

Risk preferences implied by synthetic options

Ian Dew-Becker and Stefano Giglio*

May 28, 2022

Abstract

A large literature finds that equity index options are overpriced in the historical data (starting in the 1980s), earning low returns and strongly negative alphas. In a representative agent framework, that fact implies that marginal utility is convex relative to stock market wealth – risk aversion rises as wealth falls. This paper provides novel evidence on risk aversion and the shape of marginal utility using nearly a century of data on *synthetic* options, constructed from dynamic portfolios of the market and the riskless asset. In contrast to exchange-traded options, synthetic options show no evidence of overpricing (i.e., of negative alpha), and, therefore, no evidence of risk aversion rising as wealth falls. The divergence in results between listed and synthetic options can be explained by segmentation in the derivatives market, due to regulatory constraints and illiquidity. As those constraints have relaxed in the past 10–20 years, the returns on listed options have converged to those on synthetic options.

1 Introduction

A central empirical fact of financial markets is that equity index options have been overpriced historically. Investors who purchase options have, on average, earned significant negative returns and negative alphas, of a magnitude that is difficult to rationalize in structural models. The high price of those options has been used to support two basic claims: that risk aversion rises as wealth falls, so that investors are much more averse to large than small declines, and that stock market volatility is priced. Those ideas, in turn, have driven theoretical research on preferences that might explain aversion to large declines or to fluctuations in volatility.

*Dew-Becker: Northwestern University and NBER. Giglio: Yale University and NBER. We appreciate helpful comments from seminar participants at ITAM, Yale, and the University of Connecticut.

At the same time, there is a well developed literature suggesting that the options market – and those for derivatives more generally – may be segmented from the broader financial market. Not all investors can trade options, and even those who can often face various constraints on their participation, whereas participation in simple equity investments is easy (and cheaper) for the vast majority of investors. If options markets are segmented, the price of options may not measure the preferences of the *average* investor, but instead reflect those of the relatively small number of agents trading options. In addition, options markets are still, in a historical sense, relatively young, having existed in their current form for only about 1/3 as long as we have data on equities.

Nevertheless, the questions that options have been used to address are important. A large literature asks how risk aversion varies over time and with wealth because it determines not just the behavior of asset prices (which then can affect the economy more broadly) but also because the objectives of macroeconomic policy depend on agents' utility functions. If utility is quadratic – which would imply linear marginal utility – then the variance of wealth is a sufficient statistic for welfare. But if marginal utility is convex in wealth, then policy should focus more strongly on controlling large declines in the economy.

This paper develops a novel approach to measuring the average investor's risk preferences by studying synthetic options. Our analysis focuses, like the almost entirety of the existing literature, on monthly option returns. The monthly return on an option can be well approximated (with an R^2 of about 80%) by a dynamic trading strategy that holds the underlying asset (in our case, the overall stock market) and the risk-free asset in amounts that can change on a daily basis. As an example, to synthesize a put option, one would initially hold a short position in the market, and a long position in the risk-free asset. If the market falls, the short exposure will grow, while if the market rises, the short position will shrink. In that way, holding a synthetic option is essentially a bet on mean reversion.

The paper's theoretical contribution is to show how and when such a method is able to measure features of marginal utility (and also, together with it, subjective probabilities). The basic assumption is that agents all agree on the price of the stock market. That is, agents may be heterogeneous, and markets may be incomplete, but under the assumption that they can all freely trade the market, they will agree on its equilibrium price. That assumption is in certain ways strong – especially in assuming that this condition holds every day – and so we show how it can be relaxed. On the other hand, the idea of using the stock market to measure the behavior of marginal utility is a foundational concept in asset pricing, and the assumption that we impose is entirely standard in the literature.

If the market is fairly priced every day, then any strategy constructed from time-varying

loadings on the market is also necessarily priced fairly (even if an individual investor would face high costs in implementing such a strategy in real time). That simple fact means that returns on synthetic options can be used to measure the behavior of marginal utility (i.e. state prices). Our key result is to show that by studying the properties of synthetic puts, we can estimate the coefficients of a regression of marginal utility on the market return (capturing the linear component marginal utility) and a nonlinear function of the market return (capturing the nonlinear shape of marginal utility). We show that if options earn a negative CAPM alpha – as is observed for exchange-traded options – that implies that marginal utility is convex in the market. And the same result holds for the alpha of synthetic options.

We then estimate it in the data. We begin by measuring the returns on synthetic options historically, using almost a hundred years of data (1926-2021). The first question we investigate is simply the extent to which synthetic option returns are indeed a nonlinear function of the market. We show that they do, and in fact that the part of their variation that is independent of the market is relatively small – only about 20% of the total. That nonlinearity is the core prerequisite for the analysis to be informative about nonlinearity in marginal utility.

The paper’s key result is on the CAPM alphas for the synthetic puts, which directly measure the degree of nonlinearity of marginal utility. Whereas exchange-traded puts have historically earned highly negative alphas, synthetic puts actually earn statistically and economically significant *positive* alphas, implying that, if anything, risk aversion *falls* as stock market wealth declines. The empirical regularity driving the result is simple: the returns on synthetic puts essentially measure the presence of mean reversion or momentum, and within months the stock market has historically displayed a small amount of momentum. If risk aversion rose as the market fell (i.e. if marginal utility was convex in the market return) we should instead see mean reversion in returns.

The results are robust to a wide range of modifications: they hold across strikes and maturities for the synthetic options, across different time periods, using different methods for choosing the weights used in the synthesis, and adjusting the strategy to account for possible measurement error and certain liquidity problems. The theoretical analysis also shows how drivers of marginal utility that are independent of the stock market can also potentially affect premia for synthetic puts. We argue that this effect should likely be small, and also provide quantitative evidence to support this, including a bound on its possible effects.

Overall, the paper’s core finding is that, according to synthetic options measured over

the last hundred years, there is no evidence for convexity in marginal utility as a function of stock market wealth – the pain associated with crashes does not rise faster than linearly. And, again, that result is in conflict with evidence from exchange-traded options. But if options markets are even partially segmented from broader financial markets, it need not be surprising that option prices would disagree with implications of the behavior of stock prices. That said, the barriers to trade have declined over time. Options markets have become more liquid, exchanges now make it easier for retail investors to trade options, and hedge funds act as vehicles allowing investors, such as endowments and pensions, to effectively get short option exposure, even though they are otherwise barred from doing so.

If market segmentation has declined over time, then the difference in returns between true and synthetic options should have also. And that is in fact precisely what we see in the data. In the earlier part of the sample – 1987 up to about 2008 – exchange-traded options earned significantly more negative returns than synthetic options, as has been found in past work. Since 2008, though, that gap has shrunk to zero. There is now no difference in average returns between true and synthetic puts, and also both have CAPM alphas of approximately zero (again, consistent with linear marginal utility). To put it another way, in the last decade it has not been true that index options are overpriced.

There is a large past literature estimating the shape of marginal utility as a function of the market return, with Ait-Sahalia and Lo (2000), Jackwerth (2000), and Rosenberg and Engle (2002) being key references.¹ That work has exclusively used option prices to measure state prices. But at the same time, it is well known that option prices have puzzling implications that are difficult or impossible to represent with standard utility theory. While they do imply convexity, with a significant premium for crash insurance, they also imply in some regions actually *negative* risk aversion. Numerous papers, therefore, argue that the evidence is much more in line with the view that options markets, for reasons of liquidity, transactions costs, regulatory constraints, margin costs, etc. described above, are partially segmented from the underlying equity market (Jackwerth (2000), Bollen and Whaley (2004), Bates (2008), Han (2008), Garleanu, Pedersen, and Poteshman (2008), Jurek and Stafford (2015), Frazzini and Pedersen (2022), among many others). That view implies that while option prices still may reveal marginal utility, they reveal it at best for a particular subset of investors, which will tend to change over time. By studying the entire equity market, we argue that our results are much more likely to measure the attitudes of a typical investor.

An additional prediction of the view that option prices are distorted due to segmentation is that when the limits to arbitrage such as trading costs and illiquidity shrink, option prices

¹More recently, see Welch (2016) and Beason and Schreindorfer (2022).

should move closer in line with the behavior (i.e. the lack of a put premium) that we observe in the underlying equity market. Our final contribution is to show that the data points to just that scenario playing out. The massive premium earned from selling put options has shrunk significantly in the past 20 years, to the point that it is near zero in the most recent decade of data. While confidence bands in such short samples are inevitably extremely wide, the point estimates imply that the put premium has gone away.

This paper is most directly related to work on crash risk, in terms of utility, policy, and how it affects asset prices. There is a large literature on the effect of rare disasters (Rietz (1988), Barro (2006)), with Gabaix (2012), Seo and Wachter (2019), and Beason and Schreindorfer (2022) studying option prices specifically.² Our contribution to that literature is to explore what the time-series dynamics of returns can reveal about how marginal utility is affected by crashes. There is also a literature on the sources of crashes themselves, including Farmer (2012), Bacchetta and van Wincoop (2016), and Petrosky-Nadeau, Zhang, and Kuehn (2018). In such models, there is potentially a role for policy, and knowing investor attitudes over those events is valuable for quantifying the benefit from preventing crises.

As discussed above, this paper is also related to a large literature on estimating marginal utility, or state price densities, as a function of the stock market return. Almost that entire literature measures significant convexity, which follows from the well known fact that out-of-the-money options earn significant market-adjusted returns, especially after delta hedging.³ There is also parallel work that develops no-arbitrage models to fit option prices (e.g. Pan (2002)), but those models often have unstable parameters and their out-of-sample performance can be weak (Bollen and Whaley (2004)). In general, that class of models implies significant premia on jumps and shocks to volatility, again due directly to the fact that even delta-hedged options earn significant returns (e.g. Bakshi and Kapadia (2003)).

The remainder of the paper is organized as follows. Section 2 discusses the theoretical framework and how it is applied to the data in practice. Section 3 describes the data and empirical methods and section 4 reports our main empirical results on the returns of synthetic puts and the shape of marginal utility. In section 5 we study the different behavior of exchange-traded put returns and synthetic put returns. Section 6 concludes.

²See also Schreindorfer (2020). Drechsler and Yaron (2011) study a consumption-based model with jumps, though the jumps are smaller than in the disaster literature.

³For a few examples in addition to the other work cited above, see Coval and Shumway (2001), Bakshi and Kapadia (2003), Constantinides, Jackwerth, and Savov (2013), and Muravyev and Ni (2020).

2 Theory

2.1 Definitions and notation

The market return between periods t and $t + j$ is $R_{t,t+j}^m$. The change in marginal utility is $MU_{t,t+j}$ (i.e. u'_{t+j}/u'_t , where u is utility over consumption). We also assume that agents may price assets according to some set of subjective conditional probabilities $P_{t,t+j}^*$. We denote the ratio of the subjective to the true probabilities by $S_{t,t+j}$.

Since $MU_{t,t+j}$ is a ratio of marginal utilities, we immediately have that $MU_{t,t+2} = MU_{t,t+1}MU_{t+1,t+2}$, etc. Similarly, we assume that the subjective probabilities satisfy a basic internal consistency condition on conditional probabilities: $S_{t,t+2} = S_{t,t+1}S_{t+1,t+2}$.

We say that an investment is priced by subjective marginal utility (SMU) if

$$1 = E_t [M_{t,t+j} R_{t,t+j}] \quad \forall t, j \quad (1)$$

$$\text{where } M_{t,t+j} \equiv MU_{t,t+j} S_{t,t+j}. \quad (2)$$

$M_{t,t+j}$ is the SMU and E_t is the expectation operator conditional on information available on date t (under the true probability measure).

We do not assume that markets are complete, meaning that marginal utility need not be identical across agents. Rather, we just assume that all agents agree on equation (1) for the market and risk-free asset, which will hold if they can trade those assets freely (this assumption is discussed further and relaxed somewhat in section 2.6). That is, we might say that each investor may have a different SMU, $M_{i,t,t+j}$, but that the M_i 's all satisfy $1 = E_t [M_{i,t,t+j} R_{t,t+j}]$ for all t and j for both the market and risk-free asset. In the rest of the analysis we suppress the i notation.

We assume that the risk-free rate is equal to zero for simplicity (equivalently, all returns can be interpreted as on forward contracts). That implies that $E_t M_{t,t+j} = 1$ for all t and j . The analysis is straightforward to recapitulate in the case where the risk-free rate is nonzero, and the empirical analysis accounts for nonzero interest rates.

2.2 What can we learn from returns on put options?

Suppose for the moment that returns on put options are priced by SMU. Define the gross return on a put option with strike k percent below the initial level of the market to be

$$R_{t,t+j}^k \equiv \max \{ k - (R_{t,t+j}^m - 1), 0 \} - P_t^k + 1 \quad (3)$$

where P_t^k is the initial price of the option.⁴ For convenience, we define as $\tilde{R}_{t,t+j}^k$ the part of $R_{t,t+j}^k$ that is orthogonal to the market return along with the associated alpha:

$$\tilde{R}_{t,t+j}^k \equiv R_{t,t+j}^k - \frac{\text{cov}_t(R_{t,t+j}^k, R_{t,t+j}^m)}{\text{var}_t(R_{t,t+j}^k, R_{t,t+j}^m)} (R_{t,t+j}^m - 1) \quad (4)$$

$$\alpha_{t,t+j}^k = E_t [\tilde{R}_{t,t+j}^k - 1] \quad (5)$$

If the option is priced by SMU, the empirical validity of which we will discuss later, then

$$\alpha_{t,t+j}^k = -\text{cov}_t(\tilde{R}_{t,t+j}^k, M_{t,t+j}) \quad (6)$$

And the same fact holds for the overall market. Since average returns reveal covariances with SMU, they can be used to get a linear projection for SMU:

$$M_{t,t+j} = \text{const.} - \frac{E_t[R_{t,t+j}^m - 1]}{\text{var}_t[R_{t,t+j}^m]} R_{t,t+j}^m - \frac{\alpha_{t,t+j}^k}{\text{var}_t(\tilde{R}_{t,t+j}^k)} \tilde{R}_{t,t+j}^k + \text{resid.} \quad (7)$$

where the residual term is orthogonal to $R_{t,t+j}^m$ and $\tilde{R}_{t,t+j}^k$.

Equation (7) is a nonlinear regression for SMU in terms of the market return. $\frac{E_t[R_{t,t+j}^m - 1]}{\text{var}_t[R_{t,t+j}^m]}$ measures the average slope of SMU with respect to the market return. $\tilde{R}_{t,t+j}^k$ is a piecewise linear function of the market return (with the typical kink at the strike k that put returns feature). So its coefficient, $\frac{\alpha_{t,t+j}^k}{\text{var}_t(\tilde{R}_{t,t+j}^k)}$, measures how the slope of SMU changes across the strike k .

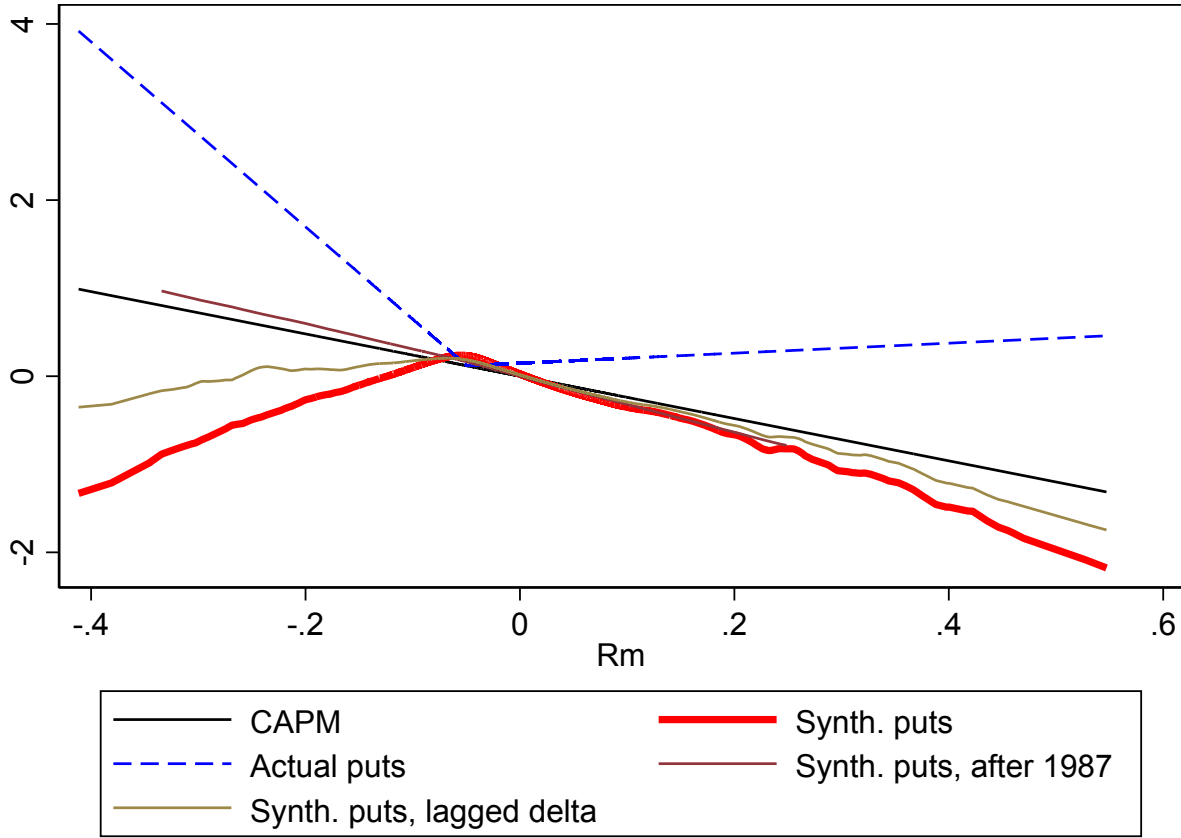
Equation (7) shows what the alpha of puts implies for marginal utility. The well known empirical result is that puts have negative alphas. Since $\tilde{R}_{t,t+j}^k$ is a convex function of the market, if puts are priced by SMU, then $\alpha_{t,t+j}^k < 0$ implies that SMU is convex in the market return.

Figure 1 illustrates that idea, plotting SMU relative to the market return under various models and assumption. Under the CAPM (black line), SMU is linear in the market return, all alphas are zero, there is no convexity, and the slope is recovered simply as $-\frac{E_t[R_{t,t+j}^m - 1]}{\text{var}_t[R_{t,t+j}^m]}$.

If puts are priced by SMU and they have a negative alpha, then one concludes that SMU is convex – in other words, crashes are especially painful, with marginal utility rising at an increasing rate as the market declines. Figure 1 illustrates that (dashed blue line) by

⁴The lack of a denominator here (i.e. some sort of date- t price) is nonstandard but convenient. Since $\max\{k - (R_{t,t+j}^m - 1), 0\} - P_t^k$ is an excess return, any scaling is also an excess return – i.e. including the price on date t or not.

Figure 1: SMU estimated using exchange-traded and synthetic puts



Note: The figure shows estimated SMU under different models and estimated in different samples. The solid black line reports the estimated SMU as a function of the market alone (as in the CAPM). The other lines model SMU as a function of the market and the orthogonalized returns on puts: listed puts (1987–2021), synthetic puts (1926–2021), synthetic puts for the post-1987 sample, and synthetic puts computed using one-day lagged beta.

plotting the SMU curve implied by the alphas observed for 5% out-of-the-money listed S&P 500 puts, during the period 1987–2021. The plot shows that put returns imply that effective risk aversion – as measured by the slope of SMU – is significantly higher when the market falls by a substantial amount.

2.3 What can we learn from returns on *synthetic* put options?

A simple summary of the previous section is that because the return on a put option is a nonlinear function of the market return, its alpha relative to the market measures the nonlinearity in SMU (the market itself captures the linear part).

It is well known that option returns can be approximated through dynamic trading in

the underlying asset – the market return in this case. If the synthetic put also has returns that are nonlinear in R^m , then it can also be used to measure nonlinearity in SMU.

Consider the following gross return on a synthetic put from t to $t + j$,

$$R_{t,t+j}^S \equiv \sum_{s=0}^{j-1} \beta_{t+s}^S (R_{t+s,t+s+1}^m - 1) + 1 \quad (8)$$

where β_t^S is a set of weights that can be a function of the values of R_s^m for $s \leq t$. Since R^m is priced by SMU, R^S is also:

$$1 = E_t [M_{t,t+j} R_{t,t+j}^S] \quad (9)$$

Just like with true puts, we can also calculate the alpha for synthetic puts, yielding

$$M_{t,t+j} = \text{const.} - \frac{E_t [R_{t,t+j}^m - 1]}{\text{var}_t [R_{t,t+j}^m]} R_{t,t+j}^m - \frac{\alpha_{t,t+j}^S}{\text{var}_t (\tilde{R}_{t,t+j}^S)} \tilde{R}_{t,t+j}^S + \text{resid.} \quad (10)$$

(where, as before, \tilde{R}^S is the part of R^S orthogonal to the market). So we again have an expression for SMU in terms of two returns, with coefficients depending on their means. In addition, if $\tilde{R}_{t,t+j}^S$ is a nonlinear function of the market return, its coefficient again measures nonlinearity in SMU.

2.4 Learning about SMU as a function of the market

Both \tilde{R}^S and M can be decomposed into parts that are nonlinear functions of the market return and residuals:

$$M_{t,t+j} = \bar{M}_{t,t+j} + \hat{M}_{t,t+j}, \text{ where } \bar{M}_{t,t+j} \equiv E [M_{t,t+j} | R_{t,t+j}^m] \quad (11)$$

$$\tilde{R}_{t,t+j}^S = \bar{R}_{t,t+j}^S + \hat{R}_{t,t+j}^S, \text{ where } \bar{R}_{t,t+j}^S \equiv E [\tilde{R}_{t,t+j}^S | R_{t,t+j}^m] \quad (12)$$

The components with overbars are those that are (in general nonlinear) functions of the market, while the circumflexes denote (uncorrelated) residuals. The reason to perform this decomposition is that it is the market-related part (the one with overbar) that allows us to learn about the part of SMU that is explicitly a function of the market return. Note that in the case of listed puts, we would have $\tilde{R}_{t,t+j}^k = \bar{R}_{t,t+j}^k$, and therefore $\hat{R}_{t,t+j}^k = 0$, since the put return depends only on the market return over the period t to $t + j$. This is not necessarily true for synthetic puts: its return may depend on the market return from t to $t + j$ as well as the path the market takes during that period: the latter component will be captured by

$\hat{R}_{t,t+j}^S$.

We now examine what synthetic puts tell us about SMU when residuals are potentially nonzero. Suppose first that $\hat{M}_{t,t+j} = 0$, so that SMU is a function *only* of the market return (as would hold under the CAPM, and more generally if the stock market return is a sufficient statistic for SMU). Then

$$M_{t,t+j} = \text{const.} - \frac{E_t [R_{t,t+j}^m - 1]}{\text{var}_t [R_{t,t+j}^m]} R_{t,t+j}^m - \frac{\alpha_{t,t+j}^S}{\text{var}_t (\bar{R}_{t,t+j}^S)} \bar{R}_{t,t+j}^S + \text{resid.} \quad (13)$$

That is, in this simple case, **we can interpret the alpha on a synthetic put in exactly the same way as that on a true put**: it measures nonlinearity in SMU as a function of the market return. Note that it does not actually matter whether R^S actually replicates the return on a put. What we need is that R^S yields nonlinear exposure to the market, so that we can measure the premium for such exposure and hence the shape of SMU.

The difference between equations (10) and (13) is that (10) involves $\tilde{R}_{t,t+j}^S$, while (13) involves $\bar{R}_{t,t+j}^S$, meaning that the latter gives SMU purely as a function of the market return.

2.4.1 A conditional CAPM interpretation

When β_t^S (the dynamic loading on the market taken by the synthetic put) is a function of the level of the market, $\alpha_{t,t+j}^S$ measures predictability in returns. Formally, one can derive from results in Lewellen and Nagel (2006) that

$$\alpha_{t,t+j}^S \approx \gamma \text{cov} \left(\beta_t^S, \frac{\gamma_t - \gamma}{\gamma} - \frac{\sigma_{M,t}^2 - \sigma_M^2}{\sigma_M^2} \right) \quad (14)$$

where $\gamma_t \equiv E_t R_{t,t+1}^m - 1$, $\gamma \equiv E[\gamma_t]$, $\sigma_{M,t}^2 \equiv \text{var}_t (R_{t,t+1}^m)$, and $\sigma_M^2 \equiv \text{var} [R_{t,t+1}^m]$ (15)

The first part of the covariance is the usual conditional CAPM intuition, which says that if β_t^S covaries positively with the market risk premium, then $\alpha_{t,t+j}^S$ will be positive. The second part is a contribution from the comovement of β_t^S with conditional volatility.

In our main analysis, we set β_t^S in order to build a replicating portfolio for puts – essentially setting β_t^S to be the Black–Scholes delta of a particular put option. For a put option, delta is negative, and becomes more negative as the market falls. So if γ_t rises when the market falls, then for a synthetic put, we will have that $\text{cov}(\beta_t^S, \gamma_t) < 0$, which, all else equal, would cause $\alpha_{t,t+j}^S$ to be negative. So that implies that when the market risk premium is countercyclical (i.e. returns are mean reverting), synthetic puts will earn a negative alpha.

The equation can also be interpreted in the opposite direction: if synthetic puts earn a negative alpha – implying that SMU is convex – then (holding volatility fixed) expected

returns must be countercyclical. That is, convexity in SMU implies countercyclical risk premia.⁵

One intuition for these results is the following. If marginal utility is convex, then a drop in the market moves the agent to a more curved part of the utility function, where risk premia γ_t are higher. This means that going forward expected returns should be higher, and it therefore generates mean-reversion in returns; in turn, this induces a negative alpha for synthetic puts. So the dynamics of returns and the alpha of synthetic puts can reveal the shape of the SMU function.

It is worth noting that this result does not hold in the same way for true puts. True puts can earn monthly alpha through the conditional CAPM mechanism, but they can also earn potentially earn alpha at the daily frequency. A synthetic put, on the other hand, must earn zero alpha at the daily frequency, because over each day the synthetic put just takes a static position. So it is possible for true puts to earn alphas without any predictability in market returns, simply due to being too expensive or too cheap, while a synthetic put can only earn an alpha at the monthly frequency through a market-timing effect.

2.4.2 The effect of non-market risk

Equation (13) is derived under the assumption that the market return is a sufficient statistic for SMU. Obviously that need not be true in general. In the fully general case, we can derive:

$$\bar{M}_{t,t+j} = \text{const.} - \frac{E_t [R_{t,t+j}^m - 1]}{\text{var}_t [R_{t,t+j}^m]} R_{t,t+j}^m - \frac{\alpha_{t,t+j}^S + \text{cov}(\hat{R}_{t,t+j}^S, \hat{M}_{t,t+j})}{\text{var}_t(\bar{R}_{t,t+j}^S)} \bar{R}_{t,t+j}^S + \text{resid.} \quad (16)$$

Since our objective is to understand how SMU is related – potentially nonlinearly – to the return on the market, we focus on just \bar{M} (there is also a regression-type equation for \hat{M} , that we ignore here). Relative to the baseline case where $\hat{M} = 0$, in the fully general case here, $\alpha_{t,t+j}^S$ must be adjusted when calculating the coefficient on the nonlinear part, \bar{R}^S .

That covariance just measures how the part of R^S that is independent of the market is priced. We address the potential presence of $\text{cov}(\hat{R}_{t,t+j}^S, \hat{M}_{t,t+j})$ in this equation in three ways: through an intuitive argument that it should be small; by formally measuring the covariance of \hat{R}^S with potential candidates for \hat{M} from the literature, and by bounding it in

⁵How does time-varying volatility affect this analysis? If the CAPM holds period-by-period with constant risk aversion (in the sense of the pricing kernel being linear in $R_{t,t+1}^m$), then $\gamma_t \propto \sigma_{M,t}^2$, which would imply that the covariance in (14) is zero. If risk aversion is countercyclical, as with a convex pricing kernel, then even if volatility rises when the market falls, γ_t will rise by enough to offset that effect, so that the covariance term is negative and $\alpha_{t,t+j}^S < 0$. In other words, if both volatility and risk aversion are countercyclical, equation (14) implies that $\alpha_{t,t+j}^S$ is negative.

the style of Cochrane and Saa-Requejo (2000).

First, the intuitive argument. $\hat{R}_{t,t+j}^S$ is the part of the return on the synthetic put that cannot be explained by *any nonlinear function* of the market return. In addition, though, R^S is constructed as a dynamic strategy on the market, so it only contains information that is in the *path* of the market return. That means, then, that $\hat{R}_{t,t+j}^S$ contains information about the path of the market return over the month, but no information about the total return over that month. It can measure, for example, whether the path to get to a given total monthly return was more versus less volatile.

There are many proposals in the literature for variables that might influence SMU beyond just the market itself, such as labor income risk, long-run growth, innovation, etc. It is difficult to imagine a scenario in which the path that the stock market takes after controlling for the total return over the month would somehow be related to labor income or innovation. For a typical investor, even if the stock market itself is not a sufficient statistic for SMU (i.e. $\hat{M} \neq 0$), it does seem as though the impact of the stock market on SMU would be very well summarized by the total return over the month, which would imply that $\text{cov}(\hat{R}_{t,t+j}^S, \hat{M}_{t,t+j}) = 0$.

To the extent that realized volatility during the month matters independently for pricing, and beyond its correlation with the market, a common argument is that that is due to market segmentation (e.g. Jurek and Stafford (2015)), where a relatively small number of dealers are net short options and charge a premium for unhedgable risk they take, which is related to volatility. That sort of premium is not due to the average investor being averse to volatility, but rather due to the cost of purely *cross-sectional* effects – some investors demand insurance and only a limited number of agents can supply it, driving up its equilibrium price. We argue in section 5 that the degree of segmentation in the options market has likely declined, and provide empirical evidence in support of that view.

All of that said, we will not ultimately rely on the assumption that $\text{cov}(\hat{R}_{t,t+j}^S, \hat{M}_{t,t+j}) = 0$ or $\hat{M} = 0$. The empirical analysis will explicitly try to measure or at least bound that covariance, to help quantify the nonlinearity of SMU.

2.4.3 What do we want in R^S ?

The analysis in this section helps show what we ideally need in R^S . First, in order to measure nonlinearity in M , we want R^S to be a nonlinear function of the market return – and ideally a nonlinear function that is economically interpretable (e.g. piecewise linear, like a put). On the contrary, if R^S were linear in the market (e.g. a static position), then it would contain exactly zero extra information compared to the market itself. At the same time, we want R^S

to depend as much as possible on the market, so that $\text{var}(\hat{R}^S)$ is small relative to $\text{var}(\bar{R}^S)$, which minimizes the potential error discussed in the previous section.

2.5 Listed puts versus synthetic puts

Given the above considerations for an ideal R^S , and given that the previous section discusses the behavior of a synthetic put, it seems natural to study the returns on traded options. Options have returns that are nonlinear in R^m and also fully spanned by it. There are two major drawbacks to options that have been raised in past work, though, that lead us to take the novel perspective of learning about the SMU by studying the dynamics of the market return instead.

The first problem is that data for options with monthly maturity only begins in 1987, whereas synthetic puts can be constructed as long as we have daily data on market returns, so at least to 1926.

The second potential concern with options is that even if the overall stock market is priced by the marginal utility of some hypothetical representative agent – which is all that is required for synthetic puts – that assumption may be far less reasonable for options. In much of the sample, retail investors could not easily trade options, when they can the options are also often illiquid, with wide bid/ask spreads and little volume and open interest, especially for relatively deep out-of-the-money strikes. Last, there is also substantial evidence that option prices are heavily influenced by intermediary frictions, implying that their prices reveal more about the marginal utility of a subset of investors (the retail investors demanding options and the dealers selling them).⁶

So while puts are theoretically ideal, characteristics of their market, segmentation in particular, lead us to study synthetic puts as an alternative. Synthetic puts have drawbacks of their own, which we discuss in the next section, but they nevertheless provide a novel perspective on investor risk attitudes. We analyze the relationship between returns on true and synthetic puts empirically in section 5, showing that they have converged over the last 10–15 years.

2.6 Transaction costs

A natural concern if we are studying the return on a synthetic put is that this is the return on a strategy that may be very expensive or simply impossible to implement, especially for a

⁶See, e.g., Bollen and Whaley (2004), Garleanu, Pedersen, and Poteshman (2008), and Frazzini and Pedersen (2022)

retail investor. But recall, though, that the analysis only relies on the simple (and standard) premise that $1 = E_t [M_{t,t+1} R_{t,t+1}^m]$. If that is true, then any linear combination of market returns is also priced by SMU, which is what $R_{t,t+j}^S$ is (equations (8) and (9)). Since we want to know about covariances with SMU, in order to obtain theoretical regression coefficients, we are using that pricing fact, which explicitly does *not* include any transaction costs (which is also why transaction costs are typically not included in empirical asset pricing studies).

The assumption that $1 = E_t [M_{t,t+1} R_{t,t+1}^m]$ is the statement that the market is correctly priced – by investors’ SMU – every day. One could imagine situations under which that may fail to hold in the data. As a simple example, suppose that in the data, on any given date, the prices of some stocks are stale due to a lack of trading. Then the market prices that are observed, and hence also the returns, have what is essentially a form of measurement error, which will be most severe at high frequencies.

A second example is the case where there are transaction costs on the market itself. Then at any given time, the market can deviate from its fair price (according to SMU) by the magnitude of the transaction cost, giving a window of potentially valid prices.

In standard applications, this issue would be often addressed by studying lower-frequency returns, like at the weekly or monthly level. For this paper’s analysis, though, daily returns are necessary in order to effectively create the synthetic puts. So suppose that while the pricing equation holds for R^m , we observe a contaminated return R^{m*} , with

$$R_{t,t+1}^{m*} = R_{t,t+1}^m + \varepsilon_{t+1} \quad (17)$$

where ε_{t+1} is the error. In the two examples above, ε_{t+1} would either come from the fact that some of the prices used in constructing $R_{t,t+1}^{m*}$ are stale, or that they deviate from the values implied by SMU due to transaction costs.

Recall that our core results are about how mean returns reveal covariances with SMU. So how does the presence of an error ε_{t+1} affect those means and their relation to the covariance with SMU? First, clearly, in order for our main results to hold, we need $E[\varepsilon_{t+1}] = 0$, so that when we estimate $E[R_{t,t+1}^{m*}]$ from the data it is also a valid estimator for $E[R_{t,t+1}^m]$. This simply requires that the stock market is not observed in a systematically biased way.

Second, in order for $E[R_{t,t+1}^S]$ to be equal to $E[R_{t,t+j}^{S*}]$ (where R^{S*} is calculated using R^{m*} in place of R^m), it must be the case that

$$\text{cov}(\beta_t^S, \varepsilon_{t+1}) = 0 \quad (18)$$

In our application, variation in β_t^S is driven by variation in past returns on the market,

so there is a bias if ε_{t+1} is correlated with lagged values of $R_{t-1,t}^{m*}$ (i.e. with lagged values of either $R_{t-1,t}^m$ or ε_t). A prime example is stale prices. Suppose in a given day, only half of the prices observed in the market are updated. Then if fundamentals (i.e. the “true” return, $R_{t-1,t}^m$) improve by 1% on day t , we would observe a return $R_{t,t+1}^{m*}$ of 0.5% on day t and then another 0.5% on day $t+1$ as the remaining stock prices are updated. That is, stale prices create positive one-day serial correlation in returns (but nothing at longer lags).

Under such a scenario, (18) being true would require β_t^S to be orthogonal to news on date t (since that predicts returns on date $t+1$). An easy way to address that is to make β_t^S a function only of information available on day $t-1$ or earlier, which we do in section 4.3.

More generally, this section shows what types of assumptions about pricing are required for the main results to hold. If the market is correctly priced relative to SMU at the end of every day (which is more reasonable in recent decades), then the baseline results are valid. If, on the other hand, the observed market price is not always correct and has some noise, then that noise needs to be orthogonal to the portfolio weights. Given a specification of that noise – e.g. single-day persistence due to stale prices – one can construct a β_t^S rule that is robust to those errors and will still measure the curvature of SMU correctly.

An alternative interpretation of the effect of transaction costs is that they might create a time-varying liquidity premium for the market. For example, perhaps when the market falls, transaction costs (e.g. through bid/ask spreads) become larger. That might cause a liquidity premium on the market to rise in bad times. Such an effect would mean that expected returns on the market would be high in bad times and we would observe mean reversion. As discussed in section 2.4.1, mean reversion will cause us to observe negative alphas for synthetic puts, and it therefore biases the analysis towards finding convexity in SMU.

3 Data and methods

3.1 Choosing time-series weights, β_t^S

We select the weights β_t^S on a daily basis to try to replicate option returns at horizons between a month and a year. Our main analysis follows the method of Hull and White (2017), which modifies the standard Black–Scholes hedging method to account for certain types of stochastic volatility. Intuitively, the weight β_t^S is equal to the delta of the option – the partial derivative of the value with respect to the price of the underlying. For puts, as the market falls, β_t^S becomes progressively more negative, but it is always bounded between 0 and -1.

In choosing which put to synthesize, we can vary both the strike and maturity. The strike price determines the point at which the slope of $R_{t,t+j}^S$ with respect to $R_{t,t+j}^m$ changes. So by varying the strike we can examine convexity at different levels of market returns.

Following past work, we focus on returns to maturity on one-month options, but we also examine results for alternative maturities. Since we are not studying traded options, we can calculate returns for overlapping periods. That is, with the convention that there are 21 trading days in every month, we calculate $R_{t,t+21}^S$ for every day t . Standard errors are then always corrected for this overlap with the Hansen-Hodrick method using 21 lags.

3.2 Synthetic options and the market return

Since we want to understand how SMU varies as a function of the market return, what matters for our results is how R^S varies with the market return. Specifically, we need to estimate the general nonlinear function $E[R_{t,t+j}^S | R_{t,t+j}^m]$. To implement that empirically, we draw on the large literature on nonparametric regression to fit the equivalent equation

$$R_{t,t+j}^S = g(R_{t,t+j}^m) + \eta_{t,t+j} \quad (19)$$

where g is an arbitrary unknown function that is uncorrelated with the mean-zero residual $\eta_{t,t+j}$. g is estimated from a local linear regression with a Gaussian kernel and the bandwidth set to 0.01 (corresponding to a one-percentage point window in the market return).

3.3 Data

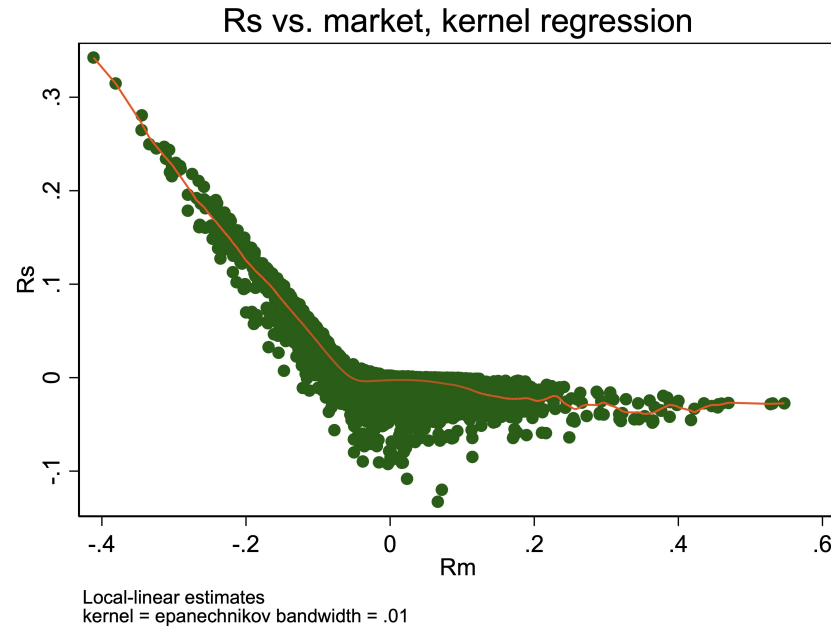
To implement the analysis described in section 2, we just need a time series for the market return, $R_{t,t+1}^m$, and the risk-free rate. We use the daily data from Kenneth French's website, and throughout the analysis t is taken to be a day.

4 Results

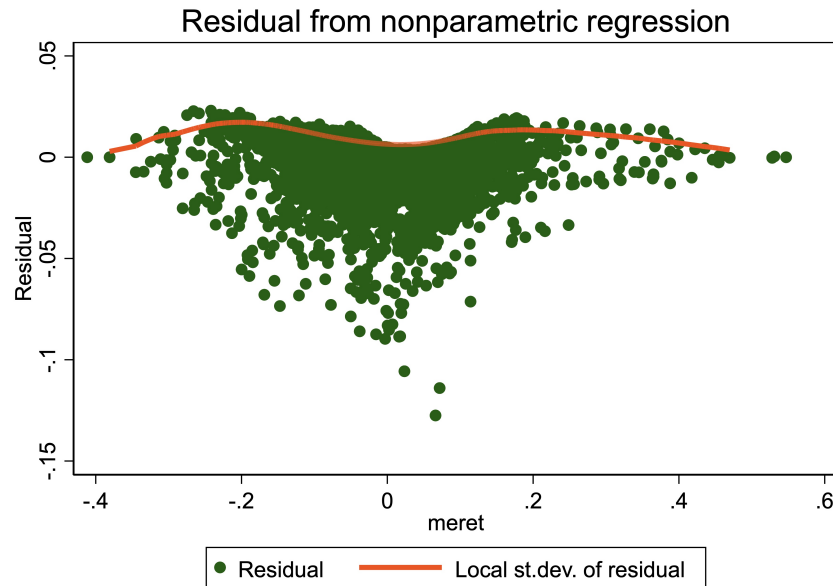
4.1 The relationship between R^S and R^m

The top panel of Figure 2 plots our benchmark $R_{t,t+j}^S$ against $R_{t,t+j}^m$, where $j = 21$ days and the strike used to construct β_t^S is 95% of the initial level of the market (corresponding to approximately a unit standard deviation decline). There is clearly significant nonlinearity – for values of the market return above the strike (-5%), the slope is near zero, while for values below it the slope is approximately -1, consistent with the fact that $R_{t,t+j}^S$ is

Figure 2: Synthetic put returns as a function of the market



(a) R^S vs. R^m



(b) Residuals and their s.d.

Note: Panel (a) shows the scatterplot of the returns to the synthetic put, R^S , against the returns of the market, R^m . The red line is a kernel estimate of the local mean. Panel (b) shows the residuals of the nonlinear fit from panel (a) against the market, and, in red, the local standard deviation of the residuals.

constructed to mimic a put option. The red points plot the nonparametric estimates of $\bar{R}_{t,t+j}^S = E[R_{t,t+j}^S | R_{t,t+j}^m]$. They formally quantify the relevant nonlinearity.

In addition, note that $\bar{R}_{t,t+j}^S$ rises with a consistent slope as $R_{t,t+j}^m$ falls, regardless of how large the decline is. If one thought that it was not possible to span large declines in the market with time-varying weights, e.g. due to large jumps, that fact would be revealed as $\bar{R}_{t,t+j}^S$ flattening out for the very negative values of $R_{t,t+j}^m$. There is no evidence of such a pattern, though. In the months with the most negative returns, the dynamic strategy continues to deliver positive returns, even including major events like the crashes in 1929 and 1987 (again, the claim here is not necessarily that an investor could have followed this strategy in real time to avoid crashes).

Figures A.1 and A.2 in the appendix replicate the top panel of Figure 2, but for the range of other strikes and maturities that we study. It shows that the pattern in Figure 2 is shared by the other synthetic puts we study.

The bottom panel of Figure 2 plots the residuals $\hat{R}_{t,t+j}^S \equiv R_{t,t+j}^S - E[R_{t,t+j}^S | R_{t,t+j}^m]$ against $R_{t,t+j}^m$ along with their local standard deviation.⁷ Due to the way the residuals are constructed, they are guaranteed to have a local mean, weighted by the kernel, equal to zero (and thus also a global mean of approximately zero). The residuals have relatively less volatility when the market return is close to zero and at its most extreme values, and greater volatility at intermediate values.

4.2 Risk premia

Table 1 reports our key results for risk premia. The risk premia – the alphas for synthetic puts – are really the central objects of interest since they measure, up to a scaling factor, the nonlinearity in SMU from equation (13). Recall that the well known empirical result for true puts is that the alphas are negative.

The top panel holds the maturity, j fixed at 21 days (one month) and varies the strike between 90% and 110% of the initial level of the market. The bottom panel varies the maturity between 21 and 252 days. The strike in those cases is set to $-5\% \times \sqrt{j/21}$, so that it is equivalent in standard deviation terms across maturities. All returns, Sharpe ratios, and information ratios are annualized.

⁷The local volatility is estimated from a second kernel regression,

$$\eta_{t,t+j}^2 = h(R_{t,t+j}^m) + residual \quad (20)$$

where for the function h we set the bandwidth to 0.05 due to there being greater variation in the squared residuals around the fitted value.

Table 1: Risk premia of synthetic puts

M	K	Mean	SR	beta	alpha	IR
21	Market	0.08 (0.02)	0.44 (0.10)			
21	0.9	0.00 (0.00)	-0.10 (0.07)	-0.08 (0.02)	0.00 (0.00)	0.07 (0.10)
21	0.95	0.00 (0.01)	0.00 (0.09)	-0.22 (0.02)	0.02 (0.01)	0.34 (0.10)
21	1	-0.01 (0.01)	-0.09 (0.10)	-0.52 (0.03)	0.03 (0.00)	0.56 (0.08)
21	1.05	-0.06 (0.02)	-0.38 (0.10)	-0.81 (0.04)	0.00 (0.00)	0.09 (0.06)
21	1.1	-0.07 (0.02)	-0.43 (0.10)	-0.91 (0.04)	0.00 (0.00)	-0.04 (0.05)
M	K	Mean	SR	beta	alpha	IR
21	0.95	0.00 (0.01)	0.00 (0.09)	-0.22 (0.02)	0.02 (0.01)	0.34 (0.10)
63	0.91	0.00 (0.00)	-0.03 (0.07)	-0.16 (0.02)	0.01 (0.00)	0.26 (0.10)
126	0.88	0.00 (0.00)	-0.08 (0.05)	-0.14 (0.02)	0.01 (0.00)	0.19 (0.08)
252	0.83	0.00 (0.00)	-0.07 (0.04)	-0.11 (0.02)	0.01 (0.00)	0.18 (0.07)

Note: Table reports characteristics of the market return (top row) and, in the other rows, returns on synthetic puts with a variety of maturities M and strikes K (as a fraction of spot). For each combination of M and K, the table reports the average return, Sharpe ratio, market beta, alpha, and information ratio, all annualized. Standard errors in parenthesis. Sample is 1926-2021.

The top row reports results for the market return, giving a measure of the average slope of SMU. The Sharpe ratio in the sample is +0.44, with a standard error of 0.10. That value is a useful benchmark that will help interpret the magnitude of the premia on other investments, and it also will be relevant in calculating the projection for SMU.

4.2.1 Varying strikes at the monthly maturity

The remaining rows of table 1 report results for synthetic puts for varying strikes and maturities. For the monthly maturity, looking across strikes, the betas are all negative and become more negative as the strike price rises, as one would expect for synthetic option returns. The average returns and Sharpe ratios for the R^S portfolios are all negative except for one, but since the betas are negative, the portfolios actually have significantly *positive* alphas and information ratios (the ratio of $\alpha_{t,t+j}^S$ to $std(\tilde{R}_{t,t+j}^S)$), with the exception of the strike of +10%.

The information ratios (IRs) vary in terms of statistical and economic significance across the strikes. The largest IR is for at-the-money options, where it is 0.56, with a t-statistic of about 7. That information ratio means that the independent part of the synthetic put actually earns a larger (positive) premium than the market itself. That said, we are not actually claiming that an investor could have earned such a premium in practice for most of the sample. Rather, this shows that there has historically been significant positive serial correlation in market returns, since the at-the-money put option has a relatively more negative exposure to the market following declines and a relatively more positive exposure following increases. In turn, this implies that SMU should be concave rather than convex (as the market drops, expected return drop as well, and vice versa).

As the strike moves away from zero in either direction, the magnitude of the IRs fall. The only other IR that is statistically significant is for the -5% strike. That having been said, in all cases but one, the point estimates are positive, and the (Hasnen-Hodrick) standard errors are small enough that we can rule out large negative values (such as those for traded puts). So whereas traded puts in the period 1987–2021 have negative alphas, synthetic puts over the long sample 1926–2021 have *positive* alphas in general, and in some cases with statistically and economically large magnitudes.

It is also worth noting that the long sample for synthetic puts gives significantly more precision than the sample for traded options – the standard error of a Sharpe ratio in a 35-year sample is 0.17, and in the post-1996 Optionmetrics sample often used for options it would be 0.20. So we get as much as double the precision compared to more recent samples.

4.2.2 Varying maturities at a fixed strike

The bottom panel of table 1 reports results for different maturities between one month and one year, holding the strike fixed in standard deviation terms equivalent to -5% at the monthly horizon. The information ratios decline, though not severely, as the maturity increases. In all cases they remain statistically significantly positive, meaning that we can rule out, at the 5% confidence level, any convexity in SMU around the -5% strike across horizons.

4.2.3 Interpretation

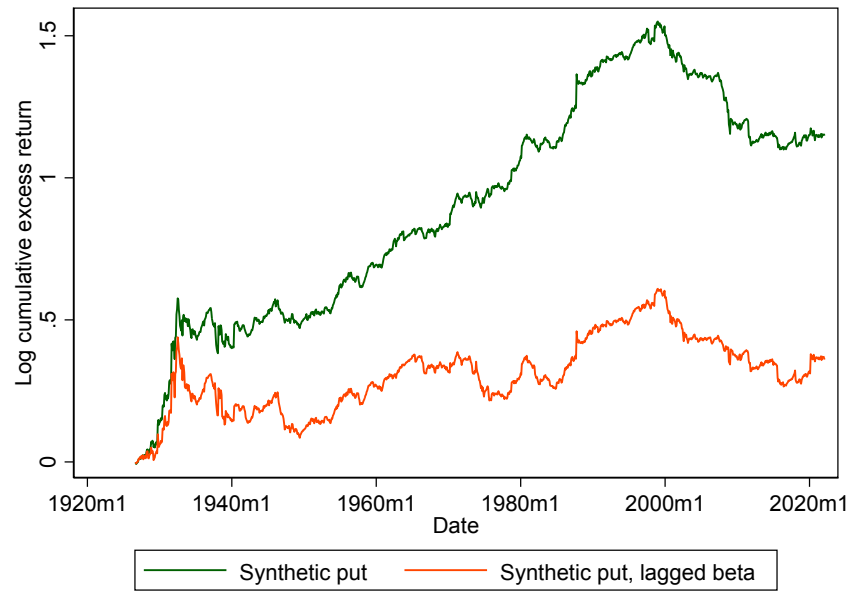
In the simple case where the market return is a sufficient statistic for SMU, the alphas here immediately imply that SMU is actually *concave* in the market return. That is, the prima facie evidence here is that rather than investors being especially strongly averse to large declines in the market, if anything they find large declines to be *less* than proportionately painful in terms of how they raise marginal utility compared to small declines.

Figure 1 plots, in red, the shape of SMU implied by the synthetic option returns for the strike of -5% and maturity of one month. Whereas the SMU curve estimated from traded options is convex, SMU based on synthetic options is significantly concave, implying that effective risk aversion actually declines as the market falls.

To help understand the implications of the pricing results, the last column of table 1 reports the R^2 from the nonlinear regression of $R_{t,t+j}^S$ on $R_{t,t+j}^m$. For a traded option return, that R^2 would be 1. Here, the values range between 0.73 and 0.92, depending on the strike. In other words, consistent with Figure 1, the vast majority of the variation in $\tilde{R}_{t,t+j}^S$, across all strikes and maturities, is explained by the market return. The smaller is that residual, the smaller is the potential bias identified in section 2.4.2. The question is just whether that residual is priced (i.e. whether $\text{cov}(\hat{R}_{t,t+j}^S, \hat{M}_{t,t+j}) \neq 0$), and the potential magnitude of that effect, which is the focus of the next section.

To see how the returns have varied over time, Figure 3 plots cumulative excess returns of $\tilde{R}_{t,t+j}^S$ over the sample (for a strike of -5% and $j = 21$). \tilde{R}^S has, by construction, zero exposure to the market return, so under the CAPM, i.e. if the pricing kernel is linear in the market return, \tilde{R}^S would have zero average excess return. The cumulative return over the sample is positive, consistent with the results in Table 1. In the time series, that positive return is driven primarily by the period between 1940 and 1980. Prior to 1940, \tilde{R}^S earned negative returns due to the fact that the market was falling over much of that period during the Great Depression. More interesting is the fact that the returns change sign on average after about 1987. That fact is well known for exchange traded options, but it is notable that

Figure 3: Log cumulative returns on \tilde{R}^S



Note: Log cumulative returns on \tilde{R}^S , for the period 1926-2021. Green line is the synthetic put, red line is the synthetic put with one-day lagged beta.

it also holds in \tilde{R}^S , which is a synthetic option return (see section 5).

4.3 Robustness

As discussed in section 2.6, if there is noise in the observed market return that is correlated with β_t^S , that can cause a bias in the results. We focus on stale prices as the most likely candidate that could cause a bias due to one-day serial correlation in returns.

The simple way to handle that scenario is to choose β_t^S to depend only on information available at date $t - 1$ or earlier. Formally, we do that by choosing β_t^S to be the hedge ratio for a synthetic option based on the level of the market on date $t - 1$. Results for that case along with the benchmark are reported in table 2. Interestingly, the alphas do generally shrink towards zero. That is due to the fact that there is in fact positive serial correlation in returns at the one-day lag in the early part of the sample (of about 10%). By lagging β_t^S , we eliminate the effect of that serial correlation. We still see no sign of a negative alpha, though, and many of the point estimates remain positive, if much smaller.⁸

Figure 3 plots the cumulative returns from this strategy. They are clearly significantly lower than the baseline in the first part of the sample, but subsequent to 1980 the two series move in parallel. Figure A.4 in the appendix also reports results lagging information by an additional day to date $t - 2$ and shows that the results are highly similar to those for the single-day lag.

Intuitively, the positive alphas are due to some persistence at the daily level, but we do not see evidence of negative serial correlation at any horizon out to one month, which is what would be necessary in order for the synthetic puts to earn negative alphas, as would be predicted if SMU were convex. To see that, figure 4 (panel a) plots the autocorrelation function for stock market returns from 1 to 21 days for the full sample along with the first and second halves (splitting in 1973), where we take the latter period as where it becomes easier to trade the market (especially as index futures appear in the 1980's).

The figure shows that there is a surprisingly large one-day autocorrelation in market returns, especially in the first half of the sample, but no notable negative autocorrelation. For context, the standard errors are approximately equal to 0.029 for the full sample and 0.042 in the two halves (under the null that returns are white noise, the standard error is $1/\sqrt{T}$). The bottom panel of the figure plots cumulative autocorrelations and also finds no evidence for reversals.

Figure 1 plots an estimate of SMU based on this alternative measure of the premium on

⁸Figure A.3 replicates figure 2 for this alternative version of β_t^S and shows the behavior is highly similar to the baseline case.

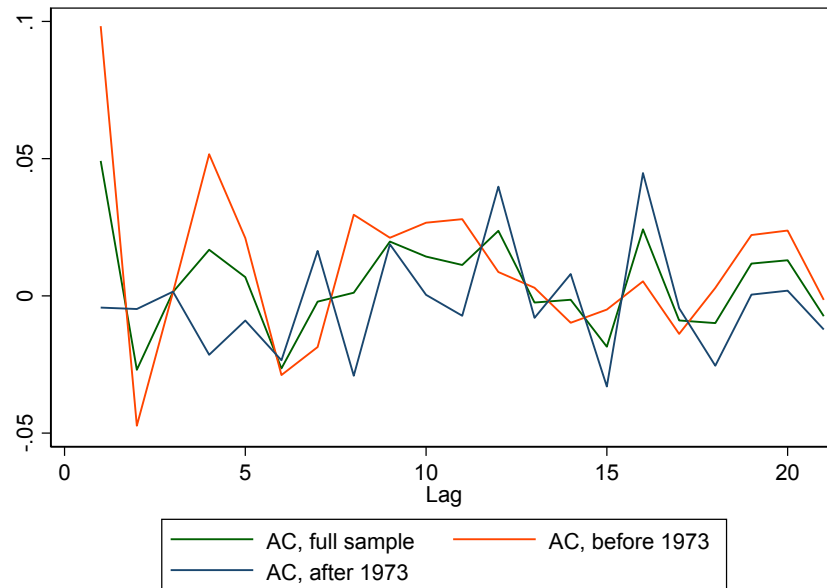
Table 2: Robustness of risk premia estimates

M	K	Information Ratio				
		Baseline	Lagged Beta	Constant IV	Before 1973	After 1973
21	0.9	0.07 (0.10)	0.03 (0.10)	0.00 (0.09)	0.16 (0.14)	-0.07 (0.13)
21	0.95	0.34 (0.10)	0.17 (0.10)	0.30 (0.09)	0.48 (0.14)	0.16 (0.14)
21	1	0.56 (0.08)	0.20 (0.08)	0.47 (0.08)	0.81 (0.12)	0.25 (0.12)
21	1.05	0.09 (0.06)	-0.02 (0.06)	0.13 (0.06)	0.14 (0.09)	0.06 (0.09)
21	1.1	-0.04 (0.05)	-0.05 (0.05)	-0.02 (0.04)	0.01 (0.08)	-0.10 (0.08)

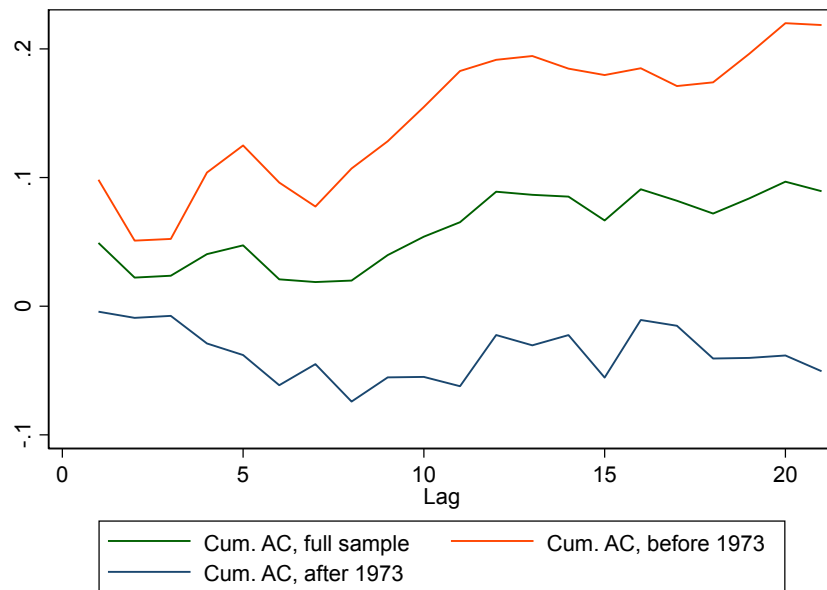
M	K	Information Ratio				
		Baseline	Lagged Beta	Constant IV	Before 1973	After 1973
21	0.95	0.34 (0.10)	0.17 (0.10)	0.30 (0.09)	0.48 (0.14)	0.16 (0.14)
63	0.91	0.26 (0.10)	0.16 (0.10)	0.20 (0.10)	0.40 (0.13)	0.02 (0.15)
126	0.88	0.19 (0.08)	0.12 (0.08)	0.12 (0.08)	0.36 (0.11)	-0.05 (0.13)
252	0.83	0.18 (0.07)	0.12 (0.07)	0.10 (0.07)	0.27 (0.10)	-0.02 (0.10)

Note: Table reports the annualized information ratio of synthetic puts of various maturities M and strikes K (as fraction of spot). Each column represents a variation in the way synthetic puts are constructed. The first column is the baseline as in table 1. The second column lags the beta by one day. The third column uses a constant IV of 0.15. The sample for the first three columns is 1926-2021. The last two columns split the sample in two halves: before and after 1973. Standard errors in parenthesis.

Figure 4: Daily autocorrelation of returns



(a) Autocorrelations



(b) Cumulated autocorrelations

Note: Panel (a) shows the autocorrelations of daily returns up to 21 lags. Panel (b) shows the cumulation of those autocorrelations. Each panel shows autocorrelations using the full sample (1926-2021), and using the pre- and post-1973 data separately.

synthetic options (light brown line). We view this as the most reasonable of the estimates, and it implies that SMU is close to linear, consistent with the CAPM.

As a second robustness test, we change the volatility model used in selecting β_t^S . The baseline results in this paper use a HAR model to forecast volatility (which is a simple and widely-used model for forecasting volatility). Table 2, though, shows that the alphas are essentially identical to those in the baseline if we simply set the volatility used to calculate the hedge weights to 0.15 on all dates (central column). That shows that the results are driven by how β_t^S depends on the level of the market, rather than its volatility.

The last two columns of table 2 split the sample in half (again in 1973).

4.4 The effect of unspanned variation – $\hat{R}_{t,t+j}^S$ and $\hat{M}_{t,t+j}$

The theoretical analysis shows that the alphas estimated for $R_{t,t+j}^S$ will be biased estimates of nonlinearity in SMU if there is a nonzero correlation between $\hat{R}_{t,t+j}^S$ and $\hat{M}_{t,t+j}$ – that is, if the part of $R_{t,t+j}^S$ not related to the market return (in any way, even nonlinearly) is related to some component of marginal utility that is also not related to market returns. That bias term, since it is unobservable, effectively represents another form of uncertainty that we must account for, in addition to standard sampling uncertainty accounted for by the usual statistical standard errors. This section examines two ways of addressing that term. First, it asks whether the residuals are correlated with any other known pricing factors. Second, it uses an analysis similar to that of Cochrane and Saa-Requejo (2000) to bound the magnitude of the premium on the unspanned component.

4.4.1 Relationship with known risk factors

A wide range of variables have been proposed that could be related to marginal utility beyond market returns. Table 3 reports correlations between \hat{R}^S and different variables. For $\hat{R}_{t,t+j}^S$ we take the value over each calendar month – i.e. the return from the beginning to the end of a month. For the variables that are measured at a fixed point in time – the Fed Funds rate, term spread, default spread, VIX, and VXO, we take the statistical innovation in the value at the end of the month relative to the lags of the variable (information available at the beginning of the month). For the other variables, which are either flows or measured over the course of each month, we take the statistical innovation in the monthly value relative to data available in the previous month. We are thus measuring the extent to which shocks to the unspanned part of returns are correlated with shocks to other variables. Given that the different time series are available for different time periods, each correlation is computed using the longest period available for that variable.

Table 3: Correlation of residuals with macro variables

	All data	Excluding 2020
Unemployment	-0.13	-0.05
Ind. Pro. Growth	0.12	0.11
Employment growth	0.13	0.09
FFR	0.04	0.02
Term Spread	-0.01	0.00
Default Spread	-0.09	-0.07
EBP	-0.21	-0.18
VIX	-0.28	-0.28
VXO	-0.16	-0.17
rv	-0.44	-0.42
Maximal corr	0.50	0.47

Note: Table reports the correlations between the residuals of the nonlinear fit of R^S onto the market and various macroeconomic variables: unemployment, industrial production growth, employment growth, the federal funds rate, the term spread (10 year minus 1 year), the default spread (BAA-AAA spread), the excess bond premium (EBP) from Gilchrist et al. (2021), the VIX, the VXO, and realized volatility. All variables are orthogonalized to the market. The last row reports the maximal correlation between any linear combination of these variables and the residuals. The second column replicates the results excluding 2020.

Since $\hat{R}_{t,t+j}^S$ is orthogonal to the market return by construction, we also orthogonalize the innovations in all of the other macro and financial time series with respect to the market return. So the correlation between $\hat{R}_{t,t+j}^S$ with innovations in the default spread represent changes in the default spread that are separate from the stock market return, and the same as true for all the other time series considered here.

Among the macro time series, the correlations are all economically small and statistically insignificant. The only notable correlations are for price series: the expected bond premium (EBP), the VIX, and realized volatility. In months in which shocks to these price series, after orthogonalizing with respect to the market return, are unexpectedly high, $\hat{R}_{t,t+j}^S$ tends to be low. If those are bad states of the world, that would make $\hat{R}_{t,t+j}^S$ risky, which we will show below would imply more convexity in SMU than our point estimates above.

It is important to emphasize again, though, that the correlations here are for series orthogonal to the market. For example, the correlation with realized volatility says that, holding the return on the market fixed, months in which the path to get to that return is more volatile are months in which $\hat{R}_{t,t+j}^S$ is lower. In order for that to affect pricing, it must be that investors' marginal utility does not depend only on some function of the market return, but also on the *path within the month* that the market takes to get to that point.

That fact therefore rules out simple models in which utility is a function of current wealth (as measured by the stock market).

The fact that realized volatility is related to $\hat{R}_{t,t+j}^S$ is well known in the literature. $\hat{R}_{t,t+j}^S$ is closely related to the delta-hedged gain on an option, which has been studied in numerous papers. While one interpretation of those results has been that marginal utility must depend on the path that returns take, we discuss an alternative below.

The bottom row of table 3 reports the maximum correlation of $\hat{R}_{t,t+j}^S$ with any linear combination of the innovations (just $\sqrt{R^2}$ from a linear regression). It is only 0.5. We take that as an upper end for a reasonable estimate of the correlation of $\hat{R}_{t,t+j}^S$ with $\hat{M}_{t,t+j}$ (though we will also explore the extreme case where the correlation is 1).

4.5 Robust uncertainty bands

Our basic goal is to fit a theoretical regression for SMU of the form

$$\bar{M}_{t,t+j} = cons + a_1 R_{t,t+j}^m + a_2 \bar{R}_{t,t+j}^S + residual \quad (21)$$

Since $R_{t,t+j}^S$ is the return on a synthetic option, its shape is, in theory, approximately piecewise linear, and figure 1 shows that holds reasonably accurately in practice. The parameter a_1 – which, recall, is estimated from the mean and variance of the market return – measures the average slope of SMU with respect to $R_{t,t+j}^m$. The parameter a_2 (which depends on $\alpha_{t,t+j}^S / \text{var}(\bar{R}_{t,t+j}^S)$ and the bias term, $\text{cov}(\hat{R}_{t,t+j}^S, \hat{M}_{t,t+j})$) measures essentially the change in the slope across the strike used in constructing $R_{t,t+j}^S$.

Table 4 reports estimates of a_1 and a_2 for a variety of strikes for the synthetic puts (across columns), and using different methods to account for the uncertainty in these parameters. It begins by reporting the average slope a_1 , equal to $-E_t[R_{t,t+j}^m - 1] / \text{var}(R_{t,t+j}^m)$. Since that depends only on the market return, it is invariant to the choices about the construction of $R_{t,t+j}^S$. Naturally, since the market return has historically been positive, we find that the slope is negative. The second row of table 4 reports a 95-percent confidence band (using Hansen-Hodrick (1980) standard errors).

The third row of table 4 reports the implied change in the slope of SMU above versus below the strike price for $R_{t,t+j}^S$, taking $a_2 = -\alpha_{t,t+j}^S / \text{var}(\bar{R}_{t,t+j}^S)$ as our “point estimate”, where, in addition to the usual statistical sampling uncertainty, we will also need to account for the $\text{cov}(\hat{R}_{t,t+j}^S, \hat{M}_{t,t+j})$ term.

Since in table 2 we find that $\alpha_{t,t+j}^S$ is consistently positive, we naturally also find that a_2 consistently implies that if anything the slope of SMU is *smaller* for below the strike price:

Table 4: Robust uncertainty bands

		K=0.9	K=0.95	K=1	K=1.05	K=1.1
a1	Estimate	-2.40	-2.40	-2.40	-2.40	-2.40
	CI (95%)	-3.54 -1.26	-3.54 -1.26	-3.54 -1.26	-3.54 -1.26	-3.54 -1.26
a2	Estimate	-2.99	-9.23	-14.15	-2.37	1.15
	CI (statistical, 95%)	-9.10 3.11	-13.41 -5.04	-18.27 -10.03	-7.14 2.41	-4.91 7.22
	CI (bias, corr=0.5)	-12.81 6.82	-14.95 -3.50	-19.92 -8.38	-6.59 1.86	-2.54 4.85
	CI (bias, corr=1)	-22.62 16.63	-20.67 2.22	-25.69 -2.61	-10.82 6.09	-6.24 8.54
	CI (both, corr=0.5)	-14.55 8.56	-16.32 -2.14	-21.24 -7.06	-8.74 4.01	-5.95 8.26

Note: Table reports the estimated coefficients a_1 (top panel) and a_2 (bottom panel) from equation (21), using data on synthetic puts with maturity of 21 days. Each column corresponds to a different strike K (as fraction of spot). The top panel also reports statistical 95% confidence intervals for a_1 . The bottom panel reports various confidence intervals for a_2 , accounting for statistical uncertainty (95% confidence level); accounting for half the maximum potential bias (i.e. assuming correlation between \hat{M} and \hat{R}^S of 0.5); accounting for the maximum potential bias (i.e. assuming correlation between \hat{M} and \hat{R}^S of 1); and accounting for both statistical uncertainty and half the maximum potential bias. Sample is 1926-2021.

SMU is estimated to be concave. In fact, that effect is so strong that in some cases it implies that SMU is actually increasing with the market return. That sort of “locally negative risk aversion” has also been found previously in studies of option returns, for example Rosenberg and Engle (2002).

The remaining rows in table 4 report various possible “uncertainty intervals” for the change in the slope above and below the strike. The first case just looks at the usual 95-percent confidence band, measuring statistical sampling uncertainty. In any case where the alpha is statistically significantly positive, the confidence band for the change in the slope will also only include negative values (i.e. implying concavity).

The next two rows incorporate potential effects of the $\text{cov}(\hat{R}_{t,t+j}^S, \hat{M}_{t,t+j})$ adjustment. We use the following simple bound,

$$\left| \text{cov}(\hat{R}_{t,t+j}^S, \hat{M}_{t,t+j}) \right| \leq \text{corr}(\hat{R}_{t,t+j}^S, \hat{M}_{t,t+j}) \text{std}(\hat{R}_{t,t+j}^S) \text{std}(\hat{M}_{t,t+j}) \quad (22)$$

For the correlation, we assume either $\left| \text{corr}(\hat{R}_{t,t+j}^S, \hat{M}_{t,t+j}) \right| < 0.5$, based on the evidence from the previous section, or the weakest possible bound of $\left| \text{corr}(\hat{R}_{t,t+j}^S, \hat{M}_{t,t+j}) \right| < 1$.

$\text{std}(\hat{R}_{t,t+j}^S)$ can be measured directly from the time series, where $\hat{R}_{t,t+j}^S$ is simply the residual from the kernel regression studied above.

Finally, to get $\text{std}(\hat{M}_{t,t+j})$ we assume that the volatility of the unspanned part of SMU is no greater than that from the part of SMU linearly spanned by the market. That is,

$std(\hat{M}_{t,t+j}) \leq E[R_{t,t+j}^m - 1]$. Intuitively, that restriction says that the Sharpe ratio available from any investment independent of the market return can be no greater than that of the market itself.

The next two rows report bounds for the change in the slope of SMU across the strike for the two bounds on $\left|corr(\hat{R}_{t,t+j}^S, \hat{M}_{t,t+j})\right|$. In general they continue to imply that we can rule out the possibility SMU is convex. And in any case, the uncertainty intervals themselves do not change the point estimates, which imply concavity.

Finally, the bottom row of table 4 constructs uncertainty intervals combining the statistical uncertainty – the usual standard error – with the uncertainty due to the pricing of unspanned risk. Rather than sum those two sources of uncertainty linearly, we take the square root of the sum of their squares, which can be thought of as capturing the idea that the bound on the price of unspanned risk is really encoding a ± 2 standard deviation range, like a prior.

5 The convergence of true and synthetic put returns

The theoretical analysis in section 2 shows that in theory options are the ideal security for estimating the shape of SMU. That fact is well known – with a full set of strikes, one can in principle measure the entire SMU curve with respect to the market return, getting measures of local risk aversion at each point (e.g. Rosenberg and Engle (2002)). Section 2 along with the introduction also discussed why options may not be the best vehicle for doing so in practice: they have a short sample, have not always been liquid, and are not easily traded by all investors. Trading frictions and market segmentation can potentially explain why true and synthetic options would have different returns over time (or, equivalently, why we would observe large gains on delta-hedged options).

But trading frictions and market segmentation have declined over time. Volume in the S&P 500 futures market, for example, is orders of magnitude larger than it was 30 years ago. Bid/ask spreads in futures are now on the order of 0.01 basis points. In addition, while endowments and pensions are legally not allowed to sell options in the US, the hedge fund sector has grown enormously – by an order of magnitude just between the years 2000 and 2007 – giving those investors a vehicle through which to get exposures that are highly similar to short options positions (Jurek and Stafford (2015)).

There are two basic explanations for why there might be a gap between returns on true and synthetic options: either the difference in their returns is a priced source of risk, or markets are segmented, so that they are not priced by the same SMU. If the latter is true,

then as segmentation has declined over time, we should also expect to see the gap between returns on true and synthetic options shrink. This section provides evidence on that point.

The top panel of Figure 5 plots 10-year moving averages of the returns from selling monthly 5% out-of-the-money true and synthetic puts, both market hedged (i.e. subtracting their beta times the market return). The solid blue line shows the well known result that selling puts was profitable for much of the sample, but those profits have actually declined significantly over time, to the point that following the crash in 2020 due to Covid, selling puts had a negative 10-year return for the first time.

The solid red line shows the behavior of synthetic puts. The returns have fluctuated around zero – sometimes positive, and sometimes negative, but with no consistent sign. The returns on the synthetic puts are, as one would expect from the results in the previous sections, highly correlated month-by-month with those on true puts, but the means are very different.

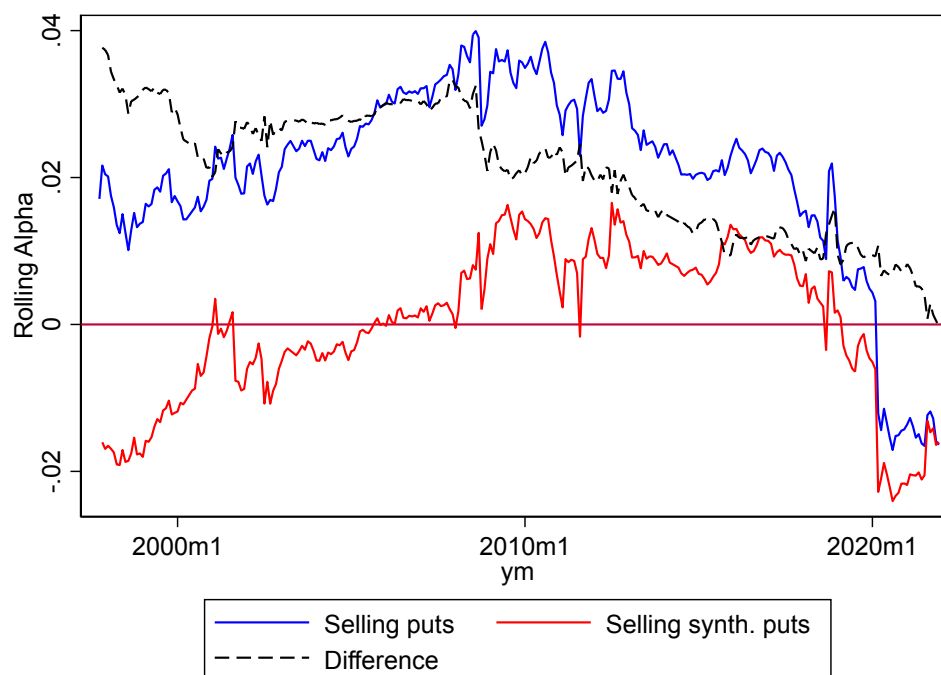
All of this can also be stated in terms of alphas. Since the returns are all market-hedged, rolling mean returns represent alphas over the same ten-year windows.

The convergence of the returns of true puts to those on synthetic puts is immediately clear from the figure. The black dashed line plots the difference between the two series, which is equivalently to the return on delta-hedged puts (e.g. Bakshi and Kapadia (2003)). That difference has fallen progressively through the sample, until by the last ten-year window it is almost exactly equal to zero.

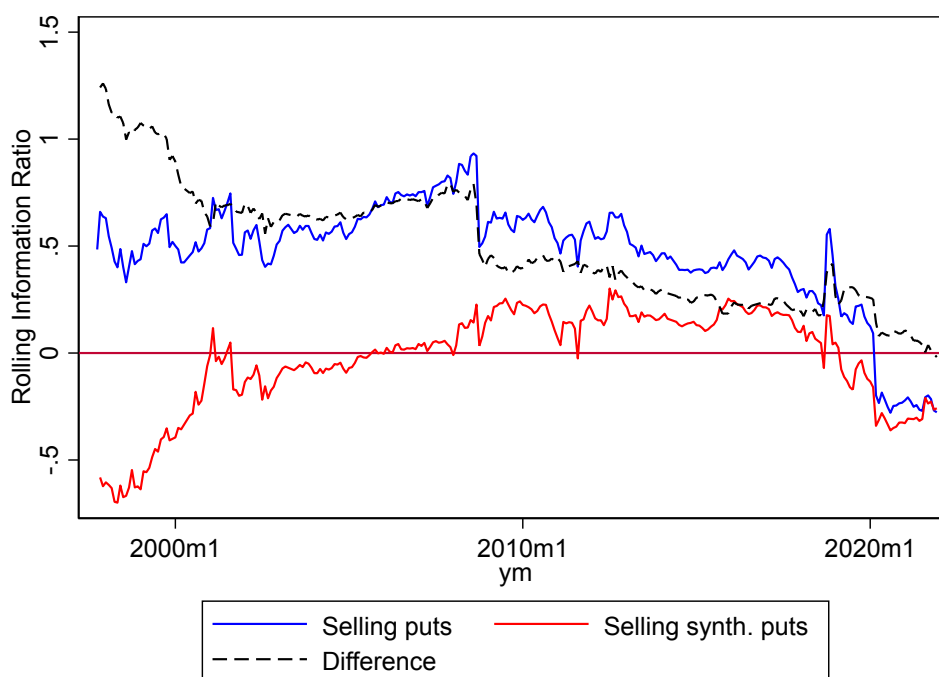
The bottom panel of figure 5 plots the Sharpe ratios of the series in the top panel (which are information ratios for the options themselves). The information ratios from selling exchange-traded puts are economically very large – about equal to the size of the market risk premium itself. That is, a mean-variance investor would be approximately indifferent between holding the market or writing market-hedged puts in the early part of the sample. By the end of the sample, the information ratios are approximately equal, consistent with the top panel, and there is no longer any evidence for outperformance.

We consider two formal tests of whether the gap between traded and synthetic option returns has shrunk over time. First, we simply regress the gap between the true and synthetic option returns on a time trend. Using Newey–West standard errors with 12 monthly lags, the coefficient on the time trend is -0.011 percent per month with a standard error of 0.004 percent (and hence significant at the 1-percent level). To put that value in context, the average value of the gap was 1.75 percent per year prior to 2000. -0.11 percent per month is enough to erase that over a period of about 13 years, consistent with what is observed in figure 5.

Figure 5: Moving average of alphas and information ratios for selling puts



(a) 10-year rolling alphas



(b) 10-year rolling information ratios

Note: Panel (a) shows the 10-year moving average of alphas to a strategy that sells puts, either traded or synthetic. The dashed line reports the difference of the two. Panel (b) reports the analogous results for information ratios.

Alternatively, we estimate a test for a structural break in the value of the gap between the two series under the null that it has a constant mean. Using the optimal exponential Wald statistic of Andrews and Ploberger (1994), we obtain a p-value of 0.0275, implying we can strongly reject the null of no break. The estimated break date (maximizing the Wald statistic) is January, 2008. This test trims the first and last 20 percent of the sample, so the value of the statistic is not affected by either the 2020 or 1987 crash.

Overall, then, this section shows that returns on exchange-listed options have converged to those on synthetic options. Over time the positive deviation of the listed puts has shrunk until by the end of the sample there is no significant difference between the two, and their alphas are approximately equal to zero. These results are consistent with the idea that in the earlier part of the sample, up to the mid-2000's, perhaps, exchange-traded options were segmented from the overall equity market and hence priced differently. As that segmentation has shrunk, the results from the two methods have converged and now both agree on the proposition that there is no particular premium for left tail risk.

6 Conclusion

The fact that options can reveal state prices is a foundational result in asset pricing, and it is well known that in the data equity index option prices imply that state prices are particularly high (relative to the associated physical probabilities) for states in which the market has significant declines. This paper takes an alternative approach to measuring the characteristics of state prices, showing that they can be recovered from the dynamics of stock market returns. The results of that method contrast starkly with those from options, with index returns implying that there is nothing particularly special about the left tail of the return distribution.

The core question one must ask in evaluating this paper's results, then, is which method is more trustworthy. Options have the advantage of giving a very direct measure, which requires only minimal assumptions, but they come with a relatively short data sample, and there is evidence that the options market is somewhat segmented from the broader equity market. Inference from the dynamics of market returns, on the other hand, requires somewhat more (though still fairly weak) assumptions, but comes with a sample three times longer than the options sample, and is one of the only markets (along with perhaps bonds) that essentially all investors participate in.

It is almost inevitable that there will be a difference in implications between the two methods, but standard models of intermediary constraints and market segmentation would

imply that as liquidity increases in the options market and it becomes better integrated and accessible, the returns in the options and equity markets should converge. We provide initial evidence that the convergence seems to be going in the direction of equities – the tail risk premium in options has been shrinking and approaching the values recovered from equity returns. One version of the question is, if delta hedging worked perfectly, would option returns converge to those implied by equity dynamics, or would equity dynamics converge to those implied by option returns?

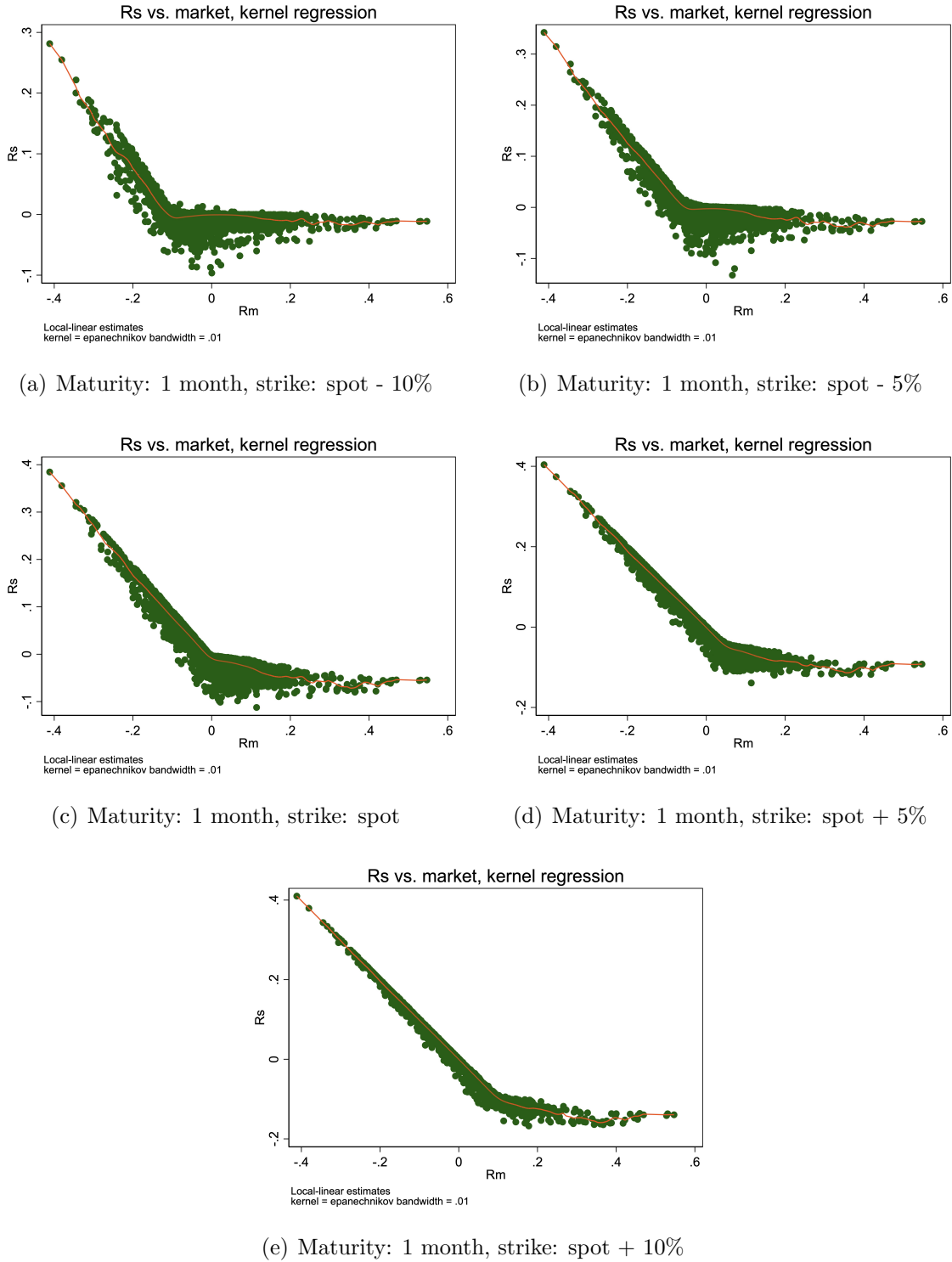
References

- Ait-Sahalia, Yacine and Andrew W Lo**, “Nonparametric risk management and implied risk aversion,” *Journal of econometrics*, 2000, *94* (1-2), 9–51.
- Andrews, Donald WK and Werner Ploberger**, “Optimal tests when a nuisance parameter is present only under the alternative,” *Econometrica: Journal of the Econometric Society*, 1994, pp. 1383–1414.
- Bacchetta, Philippe and Eric Van Wincoop**, “The great recession: a self-fulfilling global panic,” *American Economic Journal: Macroeconomics*, 2016, *8* (4), 177–98.
- Bakshi, Gurdip and Nikunj Kapadia**, “Delta-Hedge Gains and the Negative Market Volatility Risk Premium,” *The Review of Financial Studies*, 2003, *16*(2), 527–566.
- Barro, Robert J.**, “Rare Disasters and Asset Markets in the Twentieth Century,” *Quarterly Journal of Economics*, 2006, *121*(3), 823–866.
- Bates, David S**, “The market for crash risk,” *Journal of Economic Dynamics and Control*, 2008, *32* (7), 2291–2321.
- Beason, Tyler and David Schreindorfer**, “Dissecting the Equity Premium,” 2022. Working paper.
- Bollen, Nicolas PB and Robert E Whaley**, “Does net buying pressure affect the shape of implied volatility functions?,” *The Journal of Finance*, 2004, *59* (2), 711–753.
- Constantinides, George M, Jens Carsten Jackwerth, and Alexi Savov**, “The puzzle of index option returns,” *Review of Asset Pricing Studies*, 2013, *3* (2), 229–257.
- Coval, Joshua D. and Tyler Shumway**, “Expected Option Returns,” *The Journal of Finance*, 2001, *56*(3), 983–1009.

- Drechsler, Itamar and Amir Yaron**, “What’s Vol Got to Do with it?,” *The Review of Financial Studies*, 2011, *24*(1), 1–45.
- Farmer, Roger EA**, “The stock market crash of 2008 caused the Great Recession: Theory and evidence,” *Journal of Economic Dynamics and Control*, 2012, *36* (5), 693–707.
- Frazzini, Andrea and Lasse Heje Pedersen**, “Embedded leverage,” *The Review of Asset Pricing Studies*, 2022, *12* (1), 1–52.
- Gabaix, Xavier**, “Variable Rare Disasters: An Exactly Solved Framework for Ten Puzzles in Macro-Finance,” *Quarterly Journal of Economics*, 2012, *127*(2), 645–700.
- Garleanu, Nicolae, Lasse Heje Pedersen, and Allen M Poteshman**, “Demand-based option pricing,” *The Review of Financial Studies*, 2008, *22* (10), 4259–4299.
- Han, Bing**, “Investor sentiment and option prices,” *The Review of Financial Studies*, 2008, *21* (1), 387–414.
- Hansen, Lars Peter and Robert J Hodrick**, “Forward exchange rates as optimal predictors of future spot rates: An econometric analysis,” *Journal of Political Economy*, 1980, *88* (5), 829–853.
- Jackwerth, Jens Carsten**, “Recovering risk aversion from option prices and realized returns,” *The Review of Financial Studies*, 2000, *13* (2), 433–451.
- Jurek, Jakub W and Erik Stafford**, “The cost of capital for alternative investments,” *The Journal of Finance*, 2015, *70* (5), 2185–2226.
- Lewellen, Jonathan and Stefan Nagel**, “The Conditional CAPM does not Explain Asset-Pricing Anomalies,” *Journal of Financial Economics*, 2006, *82* (2), 289–314.
- Muravyev, Dmitriy and Xuechuan Charles Ni**, “Why do option returns change sign from day to night?,” *Journal of Financial Economics*, 2020, *136* (1), 219–238.
- Pan, Jun**, “The jump-risk premia implicit in options: Evidence from an integrated time-series study,” *Journal of financial economics*, 2002, *63* (1), 3–50.
- Petrosky-Nadeau, Nicolas, Lu Zhang, and Lars-Alexander Kuehn**, “Endogenous disasters,” *American Economic Review*, 2018, *108* (8), 2212–45.
- Rietz, Thomas A.**, “The Equity Risk Premium: A Solution,” *Journal of Monetary Economics*, 1988, *22*(1), 117–131.

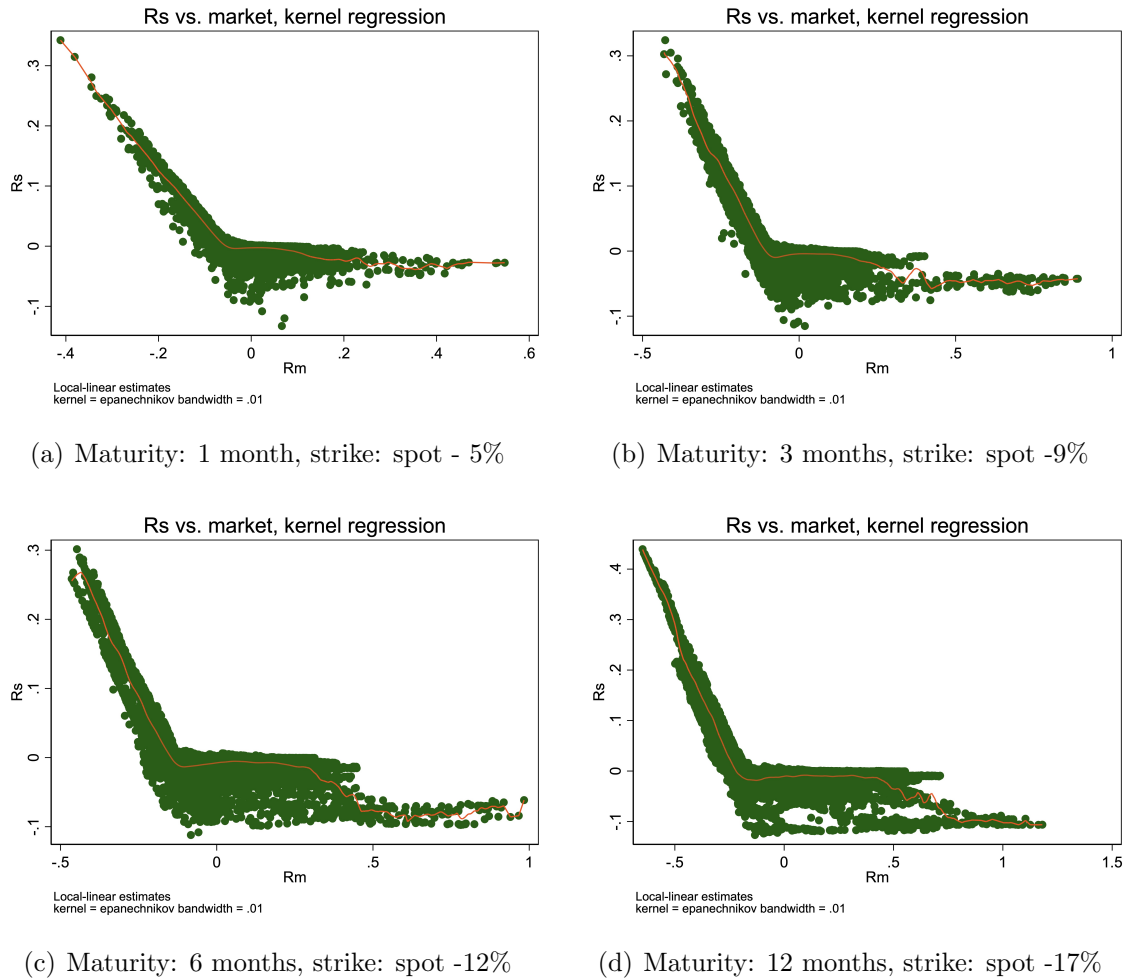
- Rosenberg, Joshua V and Robert F Engle**, “Empirical pricing kernels,” *Journal of Financial Economics*, 2002, *64* (3), 341–372.
- Schreindorfer, David**, “Macroeconomic tail risks and asset prices,” *The Review of Financial Studies*, 2020, *33* (8), 3541–3582.
- Seo, Sang Byung and Jessica A Wachter**, “Option prices in a model with stochastic disaster risk,” *Management Science*, 2019, *65* (8), 3449–3469.
- Welch, Ivo**, “The (time-varying) importance of disaster risk,” *Financial Analysts Journal*, 2016, *72* (5), 14–30.

Figure A.1: Synthetic put returns as a function of the market, various strikes



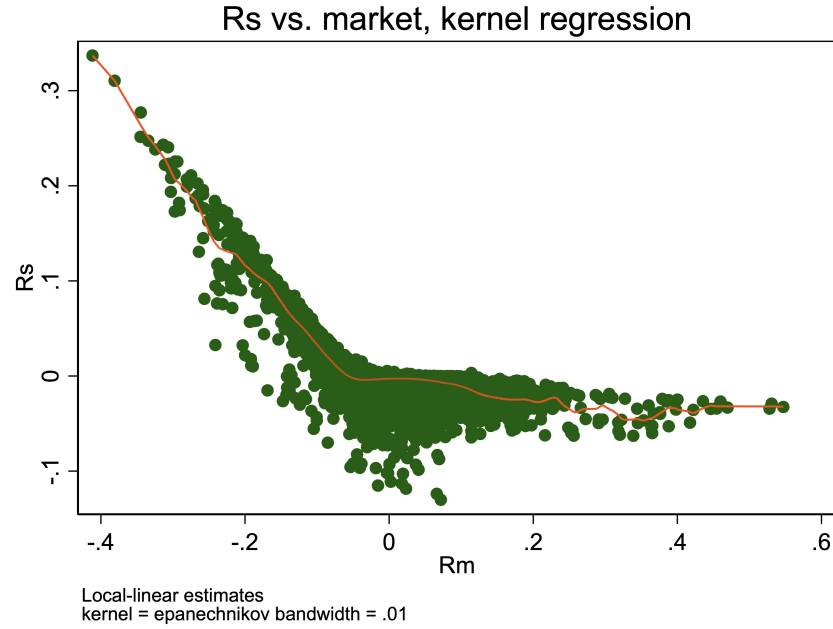
Note: The figure replicates the top panel of figure 2, using different strikes, from 10% below the spot to 10% above the spot.

Figure A.2: Synthetic put returns as a function of the market, various maturities

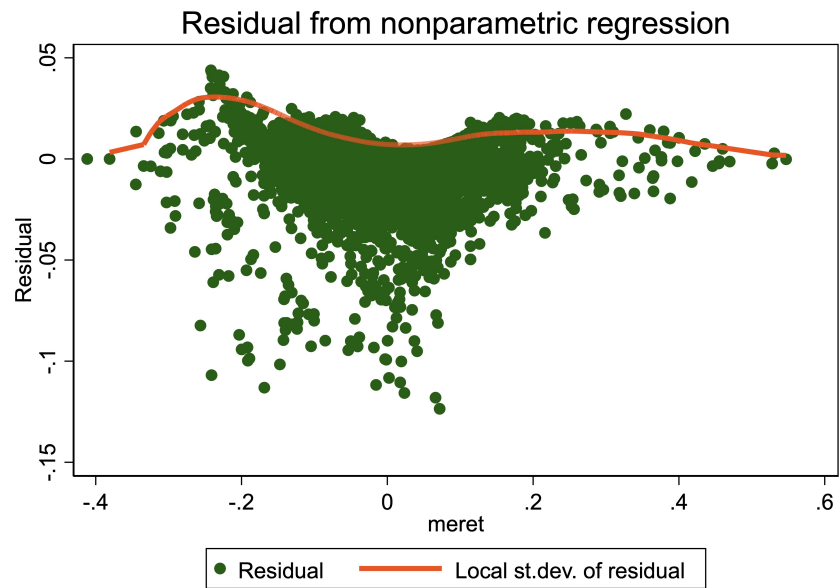


Note: The figure replicates the top panel of figure 2, using different maturities. The strike scales with the square root of time to maturity (as volatility does).

Figure A.3: Synthetic put returns as a function of the market



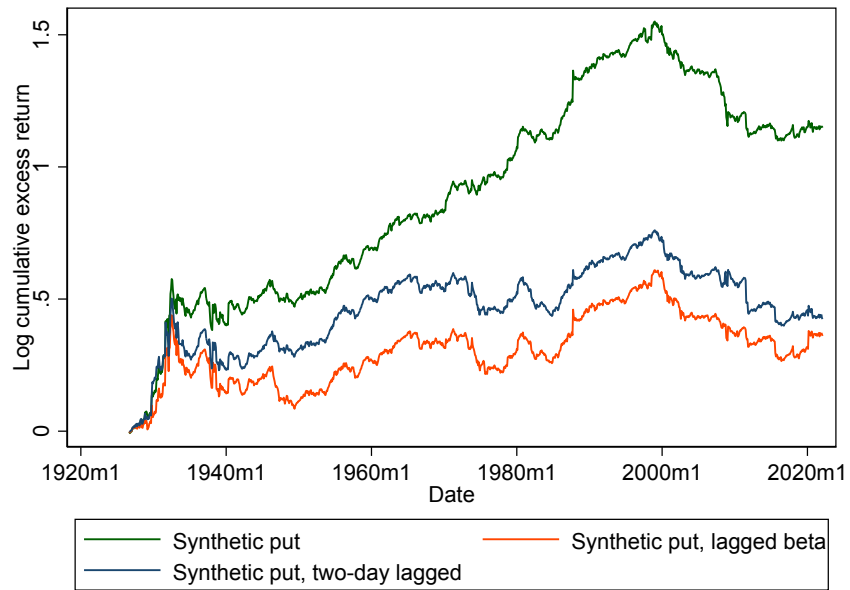
(a) R^S vs. R^m



(b) Residuals and their s.d.

Note: Same as figure 2, but using one-day lagged beta to build the synthetic put.

Figure A.4: Log cumulative returns on \tilde{R}^S



Note: Log cumulative returns on \tilde{R}^S , for the period 1926-2021. Green line is the synthetic put, red line is the synthetic put with one-day lagged beta, blue line is the synthetic put with two-day lagged beta.