04%) is only 0.62% below its data counterpart, and the standard error of the difference is 5.78. However, the model significantly underprices options across all moneyness including the reported 8% OTM put options and ATM options, which leads to its rejection at the 1% significance level. An important reason is that the estimation period covered by the Ivy DB data experiences unusually high return volatilities driven by the dot com bubble from the late 1990s to the early 2000s. Since prices of ATM options are tightly linked to the return volatility, the average prices computed using the Ivy DB data are likely to overestimate their true values in the long run.

It is worth mentioning that the implied unconditional volatility smirk from my model is quantitatively close to those implied from many previous models, including Pan (2002), who studies a reduced-form option pricing model with potential stock price jumps; Liu, Pan and Wang (2005), who model the uncertainty aversion towards the potential crashes in the aggregate endowment; and Benzoni, Collin-Dufresne and Goldstein (2005), who combines the Epstein-Zin preference with the consumption and dividend process driven by a persistent stochastic growth variable that may jump. In these three models, average prices of 8% OTM put options and ATM options are 21-24% and 13.5-15%, respectively, as compared to corresponding prices of 23.5% and 14.5% that are implied from my model. In addition, Pan (2002) reports the failure to reject that the volatility smirk implied from her model statistically matches the observed volatility smirk implied from the CBOE data.

[Table II goes approximately here]

V Alternative models

So far I've treated dividend and consumption as a single process. In the U.S., however, aggregate dividend only accounts for a small proportion of the aggregate consumption, and the two variables are only weakly correlated across time. I address this issue in this section by considering two alternative models with different and seemingly more realistic dividend processes. The two alternative models are then estimated using the similar methodology as discussed in Section III to check the robustness of the estimation results reported in Section IV.

A. Models

The simplest way to deviate from treating consumption and dividend as a single process is to assume a separate i.i.d. dividend process as follows:

$$\frac{dD_t}{D_t} = \mu_D dt + \sigma_D dB_{2t} + (e^{-b} - 1)dN_t \tag{5.1}$$

where B_{2t} is another standard Brownian whose correlation with B_{1t} is denoted by ρ_1. As in the baseline model, I assume that dividend jumps simultaneously with consumption by the same jump size of $e^{-b} - 1$. The discrete-time analog of the above dividend process without the potential jumps has been studied by CC to check the robustness of their model's stock pricing implications to different dividend specifications.

All the other assumptions remain the same as in the baseline model, and I call the new model thus generated "alternative 1". The following proposition summarizes its closed-form asset pricing implications.

Proposition 2 *Under the assumed utility in (1.1)-(1.2), the assumed dynamics of habit formation in (1.3)-(1.6), the assumed consumption processes in (1.8), and the assumed dividend process in (5.1)*

(a) The price-dividend ratio is given by

$$\frac{P_t}{D_t} = a_1 + a_2 \frac{1}{G_t} \tag{5.2}$$

where the ceofficients a_i $(i = 1, 2)$ are given by (A.35) in the Appendix.

***(b)** The risk-free rate (rf_t), the equity premium (EP_t) and the return volatility ($volR_t$) are given by:*

$$rf_t = -\mu_{\Lambda t} - \lambda J_{\Lambda t} \tag{5.3}$$

$$EP_t = \sigma_{\Lambda t} (\sigma_{Pt})_1 + \rho_1 \sigma_{\Lambda t} (\sigma_{Pt})_2 - \lambda J_{\Lambda t} J_{Pt} \tag{5.4}$$

$$volR_t = \sqrt{(\sigma_{Pt})_1^2 + (\sigma_{Pt})_2^2 + 2\rho_1 (\sigma_{Pt})_1 (\sigma_{Pt})_2 + \lambda J_{Pt}^2} \tag{5.5}$$

where $J_{Xt} \equiv \frac{X_t^+}{X_t} - 1$ ($X = \Lambda, P$) denote respectively the jump sizes of the pricing kernel and the stock price; $\mu_{\Lambda t}$, $-\sigma_{\Lambda t}$ are the drift and the diffusion terms in the process of $\frac{d\Lambda_t}{\Lambda_t}$; $(\sigma_{Pt})_i$ is the diffusion term associated with B_{it} ($i = 1, 2$) in the process of $\frac{dP_t}{P_t}$. $J_{\Lambda t}$, J_{Pt}, $\mu_{\Lambda t}$, $\sigma_{\Lambda t}$, and $(\sigma_{Pt})_i$ ($i = 1, 2$) are all closed-form functions of the state G_t, and their formulas are given by (A.39)-(A.44) in the Appendix.

I also consider a cointegrated model in which the dividend-consumption ratio, $F_t \equiv \frac{D_t}{C_t}$, is persistent. In particular, I assume F_t is mean-reverting as follows:

$$dF_t = \phi(\bar{F} - F_t)dt + F_t \sigma_F dB_{2t}, \tag{5.6}$$

where B_{2t} is again another standard Brownian. The standard deviation of the growth rate of F_t, $\frac{dF_t}{F_t}$, is assumed to be a constant σ_F, similarly to what is assumed for the consumption process. Still denote by ρ_1 the correlation between B_{2t} and B_{1t}, and combining (5.6) with (1.8) yields the dividend process:

$$\frac{dD_t}{D_t} = (\phi \frac{\bar{F} - F_t}{F_t} + \mu + \rho_1 \sigma \sigma_F)dt + \sigma dB_{1t} + \sigma_F dB_{2t} + (e^{-b} - 1)dN_t. \tag{5.7}$$

The implied dividend again jump simultaneously with the consumption by the same jump size of $e^{-b} - 1$. All the other assumptions remain the same as in the baseline model, and I call the new model thus generated "alternative 2", whose closed-form asset pricing implications are summarized in the following proposition:

Proposition 3 *Under the assumed utility in (1.1)-(1.2), the assumed dynamics of habit formation in (1.3)-(1.6), the assumed consumption processes in (1.8), and the assumed dividend-consumption ratio process in (5.6),* *(a) The price-dividend ratio is given by*

$$\frac{P_t}{D_t} = a_1 + a_2 \frac{1}{G_t} + a_3 \frac{1}{F_t} + a_4 \frac{1}{F_t \cdot G_t} \tag{5.8}$$

where the coefficients a_i $(i = 1, 2, 3, 4)$ *are given by (A.36) in the Appendix.*

(b) *The risk-free rate* (rf_t)*, the equity premium* (EP_t) *and the return volatility* $(volR_t)$ *are of the same formulas as those given by (5.3)-(5.5) for alternative 1, where the expressions for* $J_{\Lambda t}$*,* J_{Pt}*,* $\mu_{\Lambda t}$*,* $\sigma_{\Lambda t}$*, and* $(\sigma_{Pt})_i$ $(i = 1, 2)$ *are given by (A.39)-(A.41) and (A.45)-(A.47) in the Appendix.*

In both alternative models, the differences between the consumption and the dividend are interpreted as the labor income. As in the baseline model, the model-implied option prices have to be simulated based on the repeated random draws of the discounted end-of-the-period payment from the options. In particular, I need to simulate the joint distribution of G_t and F_t for computing the unconditional option pricing moments implied from alternative 2 due to the assumed persistent dividend-consumption ratios.[23]

B. Estimations

Both alternative models are estimated using the MSM in the similar way as decribed in Section III. I first estimate the non-jump dividend and consumption parameters, jointly denoted by θ_1, using the monthly consumption and dividend data. Given the estimated θ_1, I then use the joint daily data of options, stocks and the risk-free rates to estimate all the remaining parameters jointly denoted by θ_2. I use the same asset pricing moments as those with the baseline model for the estimation of θ_2. Due to the different dividend processes, I use the averages of $[\Delta c_t, \Delta c_t^2, \Delta c_t \Delta d_t, \Delta d_t, \Delta d_t^2]$ to estimate $\theta_1 = \{\mu, \sigma, \rho_1, \mu_D, \sigma_D\}$ in alternative 1, and the averages of $[\Delta c_t, \Delta c_t^2, \Delta c_t \Delta F_t, F_t F_{t-1}, F_t, \Delta F_t]$ to estimate $\theta_1 = \{\mu, \sigma, \rho_1, \phi, \bar{F}, \sigma_F\}$ in alternative 2, where $\Delta c_t \equiv \frac{C_t}{C_{t-1}} - 1$, $\Delta d_t \equiv \frac{div_t}{div_{t-1}} - 1$, $\Delta F_t \equiv \frac{F_t}{F_{t-1}} - 1$ denoting respectively the growth rates of the consumption, the growth rate of the dividend, and the growth rate of the dividend-consumption ratio.

The estimation results are reported in Table III with standard errors in the parentheses. Panel A and B report the non-jump parameters; Panel C reports the estimated potential

[23] In both the model and the data, the dividend-consumption ratio, F_t, varies within a narrow band of range and plays a minor role in the pricing of options.

disasters; and Panel D reports the goodness-of-fit tests. From Panel A, the estimated consumption and dividend with alternative 1 grow at the same speed, which justifies the cointegrated dividend-consumption ratio assumed in alternative 2. However, dividend is much more volatile than consumption, constitutes a small proportion of the aggregate consumption, and is only weakly correlated with consumption. Of the six preference parameters reported in Panel B, γ, ρ and α take similar values across all the three models. As in the baseline model, k and α are again identified with much higher precisions than $\bar{G}$ and β.

Potential economic disasters estimated with alternative 1 strike once every 64 years in the form of a 13.5% consumption contraction, which induces an average 36% stock market crash, roughly the average jump size of the two major stock market crashes documented in 1929 and 1987. Both the arrival intensity and the jump sizes are lower than those estimated with the baseline model, but the estimated consumption jumps still look dramatic compared to the no more than 10% annual consumption contractions observed during the Great Depression. Potential disasters estimated with alternative 2 strike more frequently at about once every 36 years and more severely in the form of a 15% consumption contraction, which induces a 51% stock market crash. Both numbers are close to those estimated with the baseline model, which is in fact a special case of alternative 2 by trivially fixing the dividend-consumption ratio at one.[24] Different dividend processes thus do not change the big picture of the estimated disasters when we use the identical specifications of peso problem situation and habit formation: potential economic disasters strike once every a few decades in the form of 15% or so consumption contractions, which induces even more dramatic crashes in the stock market; the implied consumption jumps and stock price jumps are both much more in line with the historical observations than those implied from peso problem models.

The Goodness-of-fit tests consist of the same four individual tests and the J_T test as those for the baseline model. All three models match well the average equity premium. The smirk premium implied from alternative 1 is about 2% lower than those from the other two but it is still not statistically different from what is observed in the data. Like the baseline model, both alternative models significantly underprice options across all moneyness due to

[24] To see it in a quick way, notice that $F_t \equiv 1$ implies $C_t = D_t$, hence from (5.8), $\frac{P_t}{C_t} = (a_1 + a_3) + (a_2 + a_4)\frac{1}{G_t}$, which is the price-dividend ratio implied from the baseline model given by (1.11) with a_3 and a_4 set to zeros.

the over-evaluation of the average option prices using the data during the Ivy DB period, hence the rejections in the J_T tests. Rejections of option pricing models, however, are not uncommon in the literature even for the reduced-form ones. For example, Bakshi et al (2000) apply MSM to estimate four reduced-form option pricing models that are commonly used in the option pricing literature using the daily option data from September 1, 1993 to August 31, 1994, and they find that all these models are rejected at the 1% significance level.

[Table III goes approximately here]

VI Asset pricing implications

This section explores the issue of model misspecification by examining my models' out-of-the-sample asset pricing implications. While the parameters are only chosen to match the unconditional moments and only short term options are used for the estimation, the three estimated models all perform well in matching many other asset pricing features, including volatilitiy smirks conditional on particular economic states implied from options with various maturities, term structure in option pricing, stock return predictability, and the historical fluctuations of smirk premia and dividend-price ratios. To my best knowledge, the matches of the term structure in option pricing and the historical smirk premia data have never been attempted before. These matches suggest that misspecification is a relatively minor issue for a setup incorporating both potential economic disasters and MSV habit specification intended to capture the joint pricing behaviors of stocks and options, hence an important support to the proposed methodology.

In the following, I use the baseline model to illustrate the out-of-the-sample matches, which I similarly find in the two alternative models. CC report that modeling a separate i.i.d. dividend process yields almost identical stock pricing implications as in the simplest case treating consumption and dividend as a single process. Whereas CC do not try a cointegrated model with a persistent dividend-consumption ratio, they comment: "any such model is likely to make the consumption and dividend claims more alike that in the simplest case, since it increases the correlation between dividends and consumption at long horizons". My findings extend their analysis by showing that in a setup with MSV habit specification, different dividend processes do not have significants effects on the implied option pricing implications as well.

A. Volatility smirks conditional on economic states

I first study volatility smirks conditional on the underlying economic states. Given the estimated structural parameters $\hat{\theta} \equiv \left(\hat{\theta}_1, \hat{\theta}_2\right)$, I first back out the daily surplus consumption ratio on each of the trading days within my option sample period using the short term (with maturites between 20 and 40 days) option data on the same day. Let N_t be the number of option prices on day t, and $O_n(t, \tau_n, M_n)$ and $\hat{O}_n(t, \tau_n, M_n, \hat{\theta}; S_t)$ $(n = 1, 2, ...N_t)$ be respectively the observed and the model price of the nth option, I next identify day t's state, $\hat{G}_t$, by minimizing the sum of the squared in-the-sample pricing errors as follows:

$$\min_{S_t} \sum_{n=1}^{N_t} \left[O_n(t, \tau_n, M_n) - \hat{O}_n(t, \tau_n, M_n, \hat{\theta}; G_t) \right]^2. \tag{6.1}$$

With the identified $\{\hat{G}_t\}$ as the inputs, I then compute the daily return volatility, $\{v\hat{o}lR_t\}$ by applying (1.14). $volR$ moves one for one with G within the range of the identified $\{\hat{G}_t\}$, hence in the following I use return volatility as the equivalent state variable, which is often exogenously assumed in the reduced-form option pricing models. (e.g., Bakshi et al, 1997; Pan, 2002).

Based on the computed $\{v\hat{o}lR_t\}$, I sort a total of 4,347 trading days within my option sample period into ten volatility deciles with days in the higher deciles experiencing higher return volatilities. To compute the average prices of options with moneyness M and maturity τ (in terms of days) conditional on a particular decile, I collect all available option contracts from days of that decile with moneyness and maturities falling into the ranges of $[M - .005, M + .005]$ and $[\tau - 10, \tau + 10]$. The data prices are computed as the average prices of all the option contracts thus collected, and different choices of the ranges are found to induce little variations in the average prices. In the following analysis, I focus on the sixth, the ninth, and the second deciles representing days of medium, high and low return volatilities, whose volatility ranges are reported in Panel A of Table IV. The corresponding model prices are simulated with the underlying state starting from the mid-point value of its range for each of the volatility deciles.[25]

Figure 4 plots together volatility smirks implied from both the (baseline) model and

[25] The computed model prices are similar whether we use G or the return volatility as the state variable.

the data during the three volatility deciles for options with maturities of 30 days, 60 days, 90 days and 120 days. The matches are generally good in terms of both the price levels of ATM options and the smirk premia, and the matches become even better with the increase of the return volatility. Take options with 90 days to expiration, for example. The smirk premium and the average price of ATM options are 4.9% and 14.0% in the model, and 6.2% and 15.3% in the data during the low volatility days. These numbers change to 6.05% and 26.1% in the model, and 6.22% and 26.6% in the data during the high volatility days, an apparently improved match. This is an attractive model feature to practitioners since options are usually more actively traded during periods of high return volatilities. In addition, the model-implied smirk premia decrease with the maturity during all the three volatility deciles, which is also consistent with the data as illustrated in Figure 4.

[Figure 4 goes approximately here]

B. Term structure of option pricing

While option pricing differentials across moneyness have been extensively studied in the literature, relatively little attention is paid to their pricing differentials across maturities, i.e., the term structure in option pricing. In the upper Panel of Figure 5, I plot model-implied option prices against their maturities under three moneyness scenarios during the medium volatility days. The observed term structure in option pricing is mostly decreasing for deep OTM put options with moneyness of .92; it becomes mostly increasing for lesser OTM put options with moneyness of .96; and the increasing pattern is even more pronounced for ATM options. To gauge this effect quantitatively, I define "term premium" as the price difference between options with 150 days to expiration and options with 30 days to expiration with the same degrees of moneyness. The data-implied term premia increase from -0.5% for deep OTM put options to 2.5% for lesser OTM put options, and even higher to 4.1% for ATM options.

My model captures this feature qualitatively, as plotted in dotted lines in the lower Panel of Figure 5. To illustrate the effects of the potential jumps, I plot in the same panel the model-implied term structures under another two jump scenarios with the two jump parameters, λ and b, taking values one standard error above, i.e., $(\hat{\lambda} + std(\hat{\lambda}), \hat{b} + std(\hat{b}))$, and one standard error below, i.e., $(\hat{\lambda} - std(\hat{\lambda}), \hat{b} - std(\hat{b}))$, their estimates of $(\hat{\lambda}, \hat{b})$. Under the scenario with more dramatic potential jumps, the generated term structure variations become quantitatively comparable to those observed in the data: the term premium starts at -1.6% for deep OTM put options; it increases to 1.2% for lesser OTM put options, and to 2.2% for ATM options. In contrast, the implied term structure is always decreasing and near flat under the other scenario with less dramatic potential jumps. I thus conclude that peso problem situation is important for generating the observed cross-sectional option pricing differentials not only along the dimension of moneyness but also along the dimension of maturity, an implication that was not explored before in the literature.

[Figure 5 goes approximately here]

C. Chi-square tests

This subsection statistically tests the matches of out-of-the-sample option prices. I first categorize options falling into each of the three volatility deciles into 8 classes according to their moneyness and maturities, hence a total of 24 option classes. I exclude short term options with maturities around 30 days, hence all options considered here are out-of-the-sample. Let I_l, $O_l(v\hat{o}l_i, \tau_i, M_i)$ and $\hat{O}_l(v\hat{o}l_i, \tau_i, M_i; \hat{\theta})$ be respectively the number of option contracts in class l, the observed and the model price of the ith option in class l $(l = 1, 2, ..., 24)$, where the model price is simulated with the underlying return volatility starting from $v\hat{o}l_i$, the realized volatility on the day when the corresponding data price is drawn, and is computed with the MSM parameter estimates, $\hat{\theta} = \{\hat{\theta}_1, \hat{\theta}_2\}$, as the inputs.

The chi-square test statistics are computed as

$$c_l = \left[\frac{1}{I_l} \sum_{i=1}^{I_l} \varepsilon_l \left(\hat{S}_i, \tau_i, M_i; \hat{\Theta} \right) \right]' \cdot \Omega_l^{-1} \left[\frac{1}{I_l} \sum_{i=1}^{I_l} \varepsilon_l \left(\hat{S}_i, \tau_i, M_i; \hat{\Theta} \right) \right] \tag{6.2}$$

where

$$\varepsilon_l \left(\hat{S}_i, \tau_i, M_i; \hat{\theta} \right) \equiv \frac{\hat{O}_l \left(\hat{S}_i, \tau_i, M_i; \hat{\Theta} \right)}{O_l \left(\hat{S}_i, \tau_i, M_i \right)} - 1 \tag{6.3}$$

is the pricing error, and the variance-covariance matrix Ω_l is estimated using 20 lags of the Newey-West. If the model is correct, the law of large numbers implies that the sample average of ε_l converges to zero. Hence for out-of-the-sample data, c_l is asymptotically distributed according to $\chi^2(1)$ for each of the 24 classes[26]. Panel B of Table IV reports the p-values of the c_l statistics together with the number of options (in parentheses) in each option class. The hypotheses that the model-implied option prices are equal to their data counterparts are not rejected with p-values all much larger than the usual significance levels. In addition, these tests formally verify the intuitive observations in Figure 4 that the matches improve with the increase of the return volatility.

[Table IV goes approximately here]

D. Historical smirk premia

In consumption-based asset pricing models, variations in asset prices are ultimately driven by consumption shocks. I test this implication applied to smirk premia in this subsection, which according to my knowledge was never tried before. I first obtain the Brownian innovations, $\{\Delta B^m_{t+1}\}$, from the historical monthly consumption data during my option sample period from April 1988 to June 2005. Next I use $\{\Delta B^m_{t+1}\}$ and the parameter estimates as the inputs to compute $\{G^m_t\}$, the time series of the equivalent economic states driven by the historical consumption shocks. The observed monthly option prices are collected in the middle of each month, and their model counterparts are simulated with the underlying state starting from the backed-out $\{G^m_t\}$ with the same moneyness and maturites as those in the observed data.

I consider three different measures of the smirk premium. Besides the usual measure as the price differentials between 8% OTM put options and ATM options, I also consider the premium demanded by 4% OTM put options and 4% OTM call options (with moneyness 1.04) relative to ATM options,

$$\left\{ \begin{array}{l} SP_1 = \text{B/S-vol}(M = .92) - \text{B/S-vol}(M = 1) \\ SP_2 = \text{B/S-vol}(M = .96) - \text{B/S-vol}(M = 1) \\ SP_3 = \text{B/S-vol}(M = 1.04) - \text{B/S-vol}(M = 1) \end{array} \right\} \tag{6.4}$$

For each of the three measures, denote by $\{SP^{data}_{i,t}\}$ and $\{SP^m_{i,t}\}$ $(i = 1, 2, 3)$ the monthly series of the smirk premia implied from the data and from the model, respectively. To see

[26] Similar tests are conducted by David and Veronesi (2002), and Buraschi and Jiltsov (2006). I also tried the tests with pricing errors measured as the absolute differences between the model prices and the observed prices, and the results are similar.

how well $\{SP^m_{i,t}\}$ capture the fluctuations in the observed $\{SP^{data}_{i,t}\}$, I run the following regressions:

$$SP^{data}_{i,t} = \beta_1 + \beta_2 SP^m_{i,t} + \epsilon_{i,t},\ i = 1, 2, 3 \tag{6.5}$$

In particular, I run each of the regressions for three different periods covered respectively by i) the combined data, ii) the CBOE data from April 1988 to December 1995, and iii) the Ivy DB data from January 1996 to June 2005. The regression results together with the correlations between $\{SP^{data}_{i,t}\}$ and $\{SP^m_{i,t}\}$ are reported in Table V. The β_2 coefficients are all economically significant with values between 20.5% and 88.1%. All but two of the coefficients are statistically significant as well. The correlations are all of impressive values and most R^2 measures are also sizable. It is acknowledged that the moneyness variations during the CBOE period are large due to the low trading of OTM put options, which can generate spurious matches to some extent. However, variations in both maturity and moneyness are moderate during the Ivy DB period and they all look random, as illustrated in Figure 3. Although for this reason the matches during the Ivy DB period are not as good as those during the CBOE period, they still indicate that a significant amount of fluctuations in the historical smirk premia are captured by the historical consumption shocks through the model.

[Table V goes approximately here]

E. Stock pricing

I list the matches of the observed stock pricing phenomena in Table VI, which repeat the well-known success of habit formation models. Panel A reports the matches of the average return volatility, the standard deviation of the risk-free rate, the average dividend-price ratio, the average risk-free rate, and the average equity premium, where their data counterparts are computed using the data during the Ivy DB period. The first two moments are out-of-the-sample, and the last three are in-the-sample, which are included to give us a complete picture of the matches of the stock pricing moments. The model captures the average values of the dividend-price ratio and the equity premium reasonably well; it avoids excess interest rate volatility sometimes present in internal habit formation models; and it somewhat under-evaluates the average values of the risk-free rate and the return volatility.

Panel B reports the model's implications about the excess stock return predictability by running the long-horizon regressions of log excess stock returns onto the log price-dividend ratio in the simulated data. We observe the classic pattern of stock return predictability firstly documented by Fama and French (1988): the coefficients are negative and significant both statistically and economically; they increase linearly at first and then less quickly; R^2 starts relatively low but then rises to impressive values.

Panel C reports the results of the test that variations in the historical dividend-price ratios are driven by the historical consumption shocks. I feed the model with historical monthly consumption data in the same way as described in Section VI.D, and then run the regression from the observed series, $\{DP_t^{data}\}$ onto the model-implied series, $\{DP_t^m\}$, for the period between Februray 1959 and December 2005. The regression coefficient is significant both economically and statisticaly with a sizable R^2, and the correlation between $\{DP_t^{data}\}$ and $\{DP_t^m\}$ is close to those reported by CC (0.32) and MSV (0.55), which are computed using the annual and the quarterly consumption shocks, respectively.

[Table VI goes approximately here]

VII Conclusions

Potential economic disasters, such as what happened during the Great Depression, can have profound effects on the empirical observations in the financial markets. However, empirically we lack direct observations to calibrate their magnitudes. Because disasters are rare, we usually have no more than one or two of such observations over a long period of time. In addition, while historically documented disasters can shed some evidence about what will happen in the future, they could be very different from the potential ones that investors are currently concerned about. Using the "peso problem" argument that potential economic disasters are likely to be factored into the observed asset prices, an alternative way is to back out their magnitudes implied from the financial data with models that link the dynamics of the real economy to those of the financial markets. The usual methodology is to add a small-probability consumption jump into the otherwise standard consumption-based asset pricing model, referred to as peso problem model in this paper, which is then calibrated to the observed equity premium. Potential jumps thus estimated, however, imply counterfactually high option pricing differentials between deep OTM put options and ATM options (known as the "smirk premium"), suggesting that the economic disasters required to rationalize the observed equity premium is too severe for the observed option prices.

In this paper, I propose a new methodology of using the S&P 500 index option data to gauge the magnitudes of the potential economic diasters, which are priced in the option markets through the induced potential stock market crashes in the form of the pricing differentials among index options across moneyness. The key point is to implement the estimation with a setup incorporating both the potential economic disasters in the form of negative consumption jumps and an external habit specification. Whereas both model elements are important devices to boost the equity premium, the former proves to be much more effective than the latter for generating the observed smirk premia. Hence, with habit formation pushing up the equity premium without affecting the smirk premium much, I can use the index option data to back out the implied economic disasters factored into the observed option pricing differentials across moneyness. Disasters thus estimated, combined with the estimated habit formation, tend to be simultaneously consistent with the observed equity premium and the observed smirk premium.

The estimated consumption jumps, with the jump sizes between 13.5% and 17.6% depending on the assumed dividend processes, are much more in line with the the historically documented no more than 10% annual consumption contraction during the Great Depression than those estimated with peso problem models in which 40-60% potential consumption contractions are usually demanded to match the observed equity premium. In addition, consumption jumps in my model induce the stock market crashes through an extra channel due the positive jumps in the agent's instantaneous risk aversion degree at the presence of habit formation, which causes excessive depressions in the stock price relative to dividends. As a result, unlike peso problem models which imply identical jump sizes in the consumption and in the stock price, the induced stock market crashes from my setup are much more dramatic than the estimated consumption jumps, which is also consistent with the empirical observations.

The paper contributes to the literature of asset pricing as well. The proposed methodology is based on the interesting new findings that peso problem models generate too pronounced volatility smirks whereas habit formation models generate too flat volatility smirks in option pricing. Combining peso problem situation with habit formation proves successful at generating a wide variety of the observed option pricing phenomena including the average volatility smirks, volatility smirks conditional on different economic states, the term structure in option pricing, and the historical fluctuations of smirk premia, as well as many stock pricing phenonema. To my best knowledge, the matches of term structure in option pricing and the historical fluctuations of smirk premia have never been attempted before in the literature. My setup significantly improves on the pricing abilities of the two previous consumption-based option pricing models, Liu et al (2005) and Benzoni et al (2005), which are silent about the observed option pricing variations across time and across different economic states. Whereas many of the matches can also be achieved using reduced-form models, it seems worthwhile to devise a setup with comparable pricing capabilities that links the dynamics of financial markets directly to those of the real economy, which according to Cochrane (2006) is "the trunk of finance". The asset pricing matches, besides being interesting of their own, relieves the concern about model misspecifications, which lends an important support to the proposed methodology

Appendix

I first prove the following lemma:

Lemma 1 *Consider a jump-diffusion process:*

$$dZ_t = (A_0 + A_1 Z_t)dt + \Sigma(t, Z_t)dB_t + A_2 Z_t(dN_t - \lambda dt) \quad \text{(A.1)}$$

where Z_t, A_0, A_1, $\Sigma(t, Z_t)$, dB_t, and A_2 are n by 1, n by 1, n by n, n by m, m by 1 and n by n, respectively; B_t and N_t are respectively an m-dimensional standard Brownian and a Poisson process with arrival intensity λ; $A_2 Z_t$ denotes the jump size of dZ_t. Under the process of (A.1) and assume certain technical conditions are satisfied[27]*, the following conditional expectations can be derived in closed-form,*

$$E_t Z_{t+\tau} = \Psi_\tau Z_t + \int_0^\tau \Psi_{\tau-s} A_0 ds, \ \tau \geq 0 \quad \text{(A.2)}$$

where $\Psi_\tau = U \exp(\omega\tau) U^{-1}$; $\exp(\omega\tau) \equiv diag[e^{\omega_1}, ..., e^{\omega_n}]^{28}$; U and $\omega \equiv [\omega_1, ..., \omega_n]'$ are eigenvectors and eigenvalues of A_1, i.e., $A_1 = U diag(\omega) U^{-1}$

[27] The technical conditions required are similar to (i)-(iii) assumed in Page 1351 of Duffie et al (2000).

[28] $diag([x_1, ...x_n])$ denotes an n by n diagonal matrix with $x_1, ...x_n$ on the principal diagonals and all off-diagonal elements set to zeros.

In particular, if $A_0 = 0_{n\times n}$,

$$E_t Z_{t+\tau} = \Psi_\tau Z_t \tag{A.3}$$

Proof. *Integrating on both sides of (A.1) yields:*

$$Z_u = Z_t + \int_t^u (A_0 + A_1 Z_s)ds + \int_t^u \Sigma(s, Z_s)dB_s + A_2 J_t \quad \text{for } u \geq t \tag{A.4}$$

where $J_t \equiv \sum_{t<\tau_i\leq u} (Z_{\tau_i} - Z_{\tau_{i-}}) - \lambda \int_t^u Z_s ds$

Using the same technique as that for the proof of lemma 1 in Duffie et al (2000), we can show that $E_t J_t = 0$.[29] *Assume technical conditions for ensuring* $E_t \int_t^u \Sigma(s, Z_s)dB_s = 0$ *are satisfied, and define* $\hat{Z}_u \equiv E_t Z_u$ *for* $u \geq t$*. Taking conditional expectation on both sides of (A.4) yields:*

$$\hat{Z}_u = \hat{Z}_t + \int_t^u (A_0 + A_1 \hat{Z}_s)ds \tag{A.5}$$

Next take derivatives with respect to u on both sides of (A.5) to yield

$$\frac{d\hat{Z}_u}{du} = A_0 + A_1 \hat{Z}_u, \text{ with } \hat{Z}_u|_{u=t} = \hat{Z}_t \tag{A.6}$$

(A.6) is a first order ordinary differential equation (ODE) in vector form, whose solution is given by

$$\hat{Z}_u = \Psi_{u-t}\hat{Z}_t + \int_t^u \Psi_{u-s}A_0 ds, \ u \geq t \tag{A.7}$$

To verify it, take derivatives with respect to u on both sides of (A.7) with the notice that $\exp(\omega(u-t))$ *is an n by n diagonal matrix,*

$$\begin{aligned}
\frac{d\hat{Z}_u}{du} &= U \exp(\omega(u-t))\,\omega U^{-1}\hat{Z}_t + A_0 + \int_t^u U\exp(\omega(u-s))\omega U^{-1}A_0 ds \\
&= A_0 + U\omega U^{-1}\left[U\exp(\omega(u-t))\omega U^{-1}\hat{Z}_t + \int_t^u U\exp(\omega(u-s))\omega U^{-1}A_0 ds\right] \\
&= A_0 + A_1\hat{Z}_u
\end{aligned}$$

where I've used the definition of Ψ *for first equality, and the definition of* U *and* ω *for the third equality. Finally substituting* $\hat{Z}_u \equiv E_t Z_u$ *into (A.7) and manipulating the time sub-indices yields (A.2).* ■

[29] Duffie et al (2000) considers a more complex case when the arrival intensity is stochastic and the jump size follows some general form of distribution, which includes my setup (constant arrival intensity, jump size proportional to the state variable) as a special case. Wheraas they only consider a scalar case, the extenstion to the vector case is straightforward.

Proof of Proposition 1. Define

$$y_t = \begin{bmatrix} y_{1t}, & y_{2t}, \end{bmatrix}' \equiv \begin{bmatrix} C_t^{1-\gamma} G_t, & C_t^{1-\gamma} \end{bmatrix}'$$

Applying Ito's lemma with jumps (e.g., Appendix F of Duffie, 2001) to y_t yields:

$$\begin{aligned} \frac{dy_{1t}}{y_{1t}} &= \left[\mu_2 + k \frac{\bar{G} - G_t}{G_t} - \alpha b \lambda \frac{G_t - \beta}{G_t} - (1-\gamma)\alpha\sigma^2 \frac{G_t - \beta}{G_t} \right] dt \\ &\quad + \left(1 - \gamma - \alpha \frac{G_t - \beta}{G_t} \right) \sigma dB_{1t} + \left[e^{-b(1-\gamma)} \left(1 + \alpha b \frac{G_t - \beta}{G_t} \right) - 1 \right] dN_t \end{aligned}$$

$$\frac{dy_{2t}}{y_{2t}} = \mu_2 dt + (1-\gamma)\sigma dB_{1t} + \left[e^{-b(1-\gamma)} - 1 \right] dN_t$$

or in vector form:

$$dy_t = A_1 y_{1t} + \Sigma(t, y_{1t}) dB_{1t} + A_2 y_{1t}(dN_t - \lambda t)$$

where $B_{2t} \equiv [B_{1t}, B_{2t}]'$; A_1 and A_2 are both 2 by 2 matrices with the non-zero elements given by

$$\begin{aligned} A_{1,11} &= \mu_2 - k - \alpha b \lambda - (1-\gamma)\alpha\sigma^2 + \lambda A_{2,11} \\ A_{1,12} &= k\bar{G} + \alpha\beta b\lambda + (1-\gamma)\alpha\beta\sigma^2 + \lambda A_{2,12} \\ A_{1,22} &= \mu_2 + \lambda A_{2,22} \end{aligned} \tag{A.8}$$

and

$$\begin{aligned} A_{2,11} &= e^{-b(1-\gamma)}(1 + \alpha b) - 1 \\ A_{2,12} &= -\alpha\beta b e^{-b(1-\gamma)}; \; A_{2,22} = e^{-b(1-\gamma)} - 1; \end{aligned} \tag{A.9}$$

where

$$\mu_2 \equiv (1-\gamma)\mu - \frac{1}{2}\gamma(1-\gamma)\sigma^2 \tag{A.10}$$

Decompose $A_1 = U\omega U^{-1}$, where $\omega \equiv diag[\omega_1, \omega_2]$ is a 2 by 2 diagonal matrix. For notational simplicity, in the following I use $f(\omega)$ to denote a 2 by 2 diagonal matrix $diag[f(\omega_1), f(\omega_2)]$ for any functional form $f(.)$.

From lemma 1:

$$E_t y_s = U e^{\omega(s-t)} U^{-1} \text{ for } s \geq t \tag{A.11}$$

Hence,

$$
\begin{aligned}
P_t &= E_t(\int_t^\infty \frac{\Lambda_s}{\Lambda_t} C_s ds) \\
&= C_t^\gamma G_t^{-1} E_t \int_t^\infty e^{-\rho(s-t)} C_s^{1-\gamma} G_s ds \\
&= C_t^\gamma G_t^{-1} \int_t^\infty E_t[e^{-\rho(s-t)} y_s^1] ds \\
&= C_t^\gamma G_t^{-1} \int_t^\infty e_1' E_t[e^{-\rho(s-t)} y_s] ds \\
&= C_t^\gamma G_t^{-1} \left(\int_t^\infty e_1' U e^{(\omega-\rho)(s-t)} U^{-1} ds \right) y_{1t} \\
&= C_t^\gamma G_t^{-1} \left(e_1' U \frac{1}{\rho-\omega} U^{-1} \right) y_{1t} \\
&= C_t^\gamma G_t^{-1} \sum_{i=1}^{2} a_i y_t^i = a_1 C_t + a_2 \frac{C_t}{G_t} \qquad \text{(A.12)}
\end{aligned}
$$

where $a_i \equiv e_1' U \frac{1}{\rho-\omega} U^{-1} e_i$, and e_i is a 2 by 1 vector with 1 in the ith $(i = 1, 2)$ entry and zero in the other; I've used the Fubini theorem for the third equality, (A.11) for the fifth equality, and the definitions of y_t^i $(i = 1, 2)$ for the last equality.

(A.12) gives the closed-form stock price in (1.10). To derive the expressions for a_i $(i = 1, 2)$, notice that $I\rho - A_1 = U(\rho - \omega)U^{-1}$ implies

$$
U \frac{1}{\rho-\omega} U^{-1} = (I\rho - A_1)^{-1}
$$

which means that $a_i \equiv e_1' U \frac{1}{\rho-\omega} U^{-1} e_i$ is just the $(1, i)$th element in $(I\rho - A_1)^{-1}$. Hence,

$$
\sum_{i=1}^{2} a_i (I\rho - A_1)_{(i,:)} = [1, 0] \qquad \text{(A.13)}
$$

where $(I\rho - A_1)_{(i,:)}$ represents the ith row of $I\rho - A_1$ $(i = 1, 2)$. (A.13) is equivalent to

$$
\begin{aligned}
&a_1(\rho - A_{1,11}) = 1 \\
&-a_1 A_{1,12} + a_2(\rho - A_{1,22}) = 0
\end{aligned}
$$

hence,

$$
\begin{aligned}
a_1 &= \tfrac{1}{\rho - A_{1,11}} \\
a_2 &= \tfrac{A_{1,12}}{(\rho - A_{1,22})(\rho - A_{1,11})}
\end{aligned} \qquad \text{(A.14)}
$$

where $A_{1,11}$, $A_{1,12}$ and $A_{1,22}$ are defined in (A.8)-(A.10). This completes the proof of part a) in Proposition 1.

To prove part b), notice that at the presence of jumps,

$$\frac{d\Lambda_t}{\Lambda_t} = \frac{d\Lambda_t^*}{\Lambda_t} + J_{\Lambda t} dN_t = \mu_{\Lambda t} dt - \sigma_{\Lambda t} dB_{1t} + J_{\Lambda t} dN_t \tag{A.15}$$

$$\frac{dP_t}{P_t} = \frac{dP_t^*}{P_t} + J_{Pt} dN = \mu_{Pt} dt + \sigma_{Pt} dB_{1t} + J_{Pt} dN_t \tag{A.16}$$

where $\frac{dX_t^*}{X_t}$ denotes the drift and the diffusion components in the dynamics of $\frac{dX_t}{X_t}$; $J_{Xt} \equiv \frac{X_t^+}{X_t} - 1$ denotes the jump sizes of X_t, where $X = \Lambda, P$. Following the asset pricing convention, the diffusion component in $\frac{d\Lambda_t}{\Lambda_t}$ is written as $-\sigma_{\Lambda t} dB_{1t}$ where $\sigma_{\Lambda t}$ is interpreted as the price of the diffusion risks.

Given the assumption that jumps are independent of the Brownian, the stock return volatility is by definition:

$$(volR_t)^2 \cdot dt = (\sigma_{P1t} dB_{1t})^2 + (J_{Pt})^2 \lambda dt \tag{A.17}$$

which yields (1.13). Similarly, the risk-free rate is by definition (e.g., Chapter 2 of Cochrane (2001)),

$$rf_t dt = -E_t \left(\frac{d\Lambda_t}{\Lambda_t} \right) = -\mu_{\Lambda t} dt - J_{\Lambda t} \lambda dt \tag{A.18}$$

which yields (1.11). At the presence of potential jumps, the formula for the equity premium is derived by (A.14) in Longstaff and Piazzesi (2004),

$$\begin{aligned} EP_t dt &\equiv -E_t \left(\frac{dP_t^*}{P_t} \frac{d\Lambda_t^*}{\Lambda_t} \right) - \lambda J_P J_\Lambda dt \\ &= \sigma_{Pt} \sigma_{\Lambda t} dt - \lambda J_P J_\Lambda dt \end{aligned} \tag{A.19}$$

which yields (1.12). Finally, applying Ito's lemma with jumps to the expressions of Λ_t and P_t in (1.8) and (1.10) yields the formulas for $\mu_{\Lambda t}$, $\sigma_{\Lambda t}$, $J_{\Lambda t}$, σ_{Pt} and J_{Pt}

$$\mu_{\Lambda t} = -\rho + \mu_1 + k \frac{\bar{G} - G_t}{G_t} - \alpha b \lambda \frac{G_t - \beta}{G_t} + \gamma \alpha \sigma^2 \frac{G_t - \beta}{G_t} \tag{A.20}$$

$$\sigma_{\Lambda t} = \sigma \left(\gamma + \alpha \frac{G_t - \beta}{G_t} \right) \tag{A.21}$$

$$J_{\Lambda t} \equiv e^{b\gamma} \left(1 + \alpha b \frac{G_t - \beta}{G_t} \right) - 1 \tag{A.22}$$

$$\sigma_{Pt} = \sigma \left[a_1 \frac{C_t}{P_t} + a_2 \left(1 + \alpha \frac{G_t - \beta}{G_t} \right) \frac{1}{G_t} \frac{C_t}{P_t} \right] \tag{A.23}$$

$$J_{Pt} = a_1 \frac{C_t}{P_t}(e^{-b} - 1) + a_2 \frac{1}{G_t} \frac{C_t}{P_t} \left(\frac{e^{-b}}{1 + \alpha b \frac{G_t - \beta}{G_t}} - 1 \right) \tag{A.24}$$

where I've used the assumed dynamics of G_t and C_t given in (1.5) and (1.7). ■

Proof of Proposition 2 and 3. The proof procedures of the these two propositions are similar to that of Proposition 1 but take more maths. Define

$$y_t = \left[\; y_{1t}, \quad y_{2t}, \; \right]' \equiv \left[C_t^{-\gamma} G_t D_t, \quad C_t^{-\gamma} D_t, \right]'$$

for alternative 1, and

$$y_t = \left[y_{1t}, \quad y_{2t}, \quad y_{1t}, \quad y_{4t} \right]' \equiv \left[\; C_t^{1-\gamma} G_t F_t, \quad C_t^{1-\gamma} F_t, \quad C_t^{1-\gamma} G_t, \quad C_t^{1-\gamma} \; \right]'$$

for alternative 2. Applying Ito's lemma with jumps to y_t yields in alternative 1,

$$\begin{aligned} \frac{dy_{1t}}{y_{1t}} \quad = \quad & \left(\mu_1 + k \frac{\bar{G} - G_t}{G_t} - \alpha b \lambda \frac{G_t - \beta}{G_t} + \mu_D + \gamma \alpha \sigma^2 \frac{G_t - \beta}{G_t} \right) dt \\ & - \left(\gamma \rho_1 \sigma \sigma_D - \alpha \rho_1 \sigma \sigma_D \frac{G_t - \beta}{G_t} \right) dt - \left(\gamma + \alpha \frac{G_t - \beta}{G_t} \right) \sigma dB_{1t} \\ & + \sigma_D dB_{2t} + \left[e^{(b-1)\gamma} \left(1 + \alpha b \frac{G_t - \beta}{G_t} \right) - 1 \right] dN_t \end{aligned}$$

$$\frac{dy_{2t}}{y_{2t}} = (\mu_1 + \mu_D - \gamma \rho_1 \sigma \sigma_D) dt - \gamma \sigma dB_{1t} + \sigma_D dB_{2t} + (e^{(b-1)\gamma} - 1) dN_t$$

or in vector form

$$dy_t = A_1 y_t + \Sigma(t, y_t) dB_t + A_2 y_t (dN_t - \lambda dt)$$

where $B_t \equiv [B_{1t}, B_{2t}]'$, and the non-zero elements in A_1 and A_2 are given by

$$\begin{aligned} A_{1,11} &= \mu_1 - k - \alpha b \lambda + \mu_D + \gamma \alpha \sigma^2 - (\alpha + \gamma) \rho_1 \sigma \sigma_D + \lambda A_{2,11} \\ A_{1,12} &= k \bar{G} + \alpha b \lambda \beta - \gamma \alpha \beta \sigma^2 + \alpha \beta \rho_1 \sigma \sigma_D + \lambda A_{2,12} \\ A_{1,22} &= \mu_1 + \mu_D - \gamma \rho_1 \sigma \sigma_D + \lambda A_{2,22} \end{aligned} \tag{A.25}$$

and

$$\begin{aligned} A_{2,11} &= e^{(b-1)\gamma}(1 + \alpha b) - 1 \\ A_{2,12} &= -\alpha \beta b e^{(b-1)\gamma}; \;\; A^1_{2,22} = e^{(b-1)\gamma} - 1; \end{aligned} \tag{A.26}$$

where

$$\mu_1 \equiv -\gamma \mu + \frac{1}{2} \gamma (\gamma + 1) \sigma^2 \tag{A.27}$$

Similarly in alternative 2,

$$
\begin{aligned}
\frac{dy_{1t}}{y_{1t}} &= \left[\mu_2 + k\frac{\bar{G}-G_t}{G_t} - \alpha b\lambda\frac{G_t-\beta}{G_t} + \phi\frac{\bar{F}-F_t}{F_t} - (1-\gamma)\alpha\sigma^2\frac{G_t-\beta}{G_t}\right]dt \\
&\left[(1-\gamma)\rho_1\sigma\sigma_F - \alpha\rho_1\sigma\sigma_F\frac{G_t-\beta}{G_t}\right]dt + \left(1-\gamma-\alpha\frac{G_t-\beta}{G_t}\right)\sigma dB_{1t} \\
&+\sigma_F dB_{2t} + \left[e^{-b(1-\gamma)}\left(1+\alpha b\frac{G_t-\beta}{G_t}\right)-1\right]dN_t
\end{aligned}
$$

$$
\begin{aligned}
\frac{dy_{2t}}{y_{2t}} &= \left[\mu_2 + \phi\frac{\hat{F}-F_t}{F_t} + (1-\gamma)\rho_1\sigma\sigma_F\right]dt \\
&+(1-\gamma)\sigma dB_{1t} + \sigma_F dB_{2t} + \left[e^{-b(1-\gamma)}-1\right]dN_t
\end{aligned}
$$

$$
\begin{aligned}
\frac{dy_{3t}}{y_{3t}} &= \left[\mu_2 + k\frac{\bar{G}-G_t}{G_t} - \alpha b\lambda\frac{G_t-\beta}{G_t} - (1-\gamma)\alpha\sigma^2\frac{G_t-\beta}{G_t}\right]dt \\
&+\left(1-\gamma-\alpha\frac{G_t-\beta}{G_t}\right)\sigma dB_{1t}B_t^1 + \left[e^{-b(1-\gamma)}\left(1+\alpha b\frac{G_t-\beta}{G_t}\right)-1\right]dN_t
\end{aligned}
$$

$$
\frac{dy_{4t}}{y_{4t}} = \mu_2 dt + (1-\gamma)\sigma dB_{1t} + \left(e^{-b(1-\gamma)}-1\right)dN_t
$$

or in vector form:

$$
dy_t = A_1 y_t + \Sigma(t, y_t)dB_t + A_2 y_t(dN_t - \lambda t)
$$

where $B_t \equiv [B_{1t}, B_{2t}]'$, and the non-zero elements in A_1 and A_2 are given by

$$
\begin{aligned}
&A_{1,11} = \mu_2 - k - \alpha b\lambda - \phi - (1-\gamma)\alpha\sigma^2 \\
&+(1-\gamma)\rho_1\sigma\sigma_F - \alpha\rho_1\sigma\sigma_F + \lambda A_{2,11} \\
&A_{1,12} = k\bar{G} + \alpha b\lambda\beta + (1-\gamma)\alpha\beta\sigma^2 \\
&+\alpha\beta\rho_1\sigma\sigma_F + \lambda A_{2,12};\ A_{1,13} = A_{1,24} = \phi\bar{F}; \\
&A_{1,22} = \mu_2 - \phi + (1-\gamma)\rho_1\sigma\sigma_F + +\lambda A_{2,22} \\
&A_{1,33} = \mu_2 - k - \alpha b\lambda - (1-\gamma)\alpha\sigma^2 + \lambda A_{2,33} \\
&A_{1,34} = k\bar{G} + \alpha b\lambda\beta + (1-\gamma)\alpha\beta\sigma^2 + \lambda A_{2,34} \\
&A_{1,44} = \mu_2 + \lambda A_{2,44}
\end{aligned}
\tag{A.28}
$$

and

$$
\begin{aligned}
&A_{2,11} = e^{-b(1-\gamma)}(1+\alpha b) - 1 \\
&A_{2,12} = -\alpha\beta b e^{-b(1-\gamma)};\ A_{2,22}^1 = e^{-b(1-\gamma)} - 1; \\
&A_{2,33} = e^{-b(1-\gamma)}(1+\alpha b) - 1;\ A_{2,34} = -\alpha\beta b e^{-b(1-\gamma)}; \\
&A_{2,44} = e^{-b(1-\gamma)} - 1;
\end{aligned}
\tag{A.29}
$$

where

$$\mu_2 \equiv (1-\gamma)\mu - \frac{1}{2}\gamma(1-\gamma)\sigma^2 \tag{A.30}$$

Decompose $A_1 = U\omega U^{-1}$, where $\omega \equiv diag[\omega_1, ..., \omega_n]$, with $n = 2$ in alternative 1 and $n = 4$ in alternative 2. From lemma 1,

$$E_t y_s = U e^{\omega(s-t)} U^{-1} \text{ for } s \geq t \tag{A.31}$$

hence,

$$\begin{aligned}
P_t &= E_t(\int_t^\infty \frac{\Lambda_s}{\Lambda_t} D_s ds) \\
&= \begin{Bmatrix} C_t^\gamma G_t^{-1} E_t[\int_t^\infty e^{-\rho(s-t)} C_s^{-\gamma} G_s D_s ds] \text{ in Proposition 2} \\ C_t^\gamma G_t^{-1} E_t \int_t^\infty e^{-\rho(s-t)} C_s^{1-\gamma} G_s F_s ds \text{ in Proposition 3} \end{Bmatrix} \\
&= C_t^\gamma G_t^{-1} \int_t^\infty E_t[e^{-\rho(s-t)} y_s^1] ds \\
&= C_t^\gamma G_t^{-1} \int_t^\infty e_1' E_t[e^{-\rho(s-t)} y_s] ds \\
&= C_t^\gamma G_t^{-1} \left(\int_t^\infty e_1' U e^{(\omega-\rho)(s-t)} U^{-1} ds \right) y_{1t} \\
&= C_t^\gamma G_t^{-1} \left(e_1' U \frac{1}{\rho - \omega} U^{-1} \right) y_{1t} \\
&= C_t^\gamma G_t^{-1} \sum_{i=1}^n a_i y_t^i \\
&= \begin{Bmatrix} a_1 D_t + a_2 \frac{D_t}{G_t} \text{ in Proposition 2} \\ a_1 D_t + a_2 \frac{D_t}{G_t} + a_3 C_t + a_4 \frac{C_t}{G_t} \text{ in Proposition 3} \end{Bmatrix}
\end{aligned} \tag{A.32}$$

where $a_i \equiv e_1' U \frac{1}{\rho-\omega} U^{-1} e_i$; e_i $(i = 1, ...n)$ is an n by 1 vector with 1 in the ith entry and zeros in all the other entries; and I've used respectively the Fubini theorem, (A.31), and the definitions of y_t^i for the third, the fifth and the last equality in (A.32).

(A.32) gives the closed-form formula for the stock price in both alternative 1 and alternative 2. Using the similar argument as that in the proof of Propostion 1,

$$\sum_{i=1}^2 a_i (I\rho - A_1)_{(i,:)} = [1, 0] \tag{A.33}$$

in alternative 1, where $(I\rho - A_1)_{(i,:)}$ represents the ith row of $I\rho - A_1$ $(i = 1, 2)$; and

$$\sum_{i=1}^{4} a_i (I\rho - A_1)_{(i,:)} = [1, 0, 0, 0] \tag{A.34}$$

in alternative 2, where $(I\rho - A_1)_{(i,:)}$ represents the ith row of $I\rho - A_1$ $(i = 1, 2, 3, 4)$. Solving (A.33) and (A.34) yields:

$$\begin{aligned} a_1 &= \tfrac{1}{\rho - A_{1,11}} \\ a_2 &= \tfrac{A_{1,12}}{(\rho - A_{1,22})(\rho - A_{1,11})} \end{aligned} \tag{A.35}$$

for alternative 1, where $A_{1,11}$, $A_{1,12}$ and $A_{1,22}$ are given by (A.25)-(A.27); and

$$\begin{aligned} a_1 &= \tfrac{1}{\rho - A_{1,11}} \\ a_2 &= \tfrac{A_{1,12}}{(\rho - A_{1,22})(\rho - A_{1,11})} \\ a_3 &= \tfrac{\phi \hat{F}}{(\rho - A_{1,33})(\rho - A_{1,11})} \\ a_4 &= \tfrac{\phi \hat{F}}{(\rho - A_{1,44})(\rho - A_{1,11})} \left(\tfrac{A_{1,12}}{\rho - A_{1,22}} + \tfrac{A_{1,34}}{\rho - A_{1,33}} \right) \end{aligned} \tag{A.36}$$

for alternative 2, where $A_{1,11}$, $A_{1,12}$, $A_{1,22}$, $A_{1,33}$, $A_{1,34}$ and $A_{1,44}$ are given by (A.28)-(A.30). This finishes the proofs of parts a) in Proposition 2 and 3.

To prove parts b), again notice that at the presence of jumps,

$$\frac{d\Lambda_t}{\Lambda_t} = \frac{d\Lambda_t^*}{\Lambda_t} + J_{\Lambda t} dN_t = \mu_{\Lambda t} dt - \sigma_{\Lambda t} dB_{1t} + J_{\Lambda t} dN_t \tag{A.37}$$

$$\frac{dP_t}{P_t} = \frac{dP_t^*}{P_t} + J_{Pt} dN = \mu_{Pt} dt + (\sigma_{Pt})_1 dB_{1t} + (\sigma_{Pt})_2 dB_{2t} + J_{Pt} dN_t \tag{A.38}$$

where $\frac{dX_t^*}{X_t}$ denotes the drift and the diffusion components in the dynamics of $\frac{dX_t}{X_t}$; $J_{Xt} \equiv \frac{X_t^+}{X_t} - 1$ denotes the jump size of X_t, where $X = \Lambda, P$. Unlike that in the baseline model, the implied stock return processes, $\frac{dP_t}{P_t}$, in both alternative models are driven by two Brownians, B_{1t} and B_{2t}, with the correlation of ρ_1, hence

$$\begin{aligned} (volR_t)^2 \cdot dt &= (\sigma_{Pt})_1 dB_{1t} + (\sigma_{Pt})_2 dB_{2t}^2 + (J_{Pt})^2 \lambda dt \\ &= (\sigma_{Pt})_1^2 + (\sigma_{Pt})_2^2 + 2\rho_1 (\sigma_{Pt})_1 (\sigma_{Pt})_2 + \lambda J_{Pt}^2 \end{aligned}$$

and

$$EP_t dt \equiv -E_t \left(\frac{dP_t^*}{P_t} \frac{d\Lambda_t^*}{\Lambda_t} \right) - J_P J_\Lambda \lambda dt = \sigma_{\Lambda t} \left[(\sigma_{Pt})_1 + \rho_1 (\sigma_{Pt})_2 \right] dt - J_P J_\Lambda \lambda dt$$

which yields expressions in (5.4)-(5.5) for both alternative 1 and alternative 2. The differently assumed dividend processes do not affect the pricing kernel, hence the risk-free rate is still given by

$$rf_t dt = -E_t\left(\frac{d\Lambda_t}{\Lambda_t}\right) = -\mu_{\Lambda t}dt - J_{\Lambda t}\lambda dt$$

for both alternative models. as well as for the baseline model.

To obtain the formulas of μ_Λ, σ_Λ, J_Λ, σ_{P_i} $(i = 1, 2)$ and J_P in alternative 1, I apply Ito's lemma with jumps to the expressions of Λ_t and P_t in (1.8) and (5.2),

$$\mu_{\Lambda t} = -\rho + \mu_1 + k\frac{\bar{G} - G_t}{G_t} - \alpha b\lambda\frac{G_t - \beta}{G_t} + \gamma\alpha\sigma^2\frac{G_t - \beta}{G_t} \tag{A.39}$$

$$\sigma_{\Lambda t} = \sigma\left(\gamma + \alpha\frac{G_t - \beta}{G_t}\right) \tag{A.40}$$

$$J_{\Lambda t} \equiv e^{b\gamma}\left(1 + \alpha b\frac{G_t - \beta}{G_t}\right) - 1 \tag{A.41}$$

$$(\sigma_{Pt})_1 = \sigma a_2\alpha\frac{D_t}{P_t}\frac{G_t - \beta}{G_t^2}; \tag{A.42}$$

$$(\sigma_{Pt})_2 = \sigma_D \tag{A.43}$$

$$J_{Pt} = \frac{a_2}{a_1 G_t + a_2}\left(\frac{1}{1 + \alpha b\frac{G_t - \beta}{G_t}} - 1\right) \tag{A.44}$$

where I've used the assumed processes of G_t, C_t and D_t in (1.5), (1.7) and (5.1). Similarly, applying Ito's lemma with jumps to the expressions of P_t in (5.8) yields

$$(\sigma_{Pt})_1 = \sigma\left[a_1\frac{D_t}{P_t} + a_3\frac{C_t}{P_t} + \left(1 + \alpha\frac{G_t - \beta}{G_t}\right)\left(a_2\frac{1}{G_t}\frac{D_t}{P_t} + a_4\frac{1}{G_t}\frac{C_t}{P_t}\right)\right] \tag{A.45}$$

$$(\sigma_{Pt})_2 = \sigma_F\left(a_1 + a_2\frac{1}{G_t}\right)\frac{D_t}{P_t} \tag{A.46}$$

$$\begin{aligned} J_{Pt} \quad = \quad & a_1\frac{D_t}{P_t}(e^{-b} - 1) + a_2\frac{1}{G_t}\frac{D_t}{P_t}\left(\frac{e^{-b}}{1 + \alpha b\frac{G_t - \beta}{G_t}} - 1\right) \\ & + a_3\frac{C_t}{P_t}(e^{-b} - 1) + a_4\frac{1}{G_t}\frac{C_t}{P_t}\left(\frac{e^{-b}}{1 + \alpha b\frac{G_t - \beta}{G_t}} - 1\right) \end{aligned} \tag{A.47}$$

for alternative 2, where I've used the assumed processes of G_t, C_t and F_t in (1.5), (1.7) and (5.6). Note that the formulas of μ_Λ, σ_Λ and J_Λ in alternative 2 remain the same as those in alternative 1, which are given in (A.39)-(A.41), because the different dividend processes do not affect the evolution of the pricing kernel. ■

References

[1] Ait-Sahalio, Yacine., and Andrew W. Lo, 1998, Nonparametric estimation of state-price densities implicit in financial asset prices, *Journal of Finance* 53(2), 499-548.

[2] Bakshi, Gurdip S., Charles Cao, and Zhiwu Chen, 1997, Empirical performance of alternative option pricing models, *Journal of Finance* 52(5), 2003-2051.

[3] Bakshi, Gurdip S., Charles Cao, and Zhiwu Chen, 2000, Pricing and hedging long-term options, *Journal of Econometrics* 94, 277-318.

[4] Bansal, Ravi, and Amir Yaron, 2004, Risks for the long run: a potential resolution of asset pricing puzzles, *Journal of Finance* 59, 1481-1509.

[5] Barro, Robert J., 2006, Rare disasters and asset markets in the twentieth century, *Quarterly Journal of Economics* 121(3), 823-866.

[6] Bates, David S., 1991, The crash of '87, was it expected? the evidence from option markets, *Journal of Finance* 46(3), 1009-1044.

[7] Bates, David S., 2000, Post -'87 crash fears in S&P 500 future options, *Journal of Econometrics* 94, 181-238.

[8] Benzoni, Luca, Pierre Collin-Dufresne, and Robert S. Goldstein, 2005, Can standard preference explain the prices of out of the money S&P 500 put options? NBER working paper, w11861.

[9] Buraschi, Andrea and Alexei Jiltsov, 2006, Model uncertainty and option markets with heterogeneous beliefs, *Journal of Finance* 61(6), 2848-2898.

[10] Buraschi, Andrea and Alexei Jiltsov, 2007, Habit formation and the term structure of interest rates, forthcoming in *Journal of Finance*.

[11] Campbell, John, and John Cochrane, 1999, By force of habit: a consumer-based explanation of aggregate stock market behavior, *Journal of Political Economy* 107, 205-251.

[12] Chen, Long, Pierre Collin-Dufresne, and Robert S. Goldstein, 2006, On the relation between credit spread puzzles and the equity premium puzzle, working paper, Michigan State University, University of California at Berkeley, and University of Minnesota.

[13] Cochrane, John, 2001, *Asset pricing*, Princeton University Press.

[14] Cochrane, John, 2005, Financial markets and the real economy, NBER working paper, w11193.

[15] Constantinides, George, 1990, Habit formation: a resolution of the equity premium puzzle, *Journal of Political Economy* 98(13), 519-543.

[16] David, Alexander, and Pietro Veronesi, 2002, Option prices with Uncertain fundamentals, working paper, Washington University in St. Louis and University of Chicago.

[17] Du, Du, 2007, Asset pricing for augmented affine jump-diffusions with regime shifts, working paper, Hong Kong University of Science and Technology.

[18] Duffie, Darrell, 2001, *Dynamic asset pricing theory*, Princeton University Press.

[19] Duffie, Darrell, Jun Pan, and Kenneth Singleton, 2000, Transform analysis and asset pricing for affine crash-diffusions, *Econometrica* 68(6), 1343-1376.

[20] Duffie, Darrell., and Kenneth Singleton, 1993, Simulated moments estimation of Markov models of asset prices, *Econometrica* 61(4), 929-952.

[21] Eraker, Bjorn, 2004, Do stock prices and volatility crash? Reconciling the evidence from spot and option prices, *Journal of Finance* 59, 1367-1404.

[22] Fama, Eugene F. and Kenneth R. French, 1988, Dividend yields and expected stock returns, *Journal of Financial economics* 96, 246-273.

[23] Gabaix, Xavier, 2007, A simple, unified, xxactly solved framework for ten puzzles in Macro-Finance, working paper, M.I.T.

[24] Hansen, Lars P., John Heaton, and Nan Li, 2004, Intangible risk, working paper, University of Chicago.

[25] Hansen, Lars P., and Kenneth Singleton, 1983, Stochastic Consumption, Risk-Aversion, and the Temporal Behavior of Asset Returns, *Journal of Political Economy* 91, 249-265 1983.

[26] Hanson, Floyd B., 2005, *Applied stochastic processes and control for crash-diffusions: models, analysis and computation*, SIAM Books.

[27] Heaton, John, 1993, The interaction between time-nonseparable preferences and time aggregation, *Econometrica* 61(2), 353-385

[28] Karatzas, Ioannis, and Steven E. Shreve, 2000, *Brownian Motion and Stochastic Calculus*, Springer.

[29] Lee, Bong-Soo, and Beth F. Ingram, 1991, Simulation Estimation of Time Series Models, *Journal of Econometrics* 47, 197-205.

[30] Liu, Jun, Jun Pan, and Francis A. Longstaff, 2003, Dynamic Asset Allocation with Event Risk, *Journal of Finance* 58(1), 231-259.

[31] Liu, Jun, Jun Pan, and Tan Wang, 2005, An equilibrium model of rare-event premia and its implication for option smirks, *Review of Financial Studies* 8(1), 131-164.

[32] Lonstaff, Francis A., and Monika Piazzesi, 2004, Corporate earnings and the equity premium, *Journal of Financial Economics* 74(3), 1-21.

[33] Mehra, Rajnish and Edward C. Prescott, 1985, The equity premium:a puzzle, *Journal of Monetary Economics* 15, 145-161.

[34] Menzly Lior, Tano Santos and Pietro Veronesi, 2004, Understanding predictability, *Journal of Political Economy* 112(1), 1-47.

[35] Pan, Jun, 2002, The crash-risk premia implicit in options: evidence from an integrated time-series study, *Journal of Financial economics* 63, 3-50

[36] Rietz, Thomas, 1988, The equity premium puzzle: a solution? *Journal of Monetary economics* 21, 117-132

[37] Rubinstein, M., 1994, Implied binomial trees, *Journal of Finance* 49, 771-818.

[38] Santos, Tano, and Pietro Veronesi, 2006, Labor income and predictable stock returns, *Review of Financial Studies* 19, 1-44.

[39] Santos, Tano, and Pietro Veronesi, 2005, Cash-flow risk, discount risk, and the value premium, working paper, Columbia University and University of Chicago.

[40] Shiller, Robert J., 1981, Do stock prices move too much to be justified by subsequent changes in dividends? *American Economic Review* 71, 421-436.

[41] Veronesi, Pietro, 2000, How does information quality affect stock returns? *Journal of Finance* 55(2) 807-837.

[42] Veronesi, Pietro, 2004, The Peso Problem Hypothesis and Stock Market Returns, *Journal of Economic Dynamics and Control*, 28(4), 707-725.

[43] Watcher, Jessica, 2005, Solving models with external habit, *Finance Research Letters* 2, 210-226

[44] Watcher, Jessica, 2006, A consumption-based model of the term structure of interest rate, *Journal of Financial economics* 79, 365-399

[45] Weil, Philippe, 1989, The equity premium puzzle and the risk-free rate puzzle, *Journal of Monetary Economics* 24, 401-421

[46] Working, Holbrook, 1960, Note on the Correlation of First Differences of Averages in a Random Chain, *Econometrica* 28, 916-918.

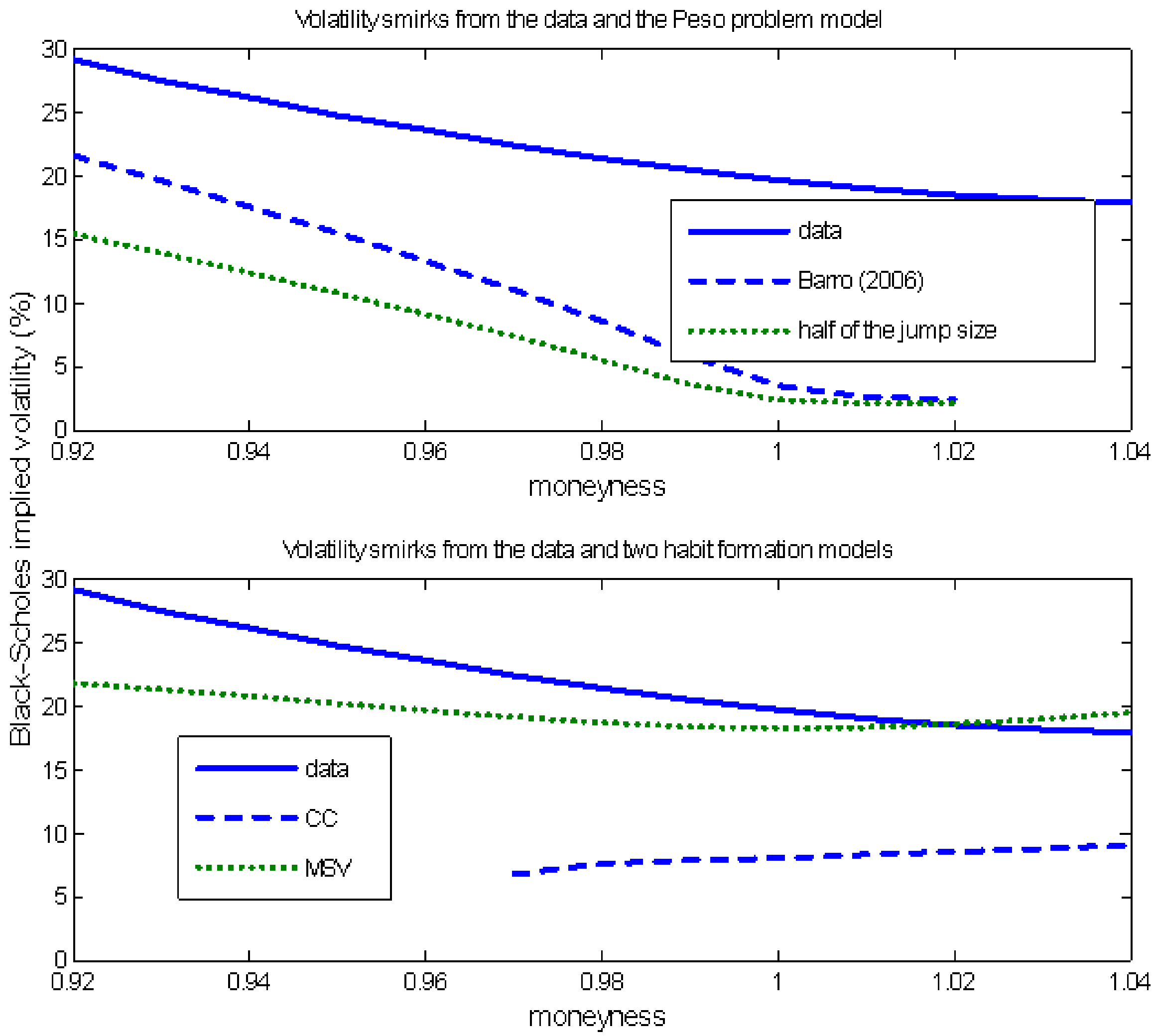


Figure 1. Volatility smirks implied from the data, the Peso problem model and the habit formation models. Figure 1 plots the average prices of options with 30 days to expiration, quoted in terms of the implied Black-Scholes volatility (B/S-vol), against the options' moneyness, i.e., the average volatility smirks. The upper panel plots together the observed volatility smirk computed using the S&P 500 index option data in the period from April 4, 1988 to June 30, 2005, and its counterpart implied from Barro (2006) in which a potential 37% consumption contraction striking once every 60 years is assumed. I also plot in the same Panel the implied volatility smirk from Barro's model with the consumption jump size cut in half and all other parameter values remaining unchanged. The lower Panel plots the volatility smirk implied from the same data together with its counterparts implied from two habit formation models: Campbell and Cochrane (1999, CC), and Menzly, Santos and Veronesi (2004, MSV), where the model prices are computed at the parameter values reported in these two papers.

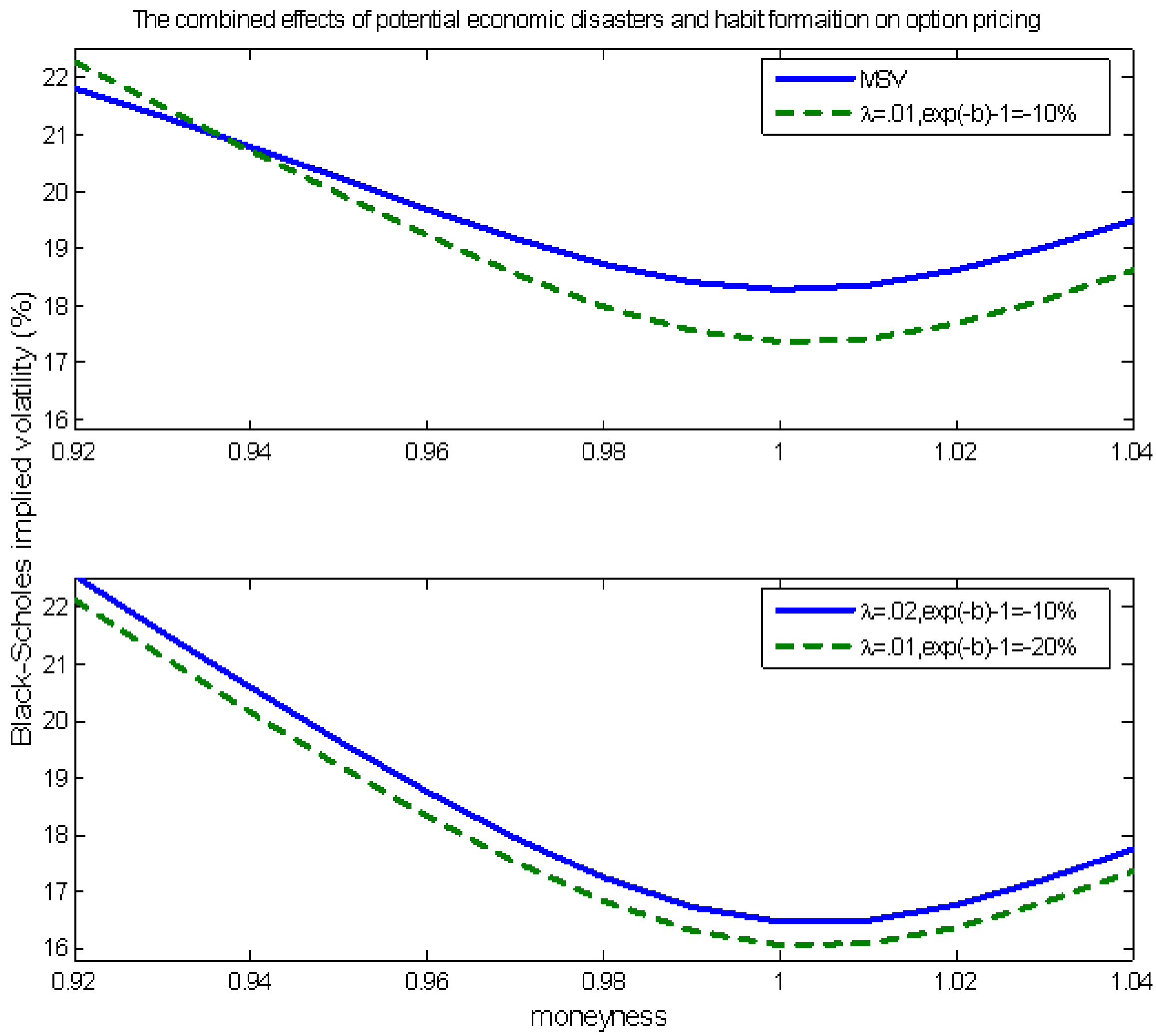


Figure 2. The combined effects of Peso problem situation and habit formation on option pricing. Figure 2 plots the average prices of options with 30 days to expiration, quoted in terms of the implied Black-Scholes volatility (B/S-vol), against the options' moneyness, i.e., the unconditional volatility smirks. The upper Panel plots together the volatility smirk implied from the habit formation model by Menzly, Santos and Veronesi (2005, MSV) and my baseline model incorporating both potential economic disasters and MSV habit specification, where the non-jump parameters are from Table 1 of MSV, and the assumed potential disaster strikes once every 100 years in the form of a 10% consumption contraction, i.e., $\lambda = .01$, $e^{-b} - 1 = -10\%$. The lower Panel plots the volatility smirks implied from my baseline model under another two jump scenarios: i) $\lambda = .02$, $e^{-b} - 1 = -10\%$; and ii) $\lambda = .01$, $e^{-b} - 1 = -20\%$.

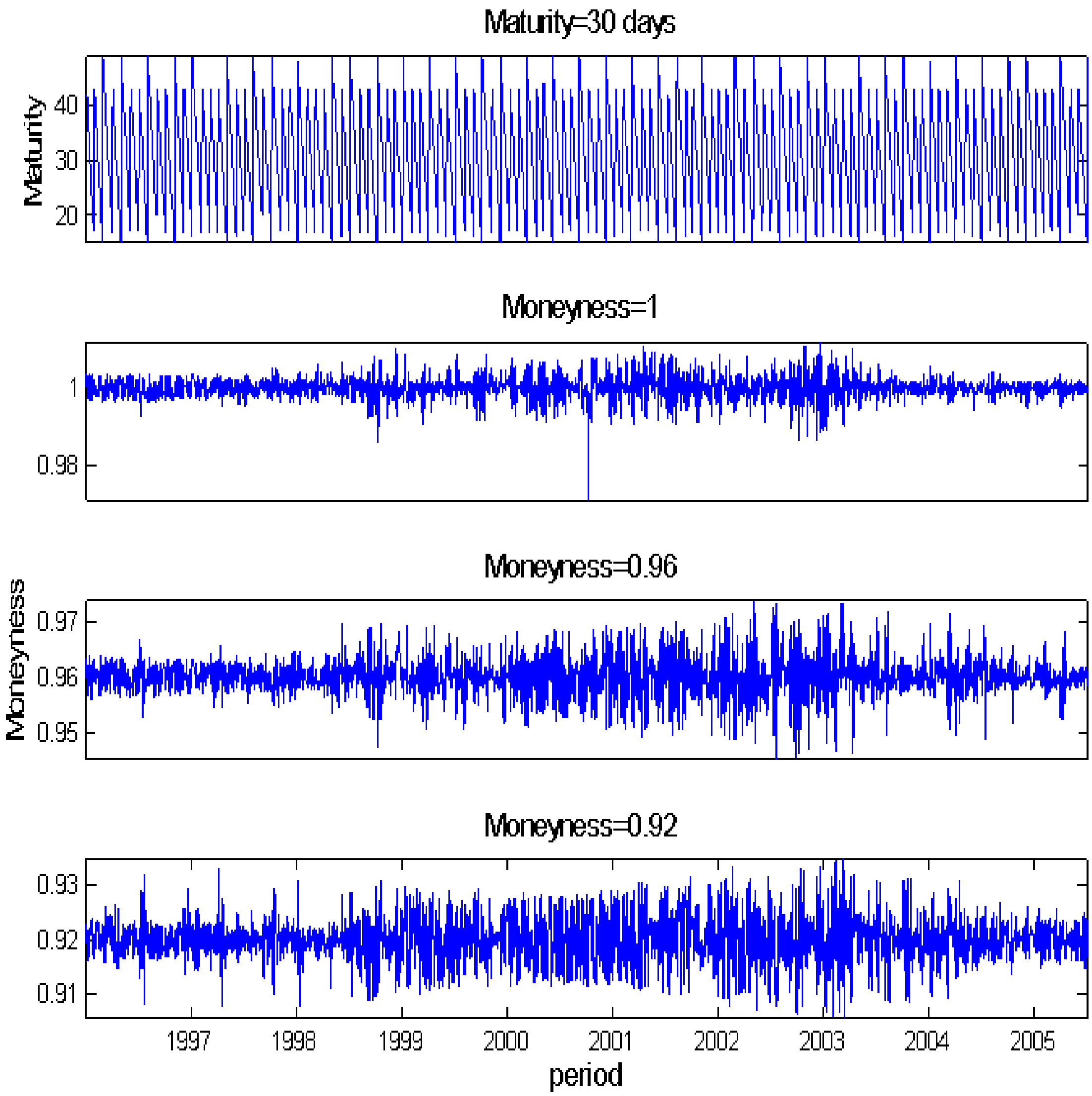
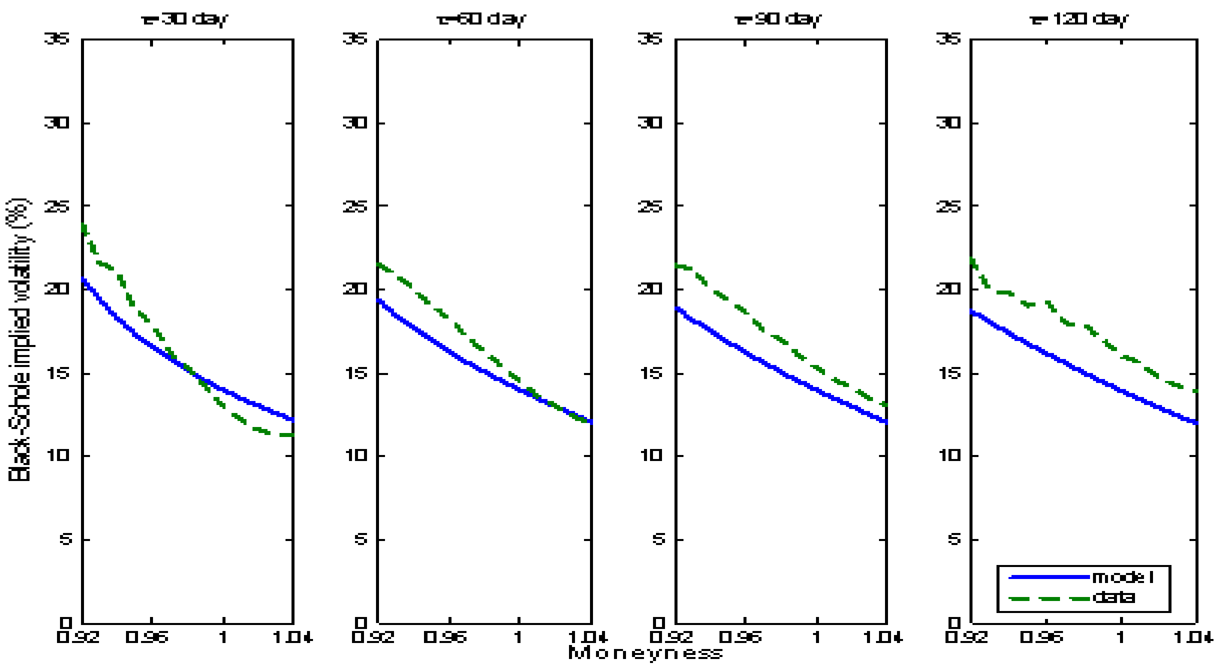
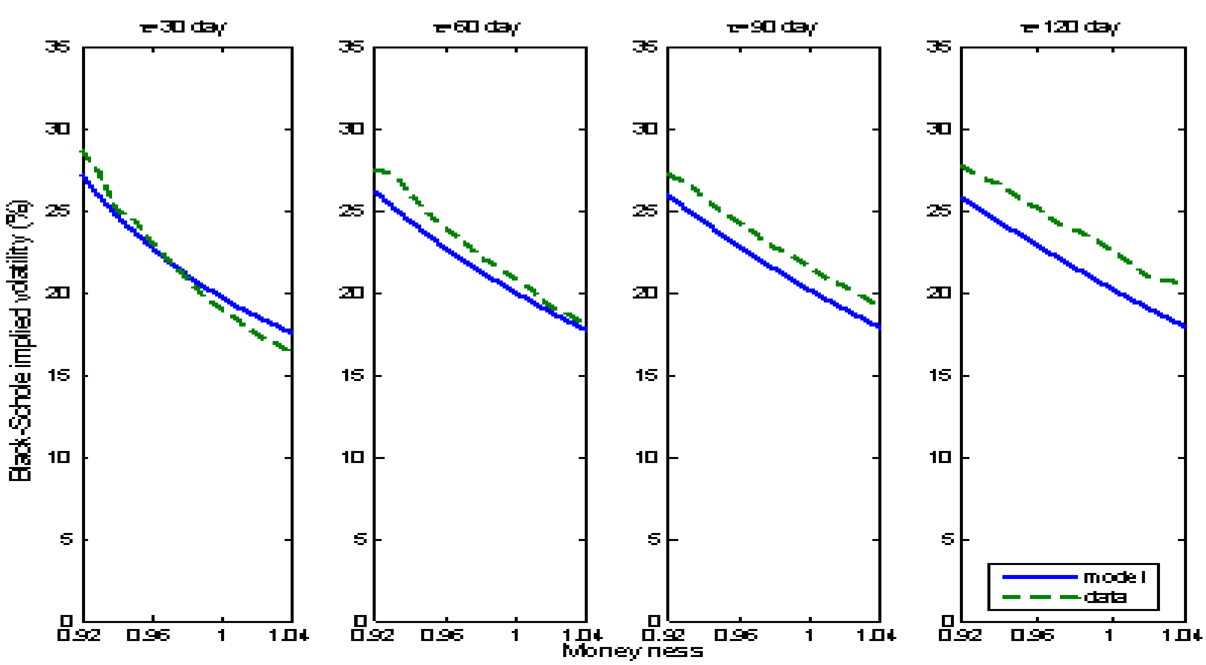


Figure 3. Time Series of the maturity and moneyness of options used in the estimations. Figure 3 plots the time series of the maturity and selected degrees of moneyness of options that are used for the model estimations in the period from January 4, 1996 to June 30, 2005. Options are all short term ones with 30 days to expiration, and their moneyness is between 0.92 and 1. I use the following policy to select the option data: among all available options on the nth trading day, I first select those with maturities closest to 30 days. From the pool of options thus chosen, I then select the one with moneyness closest to each of the elements in $\bar{M} \equiv [.92, .93, .94, .95, .96, .97, .98, .99, 1]'$.

a: Low volatility days



b: Medium volatility days



c: High volatility days

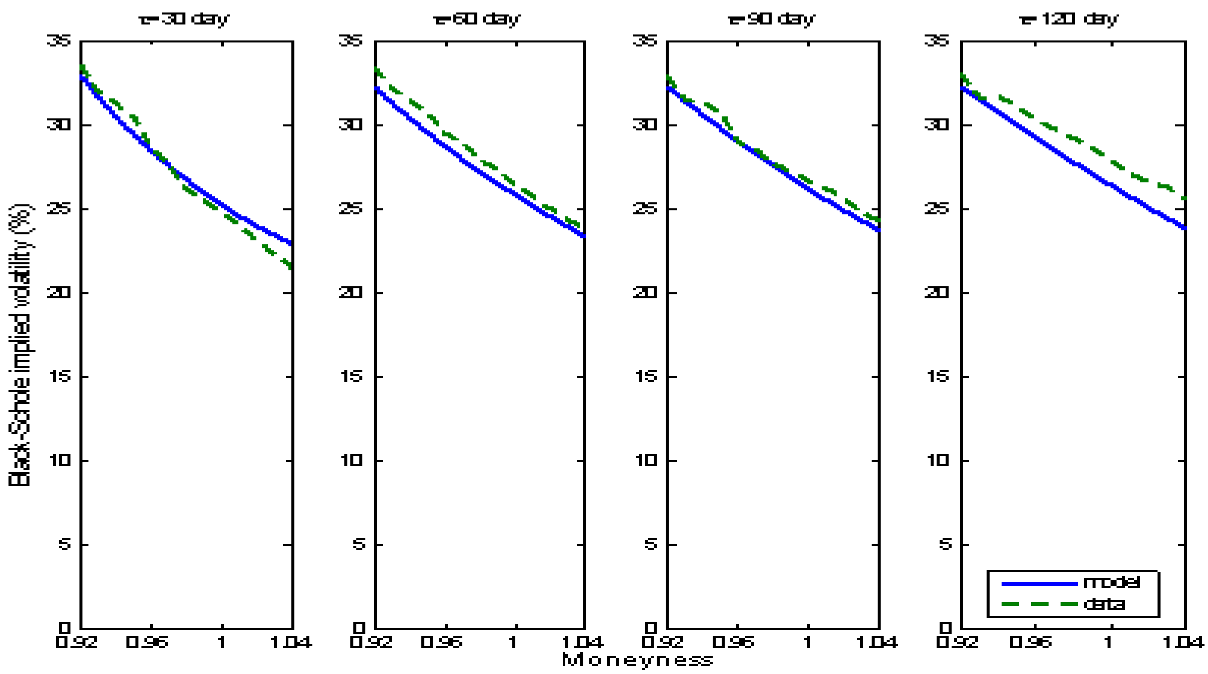


Figure 4. Conditional volatility smirks. Figure 4 plots together volatility smirks implied from both the (baseline) model and the data during the three volatility deciles for options with maturities of 30 days, 60 days, 90 days and 120 days. I first back out the daily return volatility during my option sample period using the short term option data on the same day with the parameter estimates as the inputs. I then sort the total of 4,347 trading days into ten deciles according to the backed-out daily return volatilities, where days in the higher deciles experience higher volatilities. The data prices are computed as the average prices of options selected from days falling into a particular volatility decile whose moneyness and maturities fall into the ranges of $[M - .005, M + .005]$ and $[\tau - 10, \tau + 10]$, where M and τ are the target values of moneyness and maturity, respectively. The model prices in each decile are simulated with the underlying state starting from the mid-point value of the corresponding volatility ranges. The analysis focuses on the sixth, the ninth, and the second deciles representing respectively medium volatility days, high volatility days, and low volatility days.

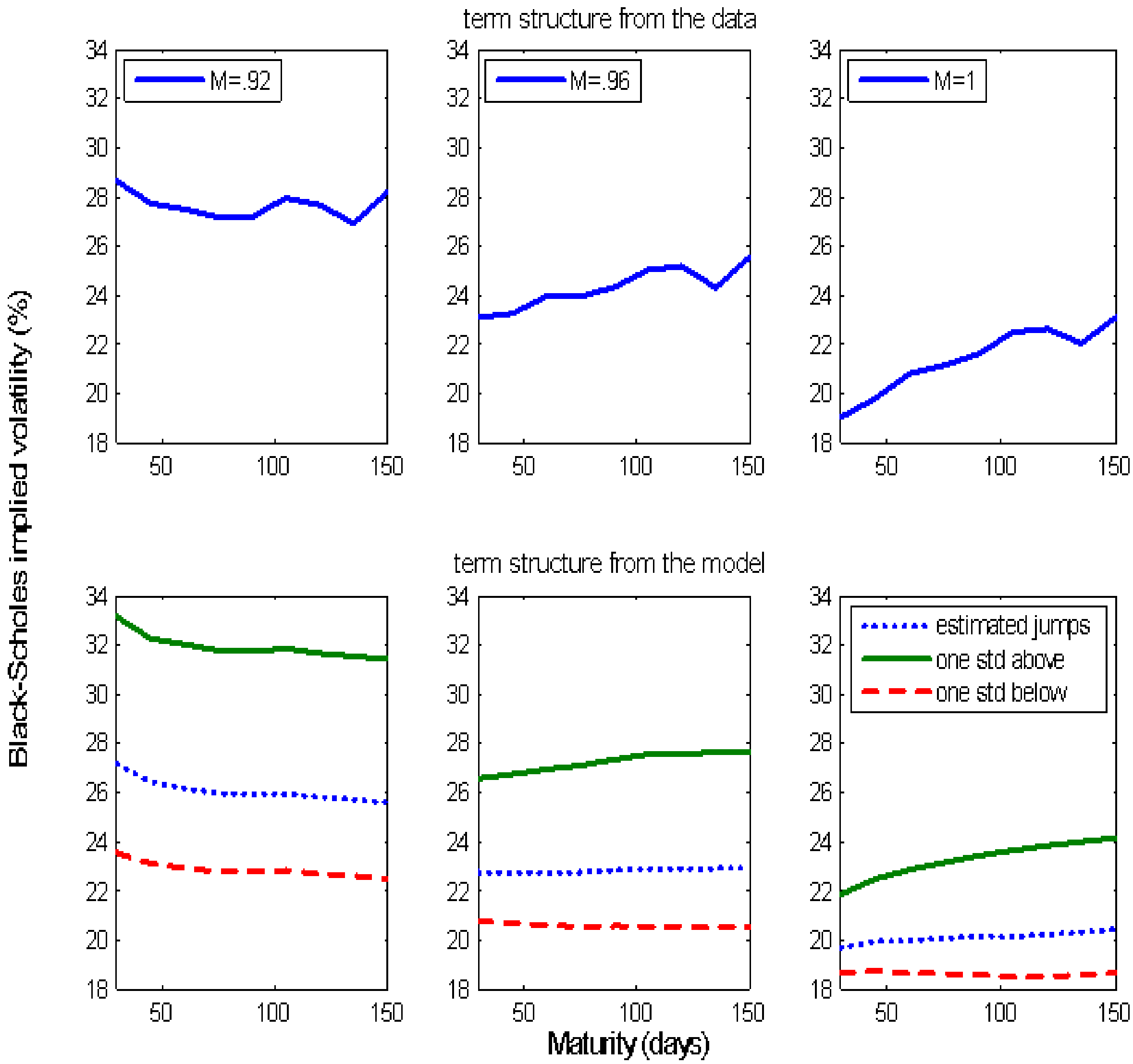


Figure 5. Term structure in option pricing. Figure 5 plots option prices against their maturities, i.e., the term structure in option pricing, during the medium volatility days under three moneyness scenarios: deep OTM put options with moneyness of 0.92, lesser OTM put options with moneyness of 0.96, and ATM options with moneyness of 1. The three upper panels plot the variations of term structure in option pricing implied from the data. The three lower panels plot their counterparts implied from the baseline model under three jump scenarios: i) the estimated jumps at $(\hat{\lambda}, \hat{b})$, ii) jumps one standard error above their estimates at $(\hat{\lambda} + std(\hat{\lambda}), \hat{b} + std(\hat{b}))$, and iii) jumps one standard error below their estimates at $(\hat{\lambda} - std(\hat{\lambda}),\ \hat{b} - std(\hat{b}))$. All other parameters remain at their estimated values

Table I.
Summary statistics of the option data

Table I reports the summary statistics of the option data collected from two datasets, CBOE and Ivy DB, under various maturity and moneyness categories. CBOE data cover the period from April 4, 1988 to December 29, 1995, and Ivy DB data cover the period from January 4, 1996 to June 30th, 2005. Both the average option prices, quoted in term of Black-Scholes implied volatilities (B/S-vols), and the number of option contracts (in parentheses) are reported for each of the option categories.

Panel A: CBOE data				
Moneyness	Days-to-Expiration			
	15-45	45-90	90-150	subtotal
.92-.94	0.284 (869)	0.267 (415)	0.271 (246)	(1530)
.94-.96	0.242 (1493)	0.239 (713)	0.251 (351)	(2557)
.96-.99	0.196 (3856)	0.201 (2169)	0.222 (816)	(6841)
.99-1.01	0.164 (3133)	0.171 (2805)	0.196 (1110)	(7048)
1.01-1.04	0.147 (3905)	0.156 (3161)	0.189 (1423)	(8489)
subtotal	(13256)	(9263)	(3946)	(26465)

Panel B: Ivy DB data				
Moneyness	Days-to-Expiration			
	15-45	45-90	90-150	subtotal
.92-.94	0.279 (5046)	0.272 (4679)	0.269 (2263)	(11988)
.94-.96	0.253 (5696)	0.255 (5061)	0.254 (2314)	(13071)
.96-.99	0.225 (10151)	0.234 (8927)	0.239 (3534)	(22612)
.99-1.01	0.203 (7296)	0.215 (6725)	0.225 (2466)	(16487)
1.01-1.04	0.187 (10242)	0.200 (9167)	0.213 (3567)	(22976)
subtotal	(38431)	(34559)	(14144)	(87134)

Table II.
Estimation results with the baseline model

Table II reports the estimation results with the baseline model treating consumption and dividend as a single process. The estimation is implemented using the method of simulated moments (MSM) in two steps. In the first step, I use the monthly consumption data to estimate the two non-jump parameters associated with the consumption process, i.e., μ, σ. Given the estimated θ_1, in the second step I use the joint daily data of options, stocks and risk-free rates to simultaneously estimate the remaining parameters associated with the potential disasters, $\{\lambda, b\}$, and the assumed preference, $\{\gamma, \rho, \bar{G}, k, \beta, \alpha\}$. Panel A reports the estimated non-jump parameters; Panel B reports the derived $S_{\max}$ and mean(η_t), denoting respectively the maximum value of the simualted surplus consumption ratio and the average local curvature of the utility function; Panel C reports the estimated potential disasters, where λ, $e^{-b}-1$ and mean(J_P) denote respectively the arrival intensity of the economic disaster, the consumption jump size and the average jump size in the stock price; Panel D reports the goodness-of-fit tests including four individual tests evaluating the differences between the model values and their data counterparts for the average prices of 8% OTM put options ($O|_{M=.92}$) and ATM options ($O|_{M=1}$), the smirk premium (SP) measured as the average price differential between 8% OTM puts and ATMs, and the average equity premium (EP), together with the J_T statistics evaluating the overal fit of the model. All standard errors are in the parenthese.

Panel A: non-jump parameters

μ	σ	γ	ρ	$\bar{G}$	k	β	α
.0205	.0182	1.21	0.0147	32.8	0.339	26.3	55.3
(.0025)	(.0009)	(0.172)	(0.0032)	(57.8)	(0.128)	(46.9)	(5.11)

Panel B: derived habit formation

	$S_{\max}$	mean(η_t)
my model	0.0671	21.7
CC	0.097	40.8
MSV	0.048	35

Panel C: potential disasters

	λ	b	$e^{-b}-1$	mean(J_P)
my model	0.0198	0.194	-0.176	-0.56
	(0.0139)	(0.0997)		
Barro (2006)	0.017		-0.37	-0.37

Panel D: goodness-of-fit tests				
$O\|_{M=.92}$ (%)	$O\|_{M=1}$ (%)	SP (%)	EP (%)	J_T stat [p-value]
-7.02	-6.88	-0.15	-0.62	469 [<.01]
(0.23)	(0.23)	(2.05)	(5.78)	

Table III.
Estimation results with the two alternative models

Table III reports the estimation results with two alternative models with different dividend processes. As in the baseline model, I apply the method of simulated moments (MSM) to implement the estimation in two steps. In the first step, I use the monthly consumption and dividend data to estimate the non-jump parameters associated with the consumption and the dividend processes and the results are reported in Panel A. Given the first step estimates, in the second step I use the same asset pricing moments as those for the baseline model to estimate the remaining parameters associated preferences (reported in Panel B) and the potential disasters (reported in Panel C), where λ, $e^{-b}-1$ and mean(J_P) denote the arrival intensity of the economic disaster, the consumption jump size, and the average jump size in the stock price, respectively. Panel D reports the goodness-of-fit tests including four individual tests evaluating the differences between the model values and their data counterparts for the average prices of 8% OTM put options ($O|_{M=.92}$) and ATM options ($O|_{M=1}$), the smirk premium (SP) measured as the average price differential between 8% OTM puts and ATMs, and the average equity premium (EP), together with the J_T statistics evaluating the overal fit of the model. All standard errors are in the parenthese.

Panel A: non-jump parameters associated with the consumption and dividend processes

	μ	σ	ρ_1	μ_D	σ_D	ϕ	$\bar{F}$	σ_F
alternative 1	0.0204	0.0183	0.07	0.0204	0.0664			
	(0.0025)	(0.0009)	(0.042)	(0.0089)	(0.01)			
alternative 2	0.0205	0.0183	-0.219			0.179	0.0254	0.0675
	(0.0025)	(0.0009)	(0.057)			(0.064)	(0.0004)	(0.01)

Panel B: parameters associated with preference

	γ	ρ	$\bar{G}$	k	β	α
alternative 1	0.931	0.0095	27.2	0.0087	5.79	40.3
	(0.25)	(0.0058)	(195)	(0.0108)	(170)	(2.22)
alternative 2	1.42	0.0123	97.1	0.302	76.0	51.1
	(0.898)	(0.0123)	(144)	(0.268)	(110)	(17.3)

Panel C: potential disasters

	λ	b	$e^{-b}-1$	mean(J_P)
alternative 1	0.0156	0.146	-0.135	-0.36
	(0.0126)	(0.116)		
alternative 2	0.0277	0.162	-0.15	-0.51
	(0.0059)	(0.0368)		

Panel D: goodness-of-fit tests					
	$O\|_{M=.92}$ (%)	$O\|_{M=1}$ (%)	SP (%)	EP (%)	J_T stat [p-value]
alternative 1	-6.5	-4.28	-2.22	-.62	912 [<.01]
	(0.24)	(0.25)	(2.33)	(5.61)	
alternative 2	-7.20	-7.31	-0.23	-0.52	474 [<.01]
	(0.18)	(0.20)	(1.78)	(5.77)	

Table IV.
Tests of the price matches for the out-of-the-sample options

Table IV reports the results of the tests that the model prices are equal to their data counterparts for the out-of-the-sample option data. I first sort a total of 4347 trading days into ten deciles according to the backed-out return volatilities using the short term option data, where days in the higher deciles experience higher volatilities. I select the sixth, the ninth, and the second deciles representing medium, high, and low volatility days, respectively, and the corresponding volatility ranges are reported in Panel A. I further divide options falling into each of the volatility days into eight classes according to their moneyness and maturities, hence a total of 24 option classes. I report in Panel B for each of the option classes the p-values of the χ^2 statistics computed as $c_l = [\frac{1}{I_l}\sum_{i=1}^{I_l} \varepsilon_l]' \cdot \Omega^{-1}[\frac{1}{I_l}\sum_{i=1}^{I_l} \varepsilon_l]$, where ε_l denotes the deviation of the model price from its data counterpart divided by the data price; Ω_l is computed using 20 lags' Newey-West for all $l = 1, 2...., 24$. Under the null hypothesis of zero pricing errors, c_l is asymptotically distributed according to $\chi^2(1)$ for each l. The numbers of option data in each of the classes are reported in the parentheses

Panel A: range of the return volatilities (%)			
	medium volatility	high volatility	low volatility
range	20.2-21.3	24.2-26.4	15.7-16.6

Panel B: Chi-square tests of the matches for the out-of-the-sample options					
	Maturity	Moneyness			
	(days)	$.92 \sim .94$	$.94 \sim .98$	$.98 \sim 1$	$1 \sim 1.04$
low	$45 \sim 75$	0.694 (247)	0.702 (711)	0.708 (672)	0.970 (1384)
volatility	$75 \sim 150$	0.694 (254)	0.699 (674)	0.698 (474)	0.705 (1054)
medium	$45 \sim 75$	0.715 (430)	0.734 (1094)	0.738 (696)	0.822 (1228)
volatility	$75 \sim 150$	0.735 (419)	0.733 (942)	0.715 (559)	0.725 (1117)
high	$45 \sim 75$	0.792 (374)	0.785 (924)	0.839 (578)	0.943 (1114)
volatility	$75 \sim 150$	0.892 (486)	0.866 (968)	0.830 (586)	0.784 (1265)

Table V.
Matches of the historical smirk premia

Table V reports the tests that fluctuations of historical smirk premia are driven by the historical consumption shocks. I feed the baseline model with the historical consumption shocks to back out the implied monthly series of $\{G_t^m\}$, where G is the equivalent state variable defined in (1.5). The observed option prices are collected in the middle of each month, and their model counterparts are simulated with the underlying state starting from the mid-point value of the corresponding volatility ranges. Denote by $\{smirk_{i,t}^{data}\}$ and $\{smirk_{i,t}^{m}\}$ the series of the smirk premia observed in the data and simulated from the model, respectively, where I consider three measures of the smirk premia demanded by 8% OTM put options (the usual measure), 4% OTM put options, and 4% OTM call options (with moneyness=1.04) relative to the prices of ATM options. Results in columns 2-4 are based on the regressions, $smirk_{i,t}^{data} = \beta_1 + \beta_2 smirk_{i,t}^{m} + \epsilon_{i,t}$, where $i = 1, 2, 3$ denote the three different measures of the smirk premia. I run the regressions for three different sample periods covered by: i) the combined data from April 1988 to June 2005, ii) the CBOE data from April 1988 to December 1995, and iii) the Ivy DB data from January 1996 to June 2005. The last column reports the correlations between $\{smirk_{i,t}^{data}\}$ and $\{smirk_{i,t}^{m}\}$.

the period covered by	β_2	$t(\beta_2)$	R^2	correlation
Measure 1, $i = 1$				
the combined data	0.592	5.05	0.240	0.490
the CBOE data	0.881	3.78	0.444	0.666
the Ivy DB data	0.406	2.24	0.156	0.396
Measure 2, $i = 2$				
the combined data	0.343	3.39	0.121	0.348
the CBOE data	0.368	2.88	0.184	0.429
the Ivy DB data	0.205	1.64	0.059	0.243
Measure 3, $i = 3$				
the combined data	0.617	3.07	0.091	0.302
the CBOE data	0.673	3.62	0.171	0.413
the Ivy DB data	0.550	1.28	0.043	0.209

Table VI
Matches of of stock pricing phenomena

Table VI reports the matches of the observed stock pricing phenomena. Panel A reports the average return volatility, the standard deviation of the risk-free rate, the average dividend-price ratio, the average risk-free rate and the average equity premium implied from both the (baseline) model and the data of the period from January 1996 to June 2005. Panel B reports the model's implications about the excess stock return predictability by running the long-horizon regressions of log excess stock returns onto the log price-dividend ratio in the simulated data. Panel C reports the test that the fluctuations in the historical dividend-price ratios are driven by the historical consumption shocks. I regress the historical dividend-price ratios series from January 1959 to December 2005, $\{DP_t^{data}\}$ onto their model counterparts, $\{DP_t^m\}$, simulated with the underlying state starting from $\{G_t^m\}$, the backed-out series of states from the historical consumption shocks, and the results are reported in the first three columns of Panel C. I report in the last column the correlation between $\{DP_t^{data}\}$ and $\{DP_t^m\}$.

Panel A: stock pricing moments (%)

	mean($volR$)	std(rf)	mean(CP)	mean(rf)	mean(EP)
model	16.3	4.27	1.74	-0.58	6.04
data	19.7	2.79	1.60	1.09	6.17

Panel B: return predictability

horizon (year)	β_2	$t(\beta_2)$	R^2
1	-0.369	-7.69	0.141
2	-0.607	-11.7	0.274
4	-0.798	-15.1	0.387
7	-0.86	-15.9	0.423

Panel C: historical dividend price ratio

regression coef.	t-stat.	R^2	correlation
0.357	5.52	0.132	0.363