

# Risk preferences implied by synthetic options

Ian Dew-Becker and Stefano Giglio\*

June 30, 2023

## Abstract

The historical returns on equity index options are well known to be strikingly negative. That is typically explained either by investors having convex marginal utility over stock returns (e.g. crash/variance aversion) or by intermediaries demanding a premium for hedging risk. This paper examines the consistency of those explanations with returns on dynamically replicated, or synthetic, options. Theoretically, it derives conditions under which convex marginal utility leads synthetic options to also have negative excess returns. Empirically, synthetic options have never earned significantly negative CAPM alphas, in stark contrast to exchange-traded options. Over the last 15 years, returns on true options have converged to those on synthetic options while various drivers of the cost and risk of hedging options exposures have shrunk, consistent with a model in which intermediaries drive option prices.

## 1 Introduction

### Background

A major empirical fact of financial markets is that equity index options have been overpriced historically relative to simple benchmark models. Investors who purchase options have, on average, earned significant negative returns and negative CAPM alphas.<sup>1</sup> As dis-

---

\*Dew-Becker: Northwestern University and NBER. Giglio: Yale University and NBER. We appreciate helpful comments from Gurdip Bakshi, Mike Chernov, Bjorn Eraker, Niels Gormsen, and seminar participants at ITAM, Yale, the University of Connecticut, Northwestern, Johns Hopkins, the University of Miami, INSEAD, HEC Paris, Berkeley, the Canadian Derivatives Institute, Boston University, the SFS Cavalcade, Booth, HBS, SAIF, Kansas University, and CFIRM.

<sup>1</sup>For analyses of returns, see Coval and Shumway (2001), Bakshi and Kapadia (2003), Broadie, Chernov, and Johannes (2007), Constantinides, Jackwerth and Savov (2013), Chambers, Foy, Liebner, and Lu (2014), Dew-Becker et al. (2017), and Muravyev and Ni (2020), among many others. For structural models, see Backus, Chernov, and Martin (2011), Drechsler and Yaron (2011), Gabaix (2012), Drechsler (2013), Seo and Wachter (2019), and Schreindorfer (2020). Note, though, that those models are almost exclusively endowment economies.

cussed in Bates (2022), there are two broad classes of explanations for the negative CAPM alpha of index options. The first is that marginal utility for some hypothetical representative investor is convex in market returns. Periods with large negative returns (and possibly also large positive returns) have state prices that are higher than would be expected just based on a model like the CAPM in which marginal utility rises linearly as the market drops. That can be due, for example, to an aversion to crashes or high volatility, or to time-varying risk aversion or other behavioral factors.<sup>2</sup> The second class of explanations focuses on intermediaries, explaining option overpricing as the result of intermediaries being net short options and charging a premium for their concentrated risk, e.g. Bollen and Whaley (2004) and Garleanu, Pedersen, and Poteshman (2009).

Option prices have many applications in finance, including measuring investor expectations of various moments of the conditional distribution of returns, investor preferences across market return states, and the drivers of risk premia (Bollerslev and Todorov (2011), Beason and Schreindorfer, (2022)).<sup>3</sup> They can also reveal potential amplification mechanisms for macroeconomic shocks (e.g., if investors become more risk averse when the market drops),<sup>4</sup> and are a key input in understanding the importance of stabilization policy – optimal policy depends on agents’ subjective valuations of different possible states of the world.<sup>5</sup>

In all of those cases, the preferences or beliefs recovered from options are implicitly assumed to be those of some typical or representative investor. But it is well known that option prices have puzzling implications that are difficult or impossible to represent with standard utility theory, for example sometimes implying *negative* risk aversion. One way to explain that is if the options market is segmented from broader financial markets.<sup>6</sup> Then option prices reveal the preferences of the specialist investors that trade in options markets, but *not* necessarily those of the typical equity investor.

---

<sup>2</sup>See Cuesdeanu and Jackwerth (2018) for an extensive review of such models and their empirical support, along with a discussion of models of intermediary frictions. On risk aversion, see Ait-Sahalia and Lo (2000), Jackwerth (2000), Rosenberg and Engle (2002), and Schreindorfer and Sichert (2022). On the variance risk premium, see, in addition to work cited elsewhere here, see Lamoreaux and Lastrapes (1993), Coval and Shumway (2001), Bakshi and Kapadia (2003), Du (2011), Dew-Becker et al. (2017), and Ait-Sahalia, Karaman, and Mancini (2020).

<sup>3</sup>On preferences over market states, see Ait-Sahalia and Lo (2000), Jackwerth (2000), and Rosenberg and Engle (2002). For conditional moments, see the CBOE VIX index and, among many others, Carr and Madan (1998), Carr and Wu (2009), and Martin (2017). Options have also been used to measure jump risk (e.g. Bollerslev and Todorov (2014)), micro uncertainty (Dew-Becker and Giglio (2020)), and option implied skewness (Kozhan, Neuberger, and Schneider (2013), Dew-Becker (2022)).

<sup>4</sup>E.g. He and Krishnamurthy (2013), Hall (2017) , and Muir (2017).

<sup>5</sup>Alvarez and Jermann (2004) is an example of how asset prices can be used to measure the cost of fluctuations. De Paoli and Zabczyk (2013) study optimal policy under time-varying risk aversion.

<sup>6</sup>See Jackwerth (2000), Bollen and Whaley (2004), Bates (2008), Han (2008), Garleanu, Pedersen, and Poteshman (2008), Jurek and Stafford (2015) , Haddad and Muir (2021), Frazzini and Pedersen (2022), among many others

## Contribution

This paper develops a novel approach to measuring the average investor’s risk preferences by studying *synthetic* options – dynamic portfolios that attempt to replicate returns on traded options from returns on more liquid investments. The paper first establishes conditions under which recovery of preferences from synthetic options is possible, and then measures returns on synthetic options over nearly a century of data, compares them to traded option returns, and examines how well the theoretical requirements hold in the data. To interpret the results and disentangle the preference- and intermediary-based theories of option pricing, the paper extends the model of Garleanu, Pedersen, and Poteshman (2009) to allow for multiple frictions in intermediary hedging and examines how the gap in returns between true and synthetic options relates to the magnitude of the frictions in the data.

Even without frictions, there are always models that can rationalize any deviation in the pricing of traded and synthetic options. At a high level, though, this paper’s contribution is simply to ask whether any conclusions are changed if one focuses on synthetic options – whose pricing depends only on behavior of equity prices – instead of traded options, whose pricing depends on a separate derivatives market.

## Methods

The paper first shows how the CAPM alpha of a traded option on the stock market measures curvature in marginal utility with respect to (i.e. projected onto) the market return. Since options have payoffs that are convex in the market return, if marginal utility is also convex, options have relatively high prices and consequently low returns (and in fact negative CAPM alphas). The more important theoretical contribution, though, is to give circumstances under which convexity in marginal utility also implies negative CAPM alphas for *synthetic* options.

Synthetic options are constructed by taking positions in the underlying to match the exposure (i.e. the delta) that a true traded option would have. The analysis formalizes three key conditions that are needed for convex marginal utility to imply negative alphas on synthetic options. The two most important are that the equity market index is correctly priced every day – which one might question in periods when trading frictions were more severe – and that the part of synthetic option returns that is unspanned (nonlinearly) by the market return is not priced.<sup>7</sup> The latter condition is where, for example, the pricing of jump risk would appear, and it very plausibly could be violated. On the other hand, notably, it is not required that investors are necessarily able to implement the synthesis – the theory just requires certain pricing conditions (though they certainly might be related to trading

---

<sup>7</sup>Throughout the paper, the term “span” is used in the Hilbert space sense of a conditional expectation.

frictions).

Empirically, the paper examines monthly returns on synthetic and traded options. The synthetic options can be constructed back to 1926 using data on the CRSP market return, while traded option returns are available since August, 1987, when monthly options were first introduced. The analysis measures CAPM alphas for synthetic and traded options and then examines how well the three required conditions for synthetic options hold.

## Results

The first empirical question is how well replication works. At a first pass, it appears to work well: synthetic options have returns that are over 90-percent correlated with traded option returns, hedge realized crashes in the data well, and are strongly convex in market returns, as required by the theoretical analysis. That does not mean that options could have been synthesized in real time historically, though – trading costs and other frictions would have made that infeasible.

The next question is what alphas the synthetic options earn. Unlike traded options, synthetic options have CAPM alphas that are zero or even positive, and the confidence bands are economically narrow. That result is robust over time, across strikes, across maturities, and to modifying various details in the construction.

While synthetic options have consistently earned near-zero CAPM alphas, the alphas of traded options have *not* been so consistent (see also Bates (2022)). According to various methods, there is a break in the returns somewhere around about 2010. In the period since 2010, in fact, the alphas of the traded options have converged to zero, consistent with the synthetic options.

The final section of the paper asks what might have caused such a shift. One might take the view that synthetic options, because they were not really feasible to construct for most of the sample (e.g. due to the lack of index futures), were not a good way to measure investor preferences. In that case, as those frictions decline and daily rebalancing becomes feasible, the returns on synthetic options would converge (down) to those on traded options. But that is not what occurred empirically.

A similar problem arises if one takes the view that the synthetic and traded options hedge different states. Synthetic option returns depend on the path the market takes, so the gap between true and synthetic returns is a function of realized volatility and other higher-order factors like jumps. If that gap is related to marginal utility, then it will be priced and drive a wedge between the traded and synthetic returns.<sup>8</sup> But those factors do not appear to have shrunk over time. The volatility of the gap between traded and synthetic option returns

---

<sup>8</sup>No-arbitrage option prices typically rely on this type of mechanism, e.g. Pan (2002).

has been stable, jump variation shows no trend, and skewness in market returns has become significantly *more* negative.<sup>9</sup>

An alternative explanation of the results is that the stock and options markets are segmented. We extend the segmented-markets model of Garleanu, Pedersen, and Poteshman (2009) to allow for trading costs and an index-futures basis. In the model, the alpha of synthetic options is exogenous, but that of traded options depends on the magnitude of hedging frictions that intermediaries face. When they shrink, the alpha does also. Empirically, both trading costs and risk due to the index-futures basis have declined over time. The index-futures basis is particularly notable. During the 1987 crash, S&P 500 futures were at times 20% underpriced relative to the level of the S&P 500 index, which would represent an enormous cost to a dealer with net long options positions who must sell futures as the underlying falls. In the crash in the fall of 2008, though, the basis was never more than about 5% and was centered on zero. One explanation for the decline in the option premium soon after the crash, then, is that intermediaries learned that they faced less hedging risk during crashes than they had thought based on past data.

### **Interpretation**

The paper's basic empirical results are simple to state: synthetic options bought on the aggregate stock market have never earned significantly negative returns, while traded options used to, but do not any longer. Under certain circumstances, the results for the synthetic options imply that marginal utility is not convex (and may even be concave) in the market return. And when looking at traded options, the evidence for convexity at the monthly frequency has shrunk or even disappeared. That is consistent with the option premium having been driven by intermediary frictions, which have declined. The findings have three implications.

First, significant care must be taken when using option prices or returns to estimate or test models. The importance of dealer dynamics means that equity and options are not frictionlessly integrated, which is often assumed in structural models. Second, and relatedly, the paper's results imply that derivatives prices, up until relatively recently, were distorted away from those implied by the preferences of whoever is the typical investor pricing equities. In thinking about calibration and estimation of structural models, then, one must decide whose preferences exactly are being modeled. As one alternative path forward, this paper's results could be taken as suggesting the use of synthetic options for calibration and estimation. Finally, the analysis finds clear evidence of nonstationarity. So when studying option returns, attention must be paid to the exact sample being used and how the results

---

<sup>9</sup>Instability of parameters in no-arbitrage models is a well understood issue in the option pricing literature. See, for example, Bollen and Whaley (2004) and Andersen, Fusari, and Todorov (2015).

may have changed over time. Financial markets have gone through many massive changes over decades and centuries. An interesting fact here is that through all of that, the returns implied by option replication have been largely stable, while there are much larger changes in derivatives markets.

### Outline

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.

## 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  $M_{t,t+j}$  (i.e.  $u'_{t+j}/u'_t$ , where  $u$  is utility over consumption).<sup>10</sup> Since  $M_{t,t+j}$  is a ratio of marginal utilities, we immediately have that  $M_{t,t+2} = M_{t,t+1}M_{t+1,t+2}$ , etc. Since any deviations of investors' subjective probability measure from the truth can also be incorporated into  $M_{t,t+j}$ , we refer to it more generally as *SMU*.

**Definition 1** *We say that a return  $R_{t,t+j}$  is priced by subjective marginal utility (SMU), denoted  $M_{t,t+j}$ , over  $t \rightarrow t + j$  if*

$$1 = E_t [M_{t,t+j}R_{t,t+j}] \tag{1}$$

where  $E_t$  is the expectation operator under the physical probability measure conditional on information available on date  $t$ .

Note that this definition may potentially only hold for certain  $t$  and  $j$  – i.e. only on some dates or over just some horizons. The definition does not imply that markets are complete, meaning that marginal utility need not be identical across agents. Rather, all agents must just agree on equation (1) for whichever assets are (universally) priced by marginal utility. In addition, it does not require rational expectations – irrational beliefs can be accommodated by  $M_{t,t+j}$  as long as they satisfy basic axioms for probability measures.

---

<sup>10</sup>If agents have biased probability measures, then the bias will also be part of  $M_{t,t+j}$ . We would in that case just require that there are internally consistent in that  $M_{t,t+2} = M_{t,t+1}M_{t+1,t+2}$ .

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.

The paper's goal is to understand how marginal utility varies with the state of the equity market. To that end, define, for an arbitrary variable  $X_{t,t+j}$ ,

$$\bar{X}_{t,t+j} \equiv E[X_{t,t+j} | R_{t,t+j}^m], \quad (2)$$

$$\text{and } \hat{X}_{t,t+j} = X_{t,t+j} - \bar{X}_{t,t+j} \quad (3)$$

$\bar{X}_{t,t+j}$  is the component of  $X_{t,t+j}$  that can be written as a (potentially nonlinear) function of the market return.  $\hat{X}_{t,t+j}$  is then the residual, which is uncorrelated with both  $\bar{X}_{t,t+j}$  and  $R_{t,t+j}^m$  by construction. We say  $\bar{X}_{t,t+j}$  is the **spanned** part and  $\hat{X}_{t,t+j}$  the **unspanned** part (spanning here is in the sense of Hilbert spaces for random variables).

For  $M_{t,t+j}$  the decomposition is into a part related to the total market return and a residual.  $\hat{M}_{t,t+j}$  does not affect the pricing of the market portfolio, and is not correlated with *any* function of the market return, so it also cannot affect the pricing of traded options, as shown in the next section. It can, though, affect the pricing of synthetic options and other securities.

$\bar{M}_{t,t+j}$  is the paper's primary object of interest – how marginal utility depends on the market return. It is what the past literature on option-implied pricing kernels has focused on, since, as the next section shows, it is what options carry information about.

## 2.2 Interpreting traded option returns

Define the gross return on some arbitrary option on the market (or even a portfolio of options) to be  $R_{t,t+j}^O$ . The part of that return (linearly) correlated with the market can always be subtracted, and we have

$$R_{t,t+j}^{O\perp} \equiv R_{t,t+j}^O - \frac{\text{cov}_t(R_{t,t+j}^O, R_{t,t+j}^m)}{\text{var}_t(R_{t,t+j}^O, R_{t,t+j}^m)} (R_{t,t+j}^m - 1) \quad (4)$$

$$\alpha_{t,t+j}^O = E_t [R_{t,t+j}^{O\perp} - 1] \quad (5)$$

where  $\alpha_{t,t+j}^O$  is the CAPM alpha of  $R_{t,t+j}^O$ . Note that  $R_{t,t+j}^{O\perp}$  is not a delta-hedged return; rather, one might say it is beta-hedged, where the hedge is a *fixed* position in the underlying. The hedge is conditional on date- $t$  information. Since it just adds a static position in the market,  $R_{t,t+j}^{O\perp}$  has the usual kinked relationship with  $R_{t,t+j}^m$ , just tilted compared to  $R_{t,t+j}^O$ .

**Proposition 2** *If  $R_{t,t+j}^O$  and  $R_{t,t+j}^m$  are priced over  $t \rightarrow t + j$ , then  $\bar{M}_{t,t+j}$  has the representation,*

$$\bar{M}_{t,t+j} = \text{const.} - \frac{E_t [R_{t,t+j}^m - 1]}{\text{var}_t [R_{t,t+j}^m]} R_{t,t+j}^m - \frac{\alpha_{t,t+j}^O}{\text{var}_t (R_{t,t+j}^{O\perp})} R_{t,t+j}^{O\perp} + \text{resid.} \quad (6)$$

where the residual term is orthogonal to  $R_{t,t+j}^m$  and  $R_{t,t+j}^{O\perp}$ .

Equation (6) is a regression of  $\bar{M}_{t,t+j}$  on two functions of the market return: a linear term ( $R_{t,t+j}^m$ ) and a nonlinear term ( $R_{t,t+j}^{O\perp}$ , which is, conditionally, an exact nonlinear function of the market return). The result comes from the fact that the covariances in the numerators of the regression coefficients can be replaced here by the two risk premia – e.g.  $E_t [R_{t,t+j}^m - 1] = -\text{cov}_t (R_{t,t+j}^m, M_{t,t+j})$ .

$\frac{-E_t [R_{t,t+j}^m - 1]}{\text{var}_t [R_{t,t+j}^m]}$  therefore measures the average slope of marginal utility with respect to the market return.  $R_{t,t+j}^{O\perp}$  is a piecewise linear function of the market return, so its coefficient,  $\frac{-\alpha_{t,t+j}^O}{\text{var}_t (R_{t,t+j}^{O\perp})}$ , measures how the slope of  $\bar{M}_{t,t+j}$  changes across the strike.

The top panel of Figure 1 illustrates that idea, plotting SMU, normalized to equal 1 for  $R_{t,t+j}^m = 1$ , relative to the market return. Under the CAPM (the gray line), SMU is linear in the market return, all alphas are zero, there is no convexity, and the slope is recovered simply as  $\frac{-E_t [R_{t,t+j}^m - 1]}{\text{var}_t [R_{t,t+j}^m]}$ .

The dashed blue line plots the SMU implied by the alphas observed for 5% out-of-the-money listed S&P 500 puts between 1987 and 2021. Historical put returns imply that effective risk aversion – as measured by the slope of SMU – is significantly higher when the market falls. The non-monotonicity here is a typical, if surprising, empirical finding.

## 2.3 Interpreting synthetic option returns

It is well known that option returns can be approximated through dynamic trading in the underlying asset – the market return in this case. This section gives conditions under which such synthetic returns yield information about nonlinearity in  $\bar{M}_{t,t+j}$ .

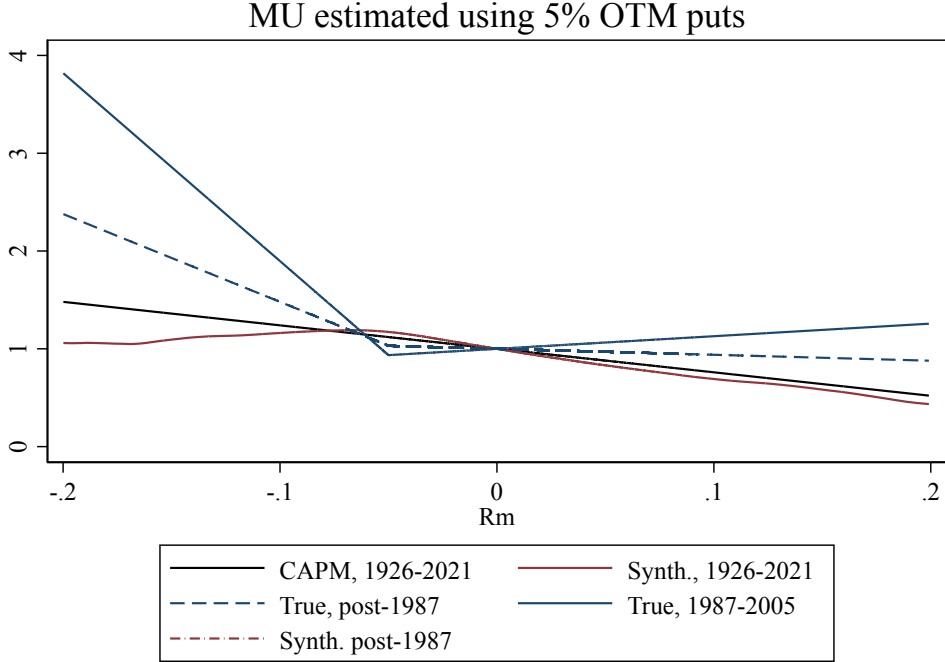
Consider the following gross return on a synthetic option from  $t$  to  $t + j$ ,

$$R_{t,t+j}^S \equiv \sum_{s=t}^{t+j-1} \delta_s^S (R_{s,s+1}^m - 1) + 1 \quad (7)$$

where  $\delta_t^S$  is a set of weights that depend on information available on date  $t$  (in the Black–Scholes (1973) replication, they are exactly the option deltas and depend on the level of the market index and its volatility). Note that in general  $R_{t,t+j}^S \neq R_{t,t+j}^O$  and the replication will



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.

not be perfect (it is not hard to construct examples in which it is, in fact, useless).

**Proposition 3** *If  $R_{t,t+1}^m$  is priced by SMU for all  $s \rightarrow s + 1$  for  $t \leq s < t + j$ , then*

$$\bar{M}_{t,t+j} = \text{const.} - \frac{E_t [R_{t,t+j}^m - 1]}{\text{var}_t [R_{t,t+j}^m]} R_{t,t+j}^m - \frac{\left( \alpha_{t,t+j}^S + \text{cov}_t \left( \hat{M}_{t,t+j}, \hat{R}_{t,t+j}^S \right) \right)}{\text{var}_t (R_{t,t+j}^{S\perp})} \bar{R}_{t,t+j}^{S\perp} + \text{resid.} \quad (8)$$

where  $\alpha_{t,t+j}^S$  is the CAPM alpha of  $R_{t,t+j}^S$  and the residual is orthogonal to  $R_{t,t+j}^m$  and  $\bar{R}_{t,t+j}^{S\perp}$ .

We again have an expression for SMU in terms of two returns, with coefficients depending on their risk premia. Proposition 3 gives three conditions under which  $\alpha_{t,t+j}^S$  can be used to measure convexity in SMU:

1.  $\bar{R}_{t,t+j}^S$  is a convex function of the market return
2.  $R_{t,t+1}^m$  is priced by SMU for all  $t \rightarrow t + 1$
3.  $\text{cov}_t \left( \hat{M}_{t,t+j}, \hat{R}_{t,t+j}^S \right)$  is zero

The first condition just says that synthetic options have returns that are convex in the market, which can be checked empirically. It will be violated if, for example, large declines in the market are driven by intraday jumps.

The second and third conditions are harder to evaluate because they are statements about SMU, which is not directly observable. The pricing condition might fail at the daily level if there are frictions allowing market prices to deviate from their fundamental values at high frequency (see section 4.4).

The third condition is the hardest to evaluate. It requires that the unspanned part of synthetic option returns not be priced. Section 4.5 discusses it extensively, both looking at what variables  $\hat{R}_{t,t+j}^S$  is correlated with and also using the method of Cochrane and Saa-Requejo (2000) to bound the covariance. Condition 3 holds in any model where the stock market return is a sufficient statistic for SMU (so that  $\hat{M}_{t,t+j} = 0$ ).<sup>11</sup> If there are other state variables that affect marginal utility and are independent of market returns (e.g. perhaps unspanned volatility), they will appear in  $\hat{M}_{t,t+j}$ . It also holds when nonlinear payoffs on the market can be replicated via dynamic trading, e.g. when the market follows a binomial tree or in continuous time (with continuous hedging) when the market is a single-dimensional diffusion, so that  $\hat{R}_{t,t+j}^S = 0$ . In reality neither of those conditions holds literally, and the question becomes how large the bias from  $\text{cov}(\hat{M}_{t,t+j}, \hat{R}_{t,t+j}^S)$  is.

### 2.3.1 Feasibility

A surprising feature of the three conditions for interpreting synthetic option returns is that none of them directly requires that options synthesis be feasible in practice. The requirement that the market be priced correctly every day, for example, does not say that every investor can trade every day. In many models, the market is correctly priced even though there is no trade in equilibrium.

There is no question that for the vast majority investors over the vast majority of the empirical sample, replicating options dynamically would have been expensive and time-consuming. And even for professionals it may not have worked at critical moments, like the 1987 crash. While that fact is not directly relevant in the analysis so far, it is central to intermediary-based models, which are analyzed in section 5. In such models, the frictions lead there to be different SMU for different agents, since they do not have access to the same investments. For example, if the typical equity investor cannot freely trade options,

---

<sup>11</sup>As examples, this would hold if marginal utility is a function of current wealth and that wealth returns are perfectly correlated with stock market returns, or if investors have Epstein–Zin preferences in the limit where the intertemporal elasticity of substitution is infinite and the stock market is equivalent to a claim on consumption (potentially levered).

then their SMU may not price options, violating the requirement of proposition 2. And if hedging options is expensive for intermediaries, that drives a wedge between traded option prices and the prices implied by the marginal utility of equity investors.

## 3 Data and methods

### 3.1 Option synthesis

Throughout the analysis,  $t$  is taken to be a day. The weight  $\delta_t^S$  is equal to the delta of the option – the partial derivative of the value with respect to the price of the underlying. Different models give different exact expressions for delta. The main analysis uses a method from Hull and White (2017) that corrects the Black–Scholes delta for the leverage effect.

Section 4.4 discusses potential effects that market microstructure biases can have on estimated mean returns and alphas. There is evidence that stale prices affect the results in the earlier part of the sample (see also Bates (2012)). The biases can be reduced or eliminated by choosing the time-series weights  $\delta_t^S$  based only on information lagged by a day, and the main results use that method.<sup>12</sup> That makes the estimates of mean returns robust, but it is also conservative in terms of fit – delta hedging is less effective when using stale information.

The return volatility needed to calculate delta is obtained from a heterogeneous autoregressive model (Corsi (2009)) that estimates expected 1-month volatility as a function of past two-week volatility and the past three months of volatility (with the lags chosen based on the Bayesian information criterion). The model is estimated on an expanding window, so that when the delta is computed only past information is used to estimate the model and forecast volatility. Robustness to the various choices here is examined in section 4.7.

The market return is measured as the (daily) CRSP value-weighted stock market return and the risk-free rate is the one-month Treasury bill rate from Kenneth French’s website.

### 3.2 Traded options

The dataset for traded options splices together data on CME futures options for the period 1987–1995 with CBOE SPX options from Optionmetrics for the period 1996–2021. Following Broadie, Chernov, and Johannes (2009), we study a monthly rolling strategy, where options

---

<sup>12</sup>That is, in all of the benchmark results  $\delta_t^S$  – the investment in the market on date  $t + 1$  – is set based on information available only at date  $t - 1$ .

are purchased on the third Friday of every month and then held to their maturity on the following month’s third Friday.

In parts of the analysis that involve direct comparisons of synthetic and traded options, we align the returns – comparing returns over the same third Friday to third Friday period. However, when looking at univariate statistics, the analysis uses 21-day overlapping windows for the synthetic options to maximize statistical power (since there is no need to only use a single return per month).

For both the synthetic and traded options, excess returns are scaled with the price of the underlying in the denominator, rather than the price of the option, as in Büchner and Kelly (2022). The scaling is the return perceived by an investor who is buying options in proportion to the underlying. It is a payoff per unit of insurance, rather than per unit of the insurance premium that is paid. See appendix B for further discussion.

## 4 Empirical results

### 4.1 The relationship between $R_{t,t+j}^S$ and $R_{t,t+j}^m$

The top panel of Figure 4.1 plots  $R_{t,t+j}^S$  for put options against  $R_{t,t+j}^m$ , where  $j = 21$  days and the strike used to construct  $\delta_t^S$  is 95% of the initial level of the market (corresponding to approximately a unit standard deviation decline). The plot for call options with the same strike is identical but rotated 45 degrees counterclockwise.

There is clearly significant nonlinearity – for values of the market return above the strike 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 constructed to mimic a put option. The red line plots the nonparametric estimates of  $\bar{R}_{t,t+j}^S \equiv E [R_{t,t+j}^S | R_{t,t+j}^m]$ .<sup>13</sup> They formally quantify the relevant nonlinearity, showing that  $\bar{R}_{t,t+j}^S$  is close to piecewise linear in the market return.

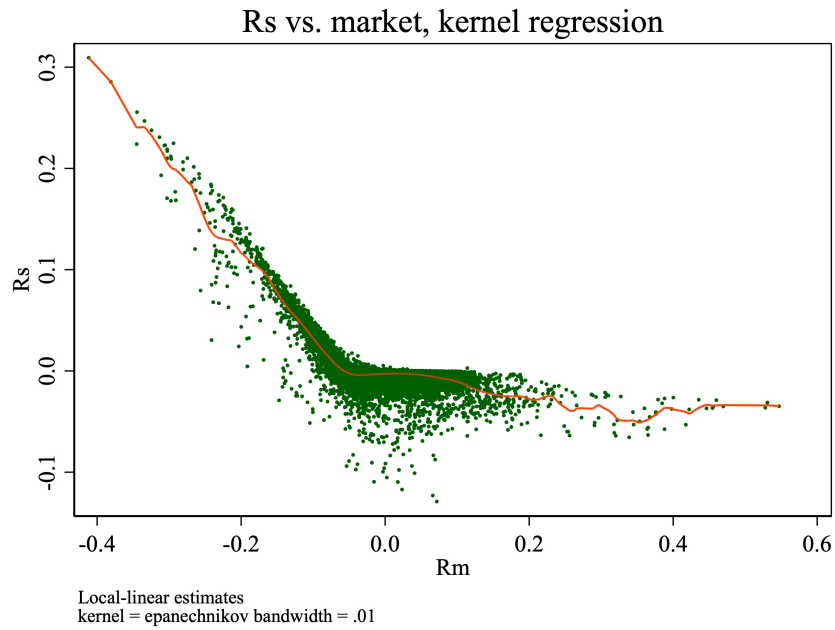
Importantly,  $\bar{R}_{t,t+j}^S$  rises with a consistent slope as  $R_{t,t+j}^m$  falls, regardless of how large the decline is. If it was not possible to span large declines in the market with time-varying weights, e.g. due to large jumps,  $\bar{R}_{t,t+j}^S$  would flatten out for the very negative values of  $R_{t,t+j}^m$ . Figure A.1 replicates figure 4.1 across strikes and shows the results are highly similar.

The table below reports the most extreme 21-day returns in the six most extreme events in the US stock market in our sample. As a benchmark, note that because the returns are meant to replicate 5% OTM puts, they would ideally generate a return that is 5% lower than the negative of the market return, minus the initial cost of the option. In all six cases,

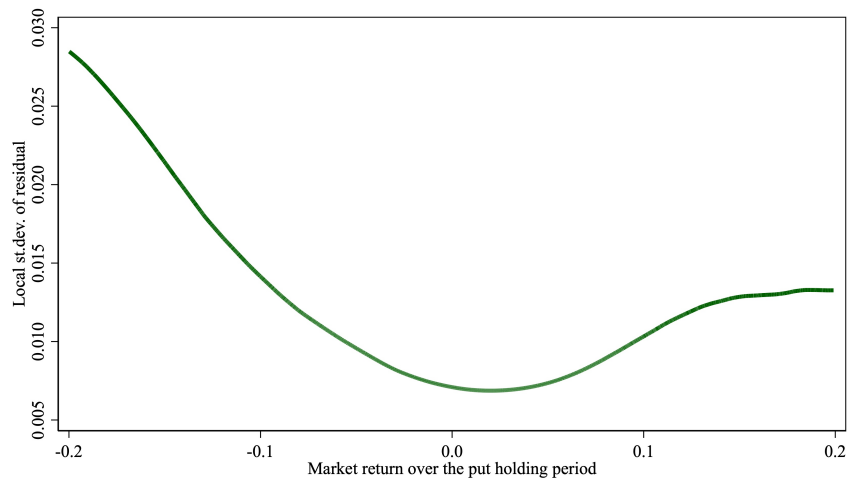
---

<sup>13</sup>Conditional expectations are calculated via a local linear regression on  $R_{t,t+j}^m$  with a Gaussian kernel and the bandwidth set to 0.01.

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



(a)  $R^S$  vs.  $R^m$



(b) Local standard deviation of residuals ( $\hat{R}_{t,t+j}^S$ )

**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.

the synthetic puts have highly positive returns, providing economically meaningful insurance against these crashes.

**Returns on the market and synthetic puts, six most extreme events**

	$R_{t,t+j}^m$	$\bar{R}_{t,t+j}^S$	$\bar{R}_{t,t+j}^S$ (no lag in $\delta_t^S$ )	Ideal return
Nov. 1929	-41%	31%	32%	35%
Mar. 2020	-33%	25%	25%	28%
Oct. 2008	-31%	22%	22%	24%
Oct. 1987	-30%	17%	23%	24%
Oct. 1931	-29%	22%	22%	24%

Recall that in the benchmark results,  $\delta_t^S$  depends on data only up to date  $t - 1$  in order to avoid microstructure biases. The third column in the table shows that when  $\delta_t^S$  uses date- $t$  information, the hedge becomes noticeably better, particularly in 1987. This shows that using lagged information for hedging is conservative for fit.

Even though the synthetic options fit well, the claim is not that the replication was implementable. The results just show that the hypothetical returns are nonlinear in the market, so that the first required condition from section 2.3 is satisfied empirically.

To begin to evaluate the third condition from section 2.3, that the residual risk is small, the bottom panel of figure 4.1 plots the conditional standard deviation of the residuals  $\hat{R}_{t,t+j}^S \equiv R_{t,t+j}^S - \bar{R}_{t,t+j}^S$ .<sup>14</sup> That standard deviation is always less than 3%, and in many cases less than 1%, especially for periods where the market return is near zero. In addition, we have the following variance decomposition:

$$\underbrace{\text{var} \left( R_{t,t+j}^S \right)}_{3.1 \times 10^{-4}} = \underbrace{\text{var} \left( \bar{R}_{t,t+j}^S \right)}_{2.5 \times 10^{-4}} + \underbrace{\text{var} \left( \hat{R}_{t,t+j}^S \right)}_{0.6 \times 10^{-4}} \quad (10)$$

About 21% of the variation in synthetic option returns in this case are unspanned by the market return. The more important question is whether that variation is priced, which we return to in section 4.5.

The amount of residual risk measured here is again a conservative estimate due to the fact that  $\delta_t^S$  is constructed using only lagged information. If  $\delta_t^S$  uses date- $t$  information, the fraction of the variance of  $R_{t,t+j}^S$  from  $\hat{R}_{t,t+j}^S$  falls to 15%.

Finally, to directly compare synthetic and traded option returns, figure A.2 compares

<sup>14</sup>The local volatility is estimated from a second kernel regression,

$$\eta_{t,t+j}^2 = h \left( R_{t,t+j}^m \right) + \text{residual} \quad (9)$$

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.

standard deviations and betas for traded and synthetic options across strikes and finds they are highly similar. Figure A.6 plots synthetic against traded option returns for different strikes and reports pairwise correlations, which range between 0.89 and 1.00.

## 4.2 Risk premia

### 4.2.1 Varying strikes at the monthly maturity

Figure 3 reports the paper’s key results for long-run average option risk premia. It includes results for three different periods: the full sample available for synthetic returns (1926–2021), the full sample available for both synthetic and traded returns (1987–2021), and the 1987–2005 sample used by Broadie, Chernov, and Johannes (BCJ; 2009), who report an extensive analysis of the performance of traded options. In all cases, we report results for put options. For alphas and information ratios, results for puts and calls are guaranteed to be identical for synthetic options at a given strike, and they are highly similar for traded options (due to put-call parity).

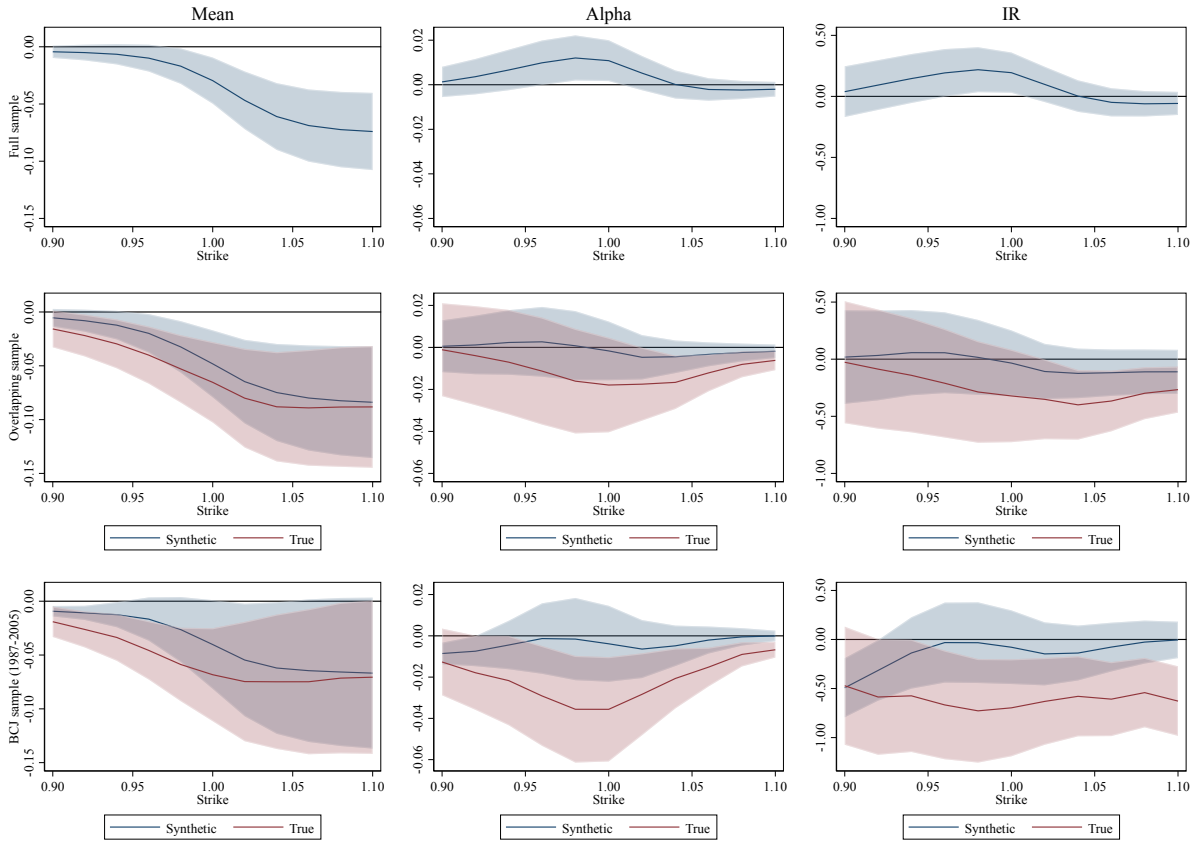
The left column of Figure 3 plots average returns. In all three panels, returns decline as the strike rises, which is to be expected as the betas also become more negative for higher strikes. In the periods of overlap between synthetic and traded options, traded options always have lower average returns than synthetic.

The middle column reports CAPM alphas, which are the paper’s key object of interest based on the theoretical analysis. The bottom two panels show that the estimated alphas of traded options are negative across all strikes in both the 1987–2021 and BCJ samples. The statistical evidence for the alphas being negative is stronger in the earlier BCJ sample, with the magnitudes falling by half in the longer sample.

In the same post-1987 period, though, and regardless of whether the post-2005 period is included, synthetic options have estimated alphas very close to zero, with no evidence of mispricing relative to the CAPM. The top panel shows that the same result holds in the full sample and is if anything actually stronger – in that case the alphas are statistically significantly *positive* for the middle strikes. The paper’s claim that synthetic options have been fairly priced historically relative to the CAPM is based on the results reported here for alpha.

It is also important to note how narrow the confidence band on alpha is in the top panel. The sample for synthetic options is about three times longer than for traded options, and the returns are available in overlapping windows, with the result that the standard errors are on average half as large in the full sample of synthetic options compared to the traded

Figure 3: Average option returns across strikes



**Note:** Cumulative log returns for returns on true and synthetic -5% puts orthogonalized with respect to the market (i.e. beta hedged). The lines are constructed to equal zero in July, 1987, when the true put options become available.

options.

Finally, the right-hand column of Figure 3 plots information ratios – the Sharpe ratio of the part of option returns uncorrelated with the market. Again, they are if anything statistically significantly *positive* in the full sample, peaking at values of about +0.2. The traded options, on the other hand, have information ratios as negative as -0.75 in the BCJ sample, which is larger than the Sharpe ratio of the overall stock market.

Figure A.3 replicates 3, but varying the maturity instead of the strike price, while figure A.4 scales the moneyness in volatility units. The results are similar to the baseline in both cases.



## 4.2.2 Cumulative alphas and variation in risk premia over time

The top panel of figure 4 plots cumulative CAPM alphas for synthetic options over the period 1926–2021 and for traded put options over the period 1987–2021 (again, for alpha the choice of put versus call is irrelevant for synthetic options; and it has only very minor effects for traded options). For readability, the cumulative returns are normalized to zero in July 1987 when the data for the traded options begins.

For synthetic options, the figure reinforces the result that over the full sample the alphas have been slightly positive. Covid jumps out in 2020 as a large positive innovation, due to the significant decline in the level of the stock market. The fact that the synthetic portfolio captures that gives clear evidence that it is able to capture economically significant large declines in the market. The bottom panel of figure 4 plots information ratios over rolling 10-year windows and again shows that the returns on synthetic options have been stable, with a brief period in the 1930’s when the returns were statistically significantly positive. In no period were they significantly negative.<sup>15</sup>

Traded put returns have two striking features. On the one hand, the month-to-month variation appears very similar to that for the synthetic options, consistent with the results presented so far. On the other hand, the average return is drastically different. The returns are highly negative, especially in the period up to 2010. The returns were roughly flat from then until the large market decline with Covid. In fact, the overall cumulative return on true puts is zero between February, 2008 and the end of the sample in December, 2021. The bottom panel of figure 4 shows how, over ten-year windows, the return on traded puts actually turned positive at the end of the sample.

Since a synthetic put is essentially a delta hedge, the difference between the returns on the traded and synthetic put returns is the return on a delta-hedged put, which is a measure of the variance risk premium (Bakshi and Kapadia (2003)). The figure therefore suggests that the alpha of the variance risk premium may have fallen to zero since 2010. Section 5.1 revisits this point in more detail and reports formal tests for a structural break.

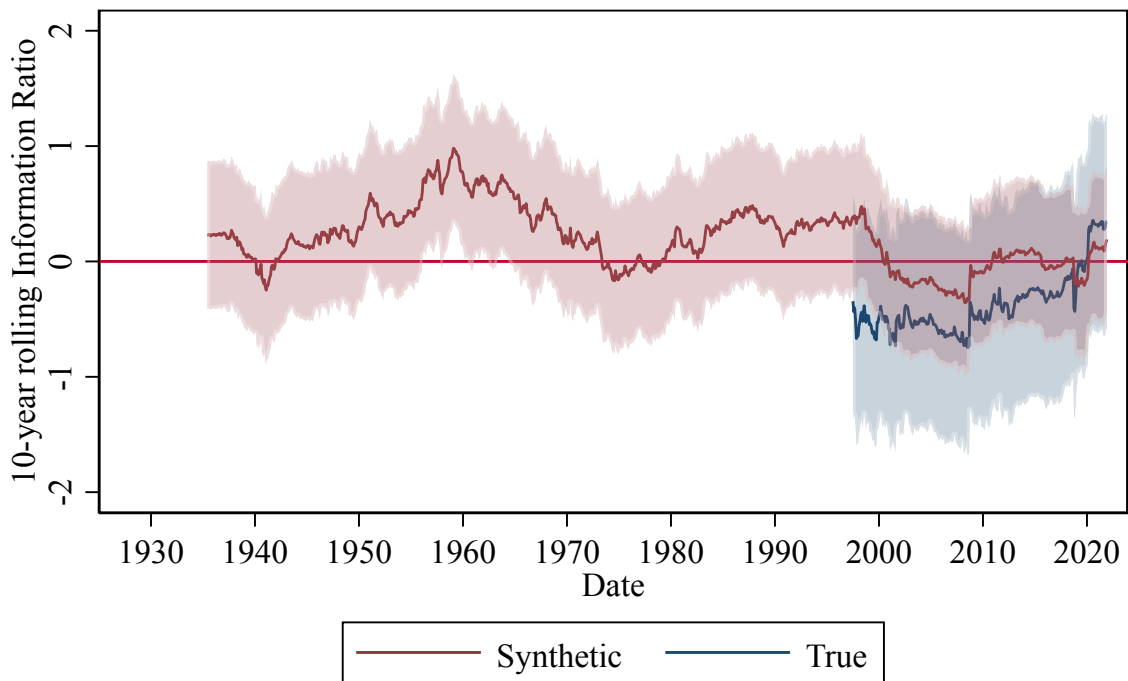
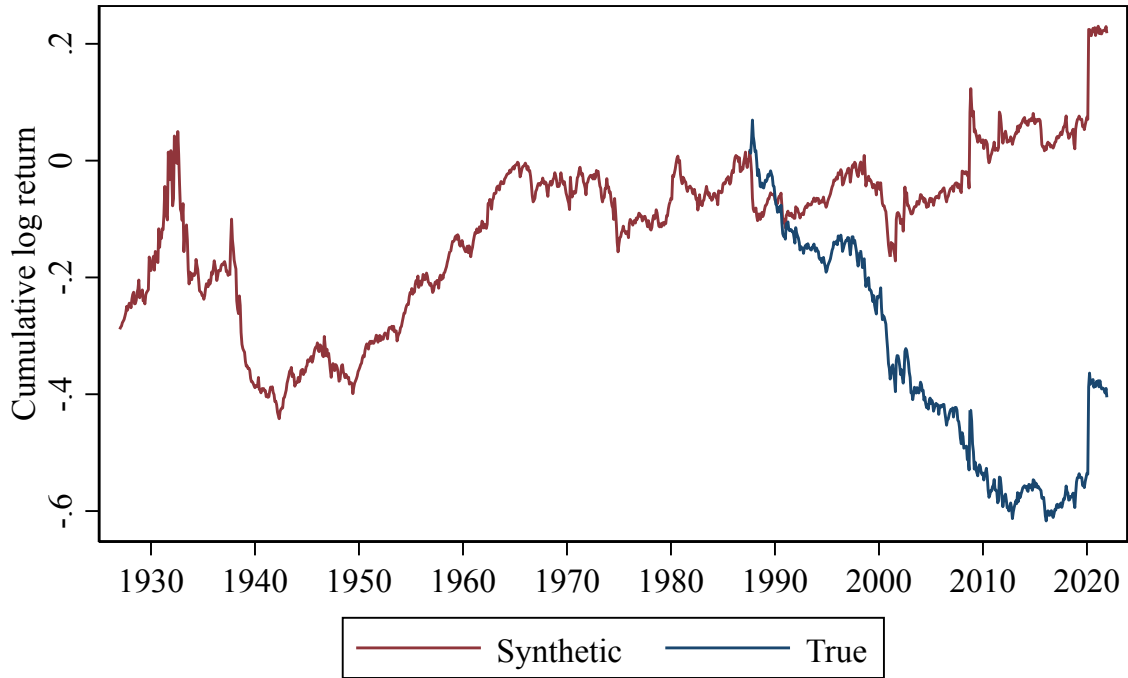
## 4.2.3 A conditional CAPM interpretation

Since synthetic options are created by trading the market dynamically, any CAPM alpha they earn has to come from timing the market. Their daily alpha is identically zero. For both synthetic puts and calls, the investment in the market,  $\delta_t^S$ , declines – becoming more

---

<sup>15</sup>The information ratios in the bottom panel deviate slightly from what is observed in the top panel because the top panel uses the monthly return series to construct cumulative returns, while the bottom panel uses the overlapping returns for synthetic options to maximize power, as discussed above.

Figure 4: Cumulative excess returns



**Note:** Cumulative log returns for returns on true and synthetic -5% puts orthogonalized with respect to the market (i.e. beta hedged). The lines are constructed to equal zero in July, 1987, when the true put options become available.

negative for a put and less positive for a call – when the market declines and rises when the market rises. Synthetic options are therefore bets on momentum. To get a negative alpha (which would be consistent with traded option returns) would require mean reversion in returns.

Formally, one can derive from results in Lewellen and Nagel (2006) that

$$\alpha_{t,t+j}^S \approx \text{cov} \left( \delta_t^S, [E_t [R_{t,t+1}^m] - E [R_{t,t+1}^m]] - \frac{E [R_{t,t+1}^m - 1]}{\text{var} (R_{t,t+1}^m)} [\text{var}_t (R_{t,t+1}^m) - \text{var} (R_{t,t+1}^m)] \right) \quad (11)$$

The first part of the covariance is the usual conditional CAPM intuition, which says that if  $\delta_t^S$  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  $\delta_t^S$  with conditional volatility – the movement of deltas with volatility is second-order (and its sign is ambiguous), so this term is relatively small quantitatively. 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.<sup>16</sup>

One intuition for these results is the following. If marginal utility is convex, then a drop in the market moves the agent to a steeper part of the SMU function, causing agents to demand higher expected returns. Returns then mean-revert somewhat on average, inducing a negative alpha for synthetic puts. Economically, in such a situation puts – even synthetic ones – are expensive because they insure against states agents are highly averse to. So the dynamics of returns and the alpha of synthetic puts can reveal the shape of the SMU function.

### 4.3 Effects of conditioning on betas

In the theoretical analysis, all alphas and betas are conditional, and hence potentially time-varying. Since option returns are non-linear functions of the market return, conditional betas of both true and synthetic options will necessarily change over time as the conditional distribution of the market return varies. That naturally then affects estimation of alphas.

Figure A.5 examines possible ways of accounting for time-variation in conditional betas.

---

<sup>16</sup>How 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  $E_t [R_{t,t+1}^m - 1] \propto \text{var}_t (R_{t,t+1}^m)$ , which would imply that the covariance in (11) is identically zero. If risk aversion is countercyclical, as with a convex pricing kernel, then even if volatility rises when the market falls,  $E_t [R_{t,t+1}^m - 1]$  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 (11) implies that  $\alpha_{t,t+j}^S$  is negative.

The left-hand column of panels plots the baseline results. In the middle column, betas are estimated from a rolling three-month window. The right-hand panels model conditional beta as a function of lagged (i.e., end of previous month) variables: the market return volatility forecast (which is most important), the market return itself, industrial production growth, and the corporate bond default spread. In both cases, the results are highly similar to the benchmark qualitatively and quantitatively.

#### 4.4 Effects of daily mispricing

Recall the second condition for using synthetic options to measure curvature in marginal utility from section 2.3 that the market is priced correctly every day. Obviously no such condition is literally true, so the question is what sort of pricing errors would create a bias in the results.

To see how the pricing condition could be violated, suppose that the “true” market return that satisfies  $1 = E_t [M_{t,t+1} R_{t,t+1}^m]$  every day,  $R_{t,t+1}^m$ , is unobservable and instead we can only see some contaminated return  $R_{t,t+1}^{m*}$ , with

$$R_{t,t+1}^{m*} = R_{t,t+1}^m + \varepsilon_{t+1} \quad (12)$$

where  $\varepsilon_{t+1}$  is the contamination. Depending on the properties of  $\varepsilon_{t+1}$  (e.g. its mean or predictability), it may be that  $1 \neq E_t [M_{t,t+1} R_{t,t+1}^{m*}]$  even if  $R_{t,t+1}^m$  is in fact priced each day.

If we define  $R_{t,t+j}^{S*}$  to be the return on an option synthesized from the contaminated market return,  $R_{t,t+1}^{m*}$ , then we have

$$E [R_{t,t+j}^{S*}] = E [R_{t,t+j}^S] + \sum_{s=0}^{j-1} E [\delta_{t+s}^S \varepsilon_{t+s+1}] \quad (13)$$

There are two conditions that must hold in order for contamination to not affect risk premia: the contamination,  $\varepsilon_t$ , must have zero mean and be uncorrelated with past values of the weights  $\delta_t^S$ . On the other hand, the errors need not be i.i.d., for example, or necessarily independent of anything else.

Recall that in the analysis above  $\delta_t^S$  is set based on information about returns only up to date  $t - 1$ . That choice is made to address the potential bias identified in this section. In particular, if  $\varepsilon_t$  is an MA(1) process, so that  $E [\varepsilon_t \varepsilon_{t-k}] = 0 \forall k > 1$ , then choosing  $\delta_t^S$  based on information from date  $t - 1$  is enough to ensure that  $E [\delta_t^S \varepsilon_{t+1}] = 0$ .

As discussed in Bates (2012), observed positive serial correlation in daily market index

returns is evidence of stale prices.<sup>17</sup> When there is a positive return in the underlying, the hedge weight  $\delta_t^S$  rises, both for puts and calls. That leads to the result that  $\text{cov}(\varepsilon_{t+1}, \delta_t^S) > 0$  when  $\varepsilon_t$  is positively serially correlated, which would lead to a positive bias,  $E[R_{t,t+j}^S] > E[R_{t,t+j}^{S*}]$ . That is why the main results use lagged information, so that  $\delta_t^S$  does not depend on  $R_t^{m*}$ .<sup>18</sup>

To examine this effect in the data, the top panel of figure 5 reports the autocorrelations of daily returns in the full sample and pre- and post-1973 separately. In the pre-1973 sample, there is clear evidence of one-day positive serial correlation, consistent with the presence of stale prices. The two-day autocorrelation, on the other hand, is negative, which is one reason why the analysis only uses a single-day lag in constructing the weights.

To see the effects of different choices for the information set for  $\delta_t^S$ , the bottom panel of figure 5 plots three versions of the cumulative alphas for synthetic options: with the baseline one-day lag, with no lag, and with a two-day lag. Switching from the baseline to no lag causes a large increase in alphas, entirely due to the positive autocorrelation, which is why it shows up only in the first half of the sample. Going from a one-day to a two-day lag also increases the alphas (by a lower amount). None of the lag choices leads to negative average alphas.

Another type of error that can arise is bid-ask bounce, which, as discussed in Jegadeesh and Titman (1995), creates negative serial correlation in returns. In contrast to stale prices, that would bias the estimate of synthetic option returns *downward*, implying the true alphas for synthetic options are even more positive and in even stronger conflict with the alphas on traded options than is suggested by the baseline results.

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

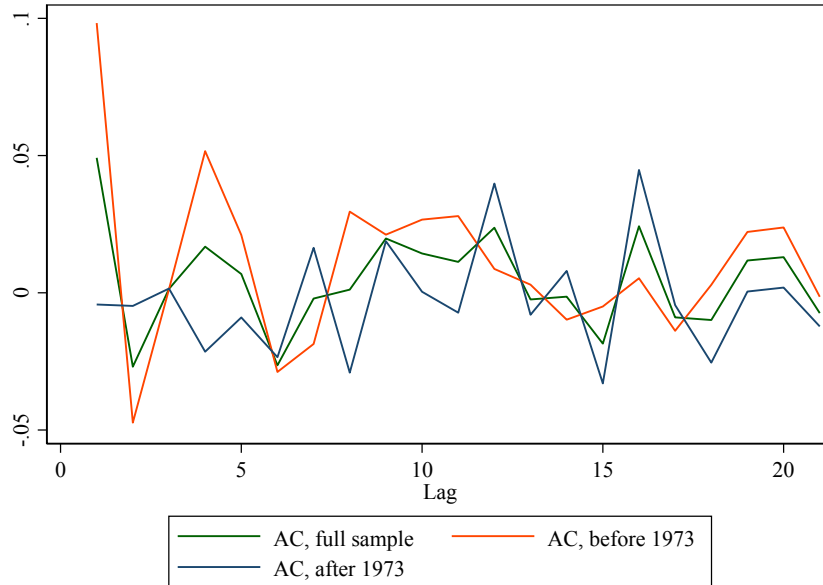
Recall from section 2.3 that the curvature of marginal utility is determined by  $\alpha_t^S + \text{cov}_t(\hat{M}_{t,t+j}, \hat{R}_{t,t+j}^S)$ . So far the analysis has ignored the covariance term. The questions are whether that covariance is zero and if not, how large it might be. This section examines what risk factors  $\hat{R}_{t,t+j}^S$  might be correlated with and then examines a bound on the magnitude of

---

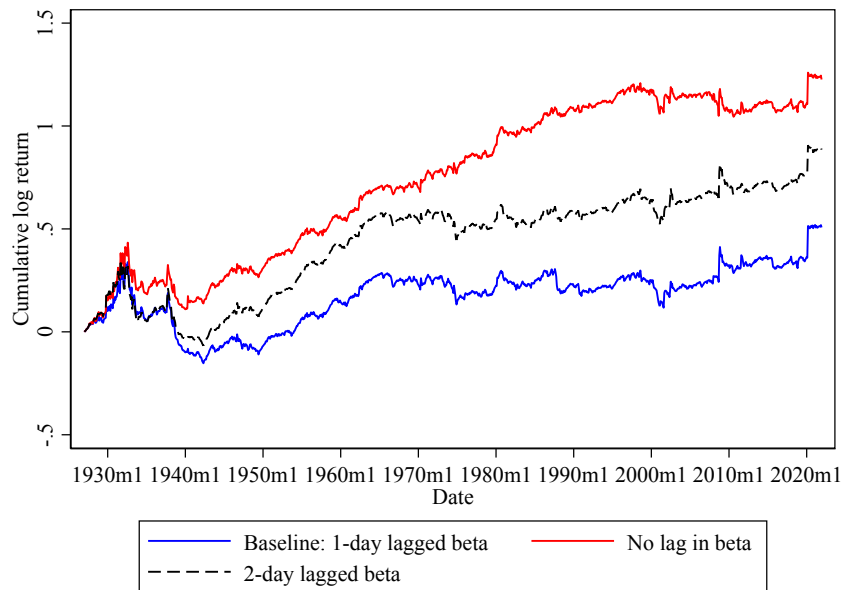
<sup>17</sup>For example, suppose on each day 50% of stock prices are updated. If there is good news on date  $t$ , that will be impounded into half of stock prices on date  $t$ , 1/4 on  $t+1$ , 1/8 on  $t+2$ , etc., inducing positive serial correlation in the “errors”  $\varepsilon_t$  relative to the “true” market return that would be observed if all stock prices were updated every day.

<sup>18</sup>One might also ask about the effect of mispricing on estimated betas. First, as an empirical matter, recall that the data shows that the betas of synthetic and traded options are highly similar, implying that there is not a severe bias. Second, while stale prices and bid-ask bounce can affect betas at high frequencies, those effects tend to shrink at lower frequencies, hence the more common focus on, say, monthly data in studies of equity returns.

Figure 5: Daily autocorrelation of returns



(a) Autocorrelations



(b) Cumulated returns of synthetic options

**Note:** Panel (a) shows the autocorrelations of daily returns up to 21 lags, using the full sample (1926-2021), and using the pre- and post-1973 data separately. Panel (b) shows the cumulative returns on the synthetic puts built using one-day lagged weights, no lag, or two-day lagged weights.

$\text{cov}_t \left( \hat{M}_{t,t+j}, \hat{R}_{t,t+j}^S \right)$  using the method of Cochrane and Saa-Requejo (2000).

#### 4.5.1 Relationship of $\hat{R}_{t,t+j}^S$ with risk factors

Table 1 reports correlations between  $\hat{R}_{t,t+j}^S$  and statistical innovations in prominent macro and financial variables.<sup>19</sup> The table thus measures the extent to which shocks to the unspanned part of returns are correlated with shocks to other variables.

Table 1: Correlation of residuals with macro variables

	All data	Excluding 2020
Unemployment	-0.13	-0.06
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	0.49	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 is true for all the other time series considered here.

<sup>19</sup>For  $\hat{R}_{t,t+j}^S$  we take the return from the beginning to the end of a month. For the variables that are measured at a fixed point in time, 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, we take the statistical innovation in the monthly value relative to data available in the previous month. The different time series are available for different time periods and each correlation is computed using the longest period available for that variable.

Among the macro time series, the correlations are all economically small and statistically insignificant. The only notable correlations are for price series: the excess 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 proposition 3 shows would imply more convexity in SMU than implied by  $\alpha_{t,t+j}^S$ .

The results suggest that options do not look risky to an investor whose marginal utility depends on either the level of the stock market or to macroeconomic variables. They do look risky, though, to an investor who cares on the path that the market return takes, suggesting that intermediaries might be relevant for pricing.

The bottom row of table 1 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 about 0.5, so we take that as an upper end for a reasonable estimate of the correlation of  $\hat{R}_{t,t+j}^S$  with  $\hat{M}_{t,t+j}$ , but we also examine results with the correlation set to 1.

#### 4.5.2 Robust uncertainty bands

If  $\text{cov}_t(\hat{M}_{t,t+j}, \hat{R}_{t,t+j}^S)$  is not equal to zero, how large of a bias does it create in measuring curvature in marginal utility? To bound the magnitude of  $\text{cov}_t(\hat{M}_{t,t+j}, \hat{R}_{t,t+j}^S)$ , start from the identity,

$$\left| \text{cov}_t(\hat{R}_{t,t+j}^S, \hat{M}_{t,t+j}) \right| = \left| \text{corr}_t(\hat{R}_{t,t+j}^S, \hat{M}_{t,t+j}) \right| \text{std}_t(\hat{R}_{t,t+j}^S) \text{std}_t(\hat{M}_{t,t+j}) \quad (14)$$

$\text{std}(\hat{R}_{t,t+j}^S)$  can be estimated based on the empirical time-series of  $\hat{R}_{t,t+j}^S$ .  $\left| \text{corr}(\hat{R}_{t,t+j}^S, \hat{M}_{t,t+j}) \right|$  is not observable, but the results in the previous section imply 0.5 as an estimate, and the upper bound is 1. Finally, to get  $\text{std}(\hat{M}_{t,t+j})$  we assume that the volatility of the unspanned part of SMU,  $\hat{M}_{t,t+j}$ , is no greater than that from the part of SMU spanned by the market,  $\bar{M}_{t,t+j}$ . That implies that

$$\text{std}(\hat{M}_{t,t+j}) \leq E[R_{t,t+j}^m - 1] / \text{std}(R_{t,t+j}^m) \quad (15)$$

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, similar to Cochrane and Saa-Requejo (2000) (see also references therein).<sup>20</sup>

<sup>20</sup>The 1 here represents the gross-risk-free rate. Again, in the empirical analysis the actual risk-free rate is used.



The parameter of interest, which measures convexity in SMU, is

$$\alpha_{t,t+j}^{S,adjusted} \equiv \alpha_{t,t+j}^S + \text{cov}_t \left( \hat{M}_{t,t+j}, \hat{R}_{t,t+j}^S \right) \quad (16)$$

It has two sources of uncertainty: estimation uncertainty in  $\alpha_{t,t+j}^S$  and the unobservable value of  $\text{cov}_t \left( \hat{M}_{t,t+j}, \hat{R}_{t,t+j}^S \right)$ . Appendix C shows how those two sources of uncertainty can be combined geometrically, essentially treating  $\text{cov} \left( \hat{R}_{t,t+j}^S, \hat{M}_{t,t+j} \right)$  as another Gaussian source of error.

Figure 6 reports an alternative version of figure 4 that now incorporates these robust uncertainty bands instead of the original confidence bands based only on estimation error. The left-hand panels assume that  $\text{corr} \left( \hat{R}_{t,t+j}^S, \hat{M}_{t,t+j} \right)$  based on the results in the previous section, while the right-hand panels use the most conservative possible value of 1.

The uncertainty bands in figure 6 are guaranteed to be wider than in the baseline. However, they can still reject information ratios of -0.5 in all but a few cases with the shortest sample. In the top-left panel, which is the most powerful case, using the longest sample and  $\text{corr} \left( \hat{R}_{t,t+j}^S, \hat{M}_{t,t+j} \right) = 0.5$ , the bound can reject even small negative information ratios. So even when accounting for unspanned risk, the curvature of SMU implied by synthetic options remains small.

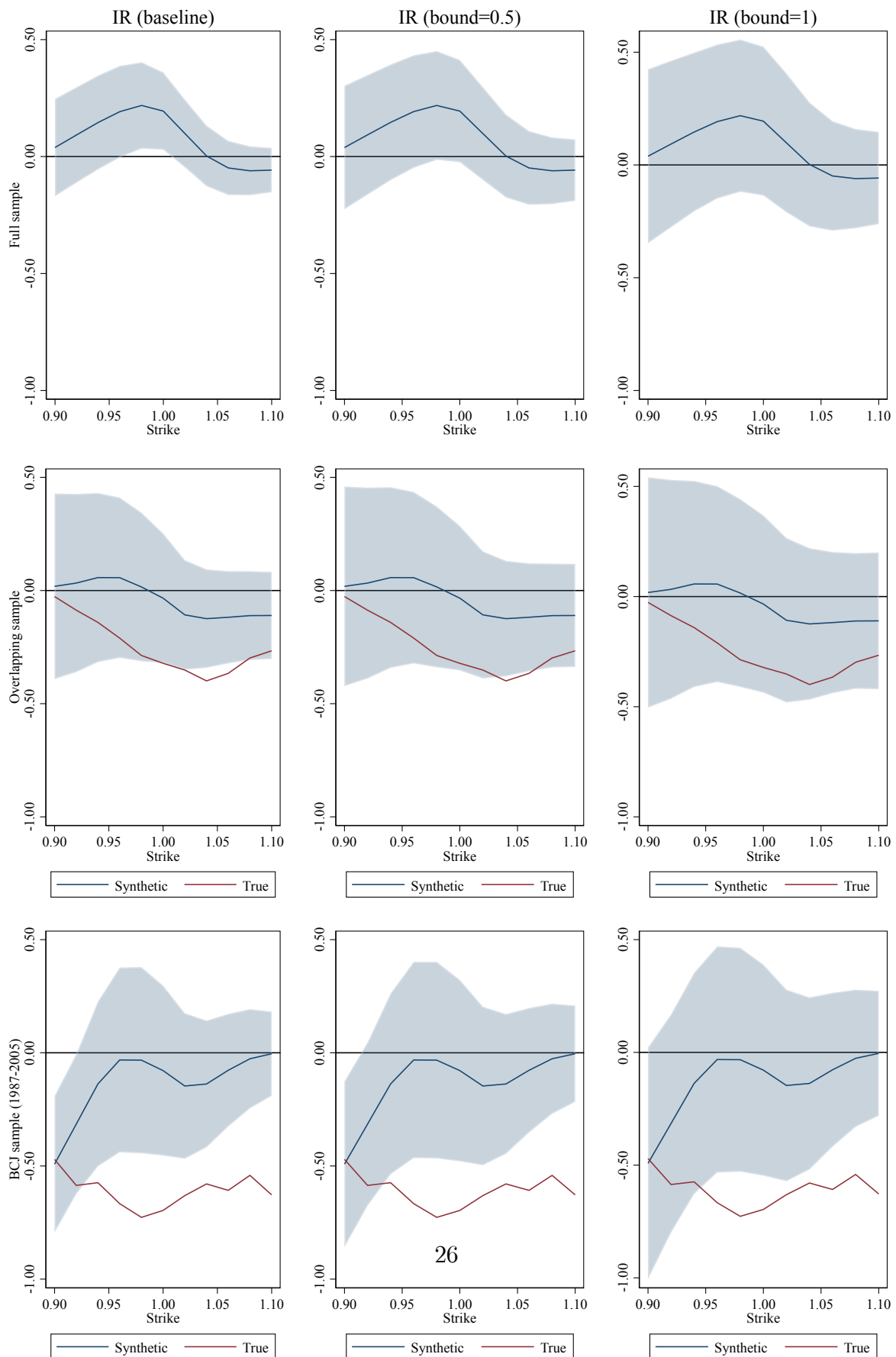
## 4.6 Implications for marginal utility

As discussed in section 2, the estimates of the CAPM alphas reported above can be used to measure curvature in marginal utility. Figure 1 plots, in red, the shape of SMU implied by the synthetic option returns for strike of -5% a maturity of one month for various samples. In all cases, the synthetic options imply that marginal utility is if anything weakly concave, consistent with the positive measured alphas. That is, the data implies that marginal utility is *less* sensitive to market returns when those returns are negative, consistent with risk aversion falling slightly in bad times. While there are no confidence bands plotted in figure 1, the change in the slope across the strike is measured by the alphas reported above, and so the confidence bands apply here.

As discussed above, the blue lines corresponding to traded options imply significant convexity due to the large negative estimated alphas, so that effective risk aversion rises strongly as market returns fall.

The analysis of how to estimate SMU based on the returns on the market and a single option is a special case of the minimum-variance SDF of Hansen and Jagannathan (1991). The logic then naturally extends to many options. Given a set of options – traded or synthetic

Figure 6: Average option returns with robust uncertainty intervals



– along with the market return, the minimum-variance SMU consistent with those returns is constructed based on the mean-variance optimal portfolio of the investments.

Figure 7 plots the minimum-variance SMU for the traded and synthetic options separately. In each case, we use strikes of 90, 95, ..., 110 percent of the spot along with the market return. The mean-variance optimal portfolio is calculated based on the full-sample estimates of mean returns and covariances, with both a lasso- and ridge-type adjustment to ensure robustness.<sup>21</sup>

The figure plots multivariate estimates of SMU for the 1926–2021 sample for synthetic options and 1987–2021 for traded options. As in the benchmark case, estimated SMU is convex for the true options and concave for the synthetic options. An advantage of the multivariate specification is that it gives the ability to potentially measure where the nonlinearity lies. For the synthetic options, the concavity is fairly consistent, though it may change signs at large positive strikes. For the traded options, the convexity appears strongest around market returns of zero. That is consistent with the finding in the previous section that it is the at-the-money true options that have the most negative alphas (see also Broadie, Chernov, and Johannes (2009) for a similar result). While traded options again imply non-monotone utility, the synthetic options do not.

## 4.7 Robustness

The baseline results use deltas from the method of Hull and White (2017). The top-right panel of figure A.7 shows that the results are highly similar simply using the standard Black–Scholes delta.

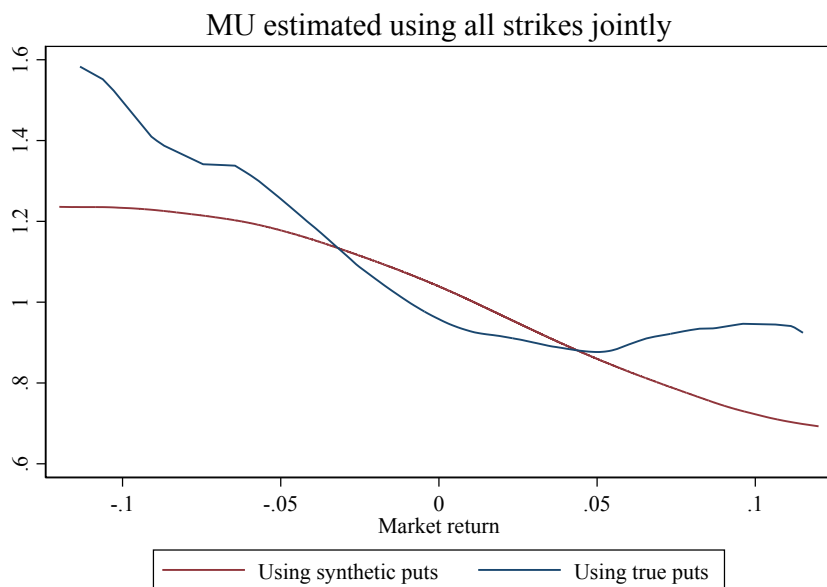
Constructing the weights for the synthesis requires a forecast of volatility. The benchmark analysis uses a recursively estimated HAR model (Corsi (2009)). The bottom-left panel of figure A.7 shows that the alphas are very similar to those in the baseline if the volatility used to calculate the hedge weights is simply set to 0.15 on all dates. That shows that the results are driven by how  $\delta_t^S$  depends on the level of the market, rather than its volatility.

The benchmark analysis uses asymptotic standard errors based on the Hansen–Hodrick method. Since option returns are highly non-normal, convergence to the asymptotic distribution might be slow, which can be addressed via bootstrapping. Results with block-bootstrapped standard errors with block length equal to two are reported in bottom-right

---

<sup>21</sup>Mean-variance optimization requires inverting the covariance matrix. We make two modifications to ensure that the inversion is well behaved. First, we inflate the main diagonal of the covariance matrix by 10%, which corresponds to a ridge-type adjustment. Second, in constructing the inverse, we drop any eigenvalue smaller than 0.01 times the largest (in practice this eliminates one eigenvalue). These adjustments only affect the weights in the tangency portfolio.

Figure 7: Estimated SMU using all options jointly



**Note:** The figure reports the estimated SMU obtained by using options of all strikes jointly. The green line uses synthetic puts (sample 1926-2021), and the red line uses actual puts (sample 1987-2021).

panel of figure A.7, and they are highly similar to the baseline.

## 5 Intermediary frictions and the change in option returns over time

The analysis so far shows theoretically that there are two potential ways to estimate curvature in marginal utility and finds empirically that they give different answers. The theory showed that there are three conditions that are needed in order to be able to estimate curvature from synthetic options, and the empirical analysis provided some evidence in their favor. And even allowing for a conservative bound on the pricing of unspanned risk in synthetic options, the implied convexity in SMU is far smaller than what is measured from traded options.

There is also a required condition for using traded options to measure the curvature of SMU, though, which is that they are priced by marginal utility. One is to look for a relatively frictionless model in which prices are driven by preferences or beliefs. In that case, to rationalize the results it must be that the unspanned risk in synthetic options is very strongly priced by investors (even more strongly than the bounds we examined allow). If,

additionally, the gap in returns between traded and synthetic options has declined, which is studied in greater detail in this section, then that must be either because preferences changed or unspanned risk shrank.

Alternatively, there are models centered on intermediary frictions, in which the marginal utility of the average equity investor need not price options. In those models, the gap between traded and synthetic option returns is driven by demand pressure and the cost of hedging for intermediaries, instead of preferences over unspanned risk. If the cost of hedging has fallen, that can explain a decline in the traded-synthetic return spread. This section presents such a model.

## 5.1 The decline of option overpricing

Figure 4 above already showed some suggestive evidence that the returns on traded options may have trended towards zero in recent years. This section analyzes that behavior in greater detail.

### 5.1.1 Rolling window estimation

The left-hand panel of figure 8 plots information ratios over rolling 10-year windows for traded and synthetic puts along with their difference for the 1987–2021 sample.<sup>22</sup> The synthetic options consistently had zero or positive returns, while the traded options consistently had strongly negative returns until declining over the past 15 years, eventually turning starkly positive in 2020. In the early periods, the information ratio for the traded options was very large – about equal to the size of the market risk premium itself.

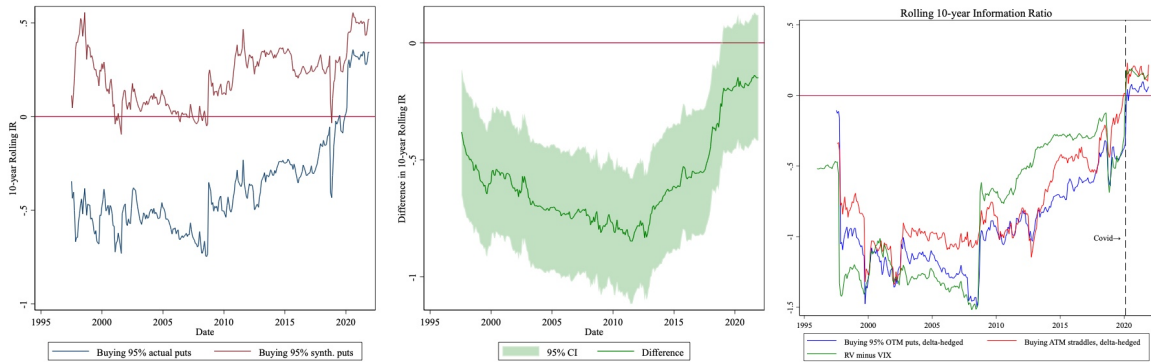
The difference between the information ratios on traded and synthetic options, plotted in the middle panel of figure 8, has also nearly converged to zero over the same period. The confidence bands show that the change in the information ratio appears to be highly statistically significant, which is tested more formally below.

The difference in the information ratios between traded and synthetic options is very closely related to the return on delta-hedged options, which have been studied widely in the past literature and used as a proxy for the variance risk premium (e.g. Bakshi and Kapadia (2003)). The right-hand panel of figure 8 plots the rolling 10-year information ratios for delta-hedged 5% OTM put options and at-the-money straddles. Both have risen over time,

---

<sup>22</sup>The main results use the overlapping 21-day returns for synthetic option. In this section, all results use synthetic option returns that match exactly the roll dates of the traded options, in order to ensure comparability. This has only minor effects, and in any case they run against the main conclusions of this section.

Figure 8: Changes in premia over time



**Note:**

converging to zero in 2020 (with the convergence for the ATM straddles coming even before Covid).

As an even simpler test, the green line in the same plot proxies for the payoff of a variance swap simply as the gap between realized variance and the squared VIX (since the VIX is, under fairly general conditions, the expectation of integrated quadratic variation, with a small error due to jump variation; see Carr and Wu (2009)). The information ratio of RV-VIX behaves highly similarly to the other series, also converging to zero over the sample.<sup>23</sup>

The results here do not mean that the variance risk premium is zero by the end of the sample. There is still a premium for variance risk, but the results imply that the premium is no larger than what would be expected from the beta. Variance rises when the market falls, so it has a negative beta and carries a negative premium. In the past, the premium from trading variance via delta-hedged options was even larger than is implied by the CAPM beta, but by the end of the sample that is no longer true. These results are consistent with those in Heston, Jacobs, and Kim (2022), who also find that there is a negative variance risk premium, but that it cannot be distinguished from simple market (beta) risk.

### 5.1.2 Formal tests for a structural break

We examine three 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,

<sup>23</sup>Specifically, define a (pseudo-) return  $R_t^{RV} = (RV_t - VIX_{t-1}^2) / VIX_{t-1}^2$ , where  $RV_t$  is (annualized) realized variance in month  $t$  and  $VIX_{t-1}$  is the level of the VIX at the end of month  $t - 1$ . The information ratio is then the CAPM alpha of  $R_t^{RV}$  divided by the residual standard deviation.

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 8.

The table below tests for changes in the difference in the information ratios for traded and synthetic options at various dates. 12/2003 and 12/2008 are chosen based on the decline in bid/ask spreads in the former case and the index-futures basis in the latter (both discussed further below). The bottom rows, 9/2011 and 1/2013, are chosen as the optimal date from a Wald test perspective. The p-values in that case are calculated using the optimal exponential Wald statistic of Andrews and Ploberger (1994), which corrects for multiple testing.

<b>Difference in information ratios, post-1987</b>			
Breakpoint	Before break	After	p-value
12/2003	0.92	0.29	0.00
12/2008	0.79	0.25	0.01
9/2011	0.81	0.07	0.00

<b>Difference in information ratios, including 1987</b>			
Breakpoint	Before break	After	p-value
12/2003	0.51	0.24	0.31
12/2008	0.48	0.18	0.19
1/2013	0.52	-0.05	0.08

The changes in the information ratios are economically significant in all cases. The p-values are very small when 1987 is excluded from the sample. When it is included, the evidence for a break is weaker, though still close to significant at the optimal break date.

Overall, then, this section shows that returns on exchange-listed options have converged to those on synthetic options in a way that is economically meaningful and on the edge of statistical significance using conventional cutoffs. Over time the positive excess return from shorting traded options has shrunk until by the end of the sample there is no significant difference between the true and synthetic options, and their alphas are both 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, 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 derivatives that pay convex functions of the market return.

## 5.2 Model

This section studies a simple extension of the model of Garleanu, Pedersen, and Poteshman (GPP; 2009) to help clarify how various frictions can affect option prices when markets are segmented. The only addition to their framework is to allow for transaction costs and index-futures basis risk. The main text describes just the key parts of the setup and predictions of the model. The details are in appendix D.

### 5.2.1 Setup

In GPP’s model, the price of the underlying – which we take here to be the market index – is determined exogenously (presumably by a much larger mass of traders), as is the demand or supply for a set of derivative claims, such as options. GPP show empirically that in general retail investors appear to be long index options, so that dealers must be net short.

The dealers are assumed to have time-additive CARA preferences over consumption with risk aversion  $\gamma$ . The key equation in the model is the dynamic budget constraint,

$$W_{t+1} = (W_t - C_t) R_f + q_t R_{t+1}^O + \theta_t R_{t+1}^F - \frac{\kappa_1}{2} \theta_t^2 \quad (17)$$

$W_t$  is wealth and  $C_t$  consumption. The risk-free rate,  $R_f$  is constant for simplicity,  $R_{t+1}^F$  is the gross return on index futures, and  $R_{t+1}^O$  is the return on the derivatives. The dealers endogenously choose consumption and the allocations to derivatives and futures  $q_t$  and  $\theta_t$ , respectively.

There are two frictions: a quadratic trading cost,  $\frac{\kappa_1}{2} \theta_t^2$ , and a wedge between the futures return,  $R_{t+1}^F$ , and the underlying index return,  $R_{t+1}^I$ ,

$$R_{t+1}^F = R_{t+1}^I + z_{t+1} \quad (18)$$

We refer to  $z_{t+1}$  as *basis risk*. Ideally the dealers would like to hedge the options they trade with the underlying, like the S&P 500. But the S&P 500 is not itself directly tradable (except at significant cost by buying 500 stocks). Instead, dealers must buy futures (or ETFs or other instruments) whose price is not guaranteed to perfectly track the index.  $z_{t+1}$  captures the risk associated with imperfect tracking.

While the dealers choose  $q_t$ , markets must clear, meaning that in equilibrium their choice of  $q_t$  must perfectly offset the (exogenous) demand from retail investors. The core idea in GPP is to understand how derivative prices, denoted by  $P_t$ , vary with quantities,  $q_t$ .



## 5.2.2 Predictions

In the model, intermediaries hedge their options each period with a position in the underlying – it can be thought of as a delta hedge that is adjusted each period. The optimal position, in the lack of any frictions, is denoted by  $\beta_t^I$  (which is simply the local sensitivity of option returns to the underlying index). The unhedgeable risk is defined as

$$\sigma_{\varepsilon,t}^2 \equiv \text{var}_t^d (R_{t+1}^O - \beta_t^I R_{t+1}^I) \quad (19)$$

where  $\text{var}_t^d$  is a variance taken under the intermediaries' pricing measure  $d$  based on date- $t$  information.

The model's key prediction is for the sensitivity of option prices,  $P_t^O$ , to demand:

**Proposition 4** *Up to first order in the transaction cost  $\kappa_1$  and the index-futures basis risk  $\text{var}_t^d(z_{t+1})$ ,*

$$\frac{\partial P_t^O}{\partial q_t} = -\gamma (R_f - 1) \left( \underbrace{\sigma_{\varepsilon,t}^2}_{\text{Unhedgeable risk}} + \underbrace{(\beta_t^I)^2 \text{var}_t^d(z_{t+1})}_{\text{Basis risk}} \right) + \underbrace{\kappa_1 R_f^2 (\beta_t^I)^2}_{\text{Imperfect hedging}} \quad (20)$$

The sensitivity is proportional to risk aversion,  $\gamma$ , and has three terms.

The first component,  $\sigma_{\varepsilon,t}^2$ , is the unhedgeable risk. Dealers hedge by taking positions exposed to the underlying, but since options have nonlinear exposure, that hedge is inevitably imperfect. Note that this unhedgeable risk is the same as the unspanned risk discussed above as a potential driver of the gap in pricing between traded and synthetic options. It is the part of the option return that cannot be spanned by a dealer replicating options with discretely updated positions.

The second term represents basis risk. When there are larger random gaps between the hedging instrument and the true underlying index, dealers face greater risk and thus demand larger premia.

Finally, the third term arises due to the quadratic trading cost  $\kappa_1$ . The trading cost causes dealers to hedge incompletely, further raising their risk from holding derivatives.

In the context of the general theoretical analysis in section 2, this is a model in which traded options are not priced by the marginal utility of equity investors. And the exogenous option demand, since it must be borne only by dealers, drives option prices up, creating negative CAPM alphas.

## 5.3 Empirics

The results in section 5.1 show that the premium for traded options has shrunk toward zero over time. The extension of the GPP model here would predict that there should be a decline in the option premium if any of the three factors in equation (20) have shrunk. This section examines the three factors and how they have changed over time.

### 5.3.1 Trading frictions

To see how trading costs have changed, the top-left panel of figure 9 plots measures of bid-ask spreads for equities. The gray line is the average spread for the Dow 30, blue is for the DIA Dow 30 ETF, and red is for the SPY S&P 500 ETF. Spreads were around 100 basis points until the late 1990's, falling with both the rise in electronic trading and decimalization.

The top-right panel of figure 9 plots effective spreads based on the Roll (1984) estimator for S&P 500 futures. Effective spreads also have consistently fallen over time. While the decline looks close to linear, note that the y-axis is on a log scale, so that in absolute terms the declines were much larger in the early part of the sample. Overall, effective spreads fell by about a factor of 100 over the 1982–2021 period.

### 5.3.2 Basis risk

Basis risk, as defined in the model, is the standard deviation of the gap between the underlying index and the futures price. That can be measured empirically by the gap between the level of the S&P 500 index and the futures price.<sup>24</sup> The middle-left panel of figure 9 plots the three-month rolling standard deviation of that gap. The y-axis is again on a  $\log_{10}$  scale.

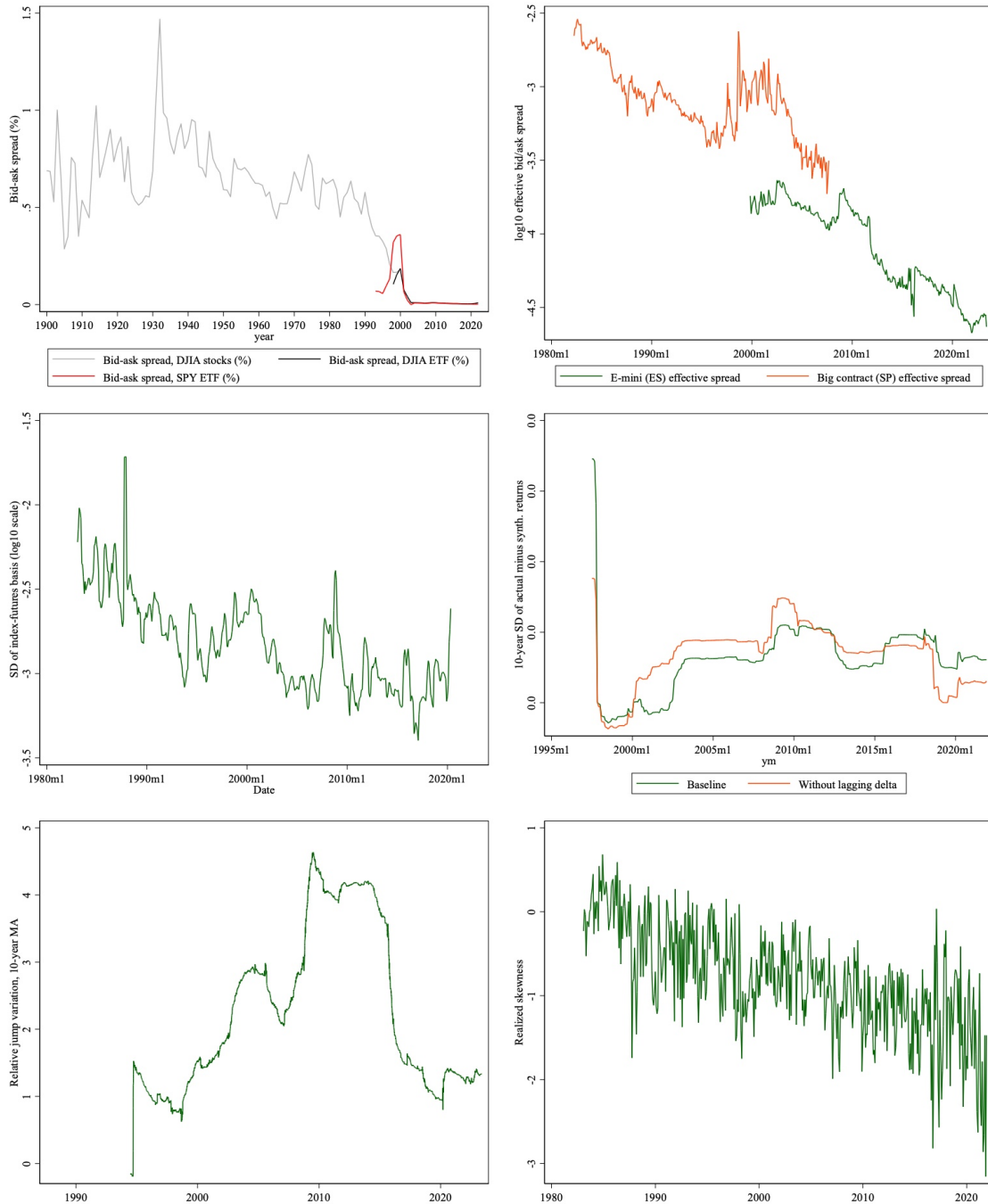
Over time, basis risk has fallen by about an order of magnitude. While there is a large decline early in the sample, similar to trading frictions, basis risk seems to settle at its current level around the early 2000's.

Also similar to trading frictions, basis risk is significantly higher during market crashes – that is clearly apparent for 1987, 2008, and 2020. However, note that in the latter two episodes basis risk was far, far smaller. During the 1987 crash, SP futures were underpriced by as much as 20% for a period of three days. During the 2008 crash, on the other hand, the mispricing was never larger than 5%. In the 2008 and 2020 crashes, the volatility of the basis is lower by an order of magnitude than in 1987. To the extent that it is basis risk in

---

<sup>24</sup>Specifically, the gap is measured for five-minute windows as the log difference in the 5-minute averages of level of the S&P 500 index and the near-month CME SP futures price. The intraday values for the S&P 500 index are obtained from tickdata.com (and we have also confirmed that they match those obtained from Cboe well).

Figure 9: Various measures of hedging risk



**Note:** These graphs replicate the main results, but the shaded uncertainty intervals here incorporate the bound on the effect from potential pricing of unspanned risk in the synthetic option returns.

crashes that dealers are worried about – futures prices separating from the underlying index at exactly the moment when the hedge is most important – that is something that can really only be learned about in a crash. In other words, 2008 represented a stress test that the futures market appears to have passed in this sense, after which the model would predict a smaller options premium.

### 5.3.3 Unhedgeable risk

As GPP discuss, the three major sources of unhedgeable risk come from discreteness in hedging (i.e. the time delay before the hedge is updated), jumps, and risks that are unspanned by the underlying (e.g. shocks to volatility that are orthogonal to the underlying return).

Figure 9 plots three measures of unhedgeable risk. The first is the standard deviation of the daily delta-hedged options return. It shows no clear trend. While the 10-year rolling standard deviation is high when the 1987 crash is included, outside that event it has been flat, with no downward trend matching the change in traded option returns.

The second measure of unhedgeable risk is based on the idea that jumps are a potential major driver. It is the measure of jumps as the gap between quadratic and bipower variation in the S&P 500 return (see Bollerslev et al. (2009)). Relative jump variation rose during the 2008 financial crisis, and has been lower subsequently, but again does not have a clear trend. The period when traded option returns were most negative does not match the period when jump variation was highest.

Finally, unhedgeable risk is, more broadly, driven by higher moments in returns, so the bottom-right panel of figure 9 plots the measure of S&P 500 return skewness from Neuberger (2012). Realized skewness has, over time, trended consistently more negative (implied skewness does the same; see the CBOE’s SKEW index and Dew-Becker (2022)?).

Overall, figure 9 shows that unhedgeable risk does not appear to have clearly fallen along with options alphas.

## 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.
- Ait-Sahalia, Yacine, Mustafa Karaman, and Lorian Mancini**, “The term structure of equity and variance risk premia,” *Journal of Econometrics*, 2020, *219* (2), 204–230.
- Alvarez, Fernando and Urban J. Jermann**, “Using Asset Prices to Measure the Cost of Business Cycles,” *Journal of Political Economy*, 2004, *112*(6), 1223–1256.
- Andersen, Torben G, Nicola Fusari, and Viktor Todorov**, “The risk premia embedded in index options,” *Journal of Financial Economics*, 2015, *117* (3), 558–584.
- 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.

- Backus, David, Mikhail Chernov, and Ian Martin**, “Disasters implied by equity index options,” *The Journal of Finance*, 2011, *66* (6), 1969–2012.
- 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.
- Bates, David S**, “The market for crash risk,” *Journal of Economic Dynamics and Control*, 2008, *32* (7), 2291–2321.
- , “US stock market crash risk, 1926–2010,” *Journal of Financial Economics*, 2012, *105* (2), 229–259.
- , “Empirical Option Pricing Models,” *Annual Review of Financial Economics*, 2022, *14*, 369–389.
- Beason, Tyler and David Schreindorfer**, “Dissecting the equity premium,” 2022.
- Black, Fischer and Myron Scholes**, “The Pricing of Options and Corporate Liabilities,” *The Journal of Political Economy*, 1973, *81* (3), 637–654.
- 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.
- Bollerslev, Tim and Viktor Todorov**, “Tails, fears, and risk premia,” *Journal of Finance*, 2011, *66*(6), 2165–2211.
- , **Uta Kretschmer, Christian Pigorsch, and George Tauchen**, “A discrete-time model for daily S & P500 returns and realized variations: Jumps and leverage effects,” *Journal of Econometrics*, 2009, *150* (2), 151–166.
- Broadie, Mark, Mikhail Chernov, and Michael Johannes**, “Model Specification and Risk Premia: Evidence from Futures Options,” *The Journal of Finance*, 2007, *62*(3), 1453–1490.
- , —, and —, “Understanding Index Option Returns,” *The Review of Financial Studies*, 2009, *22*(11), 4493–4529.
- Büchner, Matthias and Bryan Kelly**, “A factor model for option returns,” *Journal of Financial Economics*, 2022, *143* (3), 1140–1161.
- Carr, Peter and Dilips B. Madan**, *Towards a Theory of Volatility Trading*, London: Risk Books,

- **and Liuren Wu**, “Variance Risk Premiums,” *Review of Financial Studies*, 2009, *22*(3), 1311–1341.
- Chambers, Donald R, Matthew Foy, Jeffrey Liebner, and Qin Lu**, “Index option returns: Still puzzling,” *The Review of Financial Studies*, 2014, *27* (6), 1915–1928.
- Cochrane, John H and Jesus Saa-Requejo**, “Beyond arbitrage: Good-deal asset price bounds in incomplete markets,” *Journal of political economy*, 2000, *108* (1), 79–119.
- 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.
- Corsi, Fulvio**, “A simple approximate long-memory model of realized volatility,” *Journal of Financial Econometrics*, 2009, *7* (2), 174–196.
- Coval, Joshua D. and Tyler Shumway**, “Expected Option Returns,” *The Journal of Finance*, 2001, *56*(3), 983–1009.
- Cuesdeanu, Horatio and Jens Carsten Jackwerth**, “The pricing kernel puzzle: Survey and outlook,” *Annals of Finance*, 2018, *14*, 289–329.
- Dew-Becker, Ian**, “Real-time forward-looking skewness over the business cycle,” 2022. Working paper.
- **and Stefano Giglio**, “Cross-sectional uncertainty and the business cycle: evidence from 40 years of options data,” 2020. Working paper.
- , — , **Anh Le, and Marius Rodriguez**, “The price of variance risk,” *Journal of Financial Economics*, 2017, *123* (2), 225 – 250.
- Drechsler, Itamar**, “Uncertainty, Time-Varying Fear, and Asset Prices,” *The Journal of Finance*, 2013, *68* (5), 1843–1889.
- **and Amir Yaron**, “What’s Vol Got to Do with it?,” *The Review of Financial Studies*, 2011, *24*(1), 1–45.
- Du, Du**, “General equilibrium pricing of options with habit formation and event risks,” *Journal of Financial Economics*, 2011, *99*, 400–426.
- 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.
- Haddad, Valentin and Tyler Muir**, “Do intermediaries matter for aggregate asset prices?,” *The Journal of Finance*, 2021, *76* (6), 2719–2761.
- Hall, Robert E**, “High discounts and high unemployment,” *American Economic Review*, 2017, *107* (2), 305–30.
- Han, Bing**, “Investor sentiment and option prices,” *The Review of Financial Studies*, 2008, *21* (1), 387–414.
- Hansen, Lars Peter and Ravi Jagannathan**, “Implications of Security Market Data for Models of Dynamic Economies,” *Journal of Political Economy*, 1991, *99*(2), 225–262.
- 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.
- He, Zhiguo and Arvind Krishnamurthy**, “Intermediary asset pricing,” *American Economic Review*, 2013, *103* (2), 732–770.
- Heston, Steven, Kris Jacobs, and Hyung Joo Kim**, “Exploring Risk Premia, Pricing Kernels, and No-Arbitrage Restrictions in Option Pricing Models.” Working paper.
- Hull, John and Alan White**, “Optimal delta hedging for options,” *Journal of Banking & Finance*, 2017, *82*, 180–190.
- Jackwerth, Jens Carsten**, “Recovering risk aversion from option prices and realized returns,” *The Review of Financial Studies*, 2000, *13* (2), 433–451.
- Jegadeesh, Narasimhan and Sheridan Titman**, “Short-horizon return reversals and the bid-ask spread,” *Journal of Financial Intermediation*, 1995, *4* (2), 116–132.
- Jurek, Jakub W and Erik Stafford**, “The cost of capital for alternative investments,” *The Journal of Finance*, 2015, *70* (5), 2185–2226.
- Kozhan, Roman, Anthony Neuberger, and Paul Schneider**, “The skew risk premium in the equity index market,” *The Review of Financial Studies*, 2013, *26* (9), 2174–2203.



- Lamoureux, Christopher G and William D Lastrapes**, “Forecasting stock-return variance: Toward an understanding of stochastic implied volatilities,” *The Review of Financial Studies*, 1993, 6 (2), 293–326.
- Lewellen, Jonathan and Stefan Nagel**, “The Conditional CAPM does not Explain Asset-Pricing Anomalies,” *Journal of Financial Economics*, 2006, 82 (2), 289–314.
- Martin, Ian**, “What is the Expected Return on the Market?,” *The Quarterly Journal of Economics*, 2017, 132 (1), 367–433.
- Muir, Tyler**, “Financial crises and risk premia,” *The Quarterly Journal of Economics*, 2017, 132 (2), 765–809.
- 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.
- Neuberger, Anthony**, “Realized skewness,” *The Review of Financial Studies*, 2012, 25 (11), 3423–3455.
- 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.
- Paoli, Bianca De and Pawel Zabczyk**, “Cyclical risk aversion, precautionary saving, and monetary policy,” *Journal of Money, Credit and Banking*, 2013, 45 (1), 1–36.
- Roll, Richard**, “A simple implicit measure of the effective bid-ask spread in an efficient market,” *The Journal of finance*, 1984, 39 (4), 1127–1139.
- 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.
- and **Tobias Sichert**, “Volatility and the Pricing Kernel.” Working paper.
- Seo, Sang Byung and Jessica A Wachter**, “Option prices in a model with stochastic disaster risk,” *Management Science*, 2019, 65 (8), 3449–3469.

## A Proofs from section 2

### A.1 Proposition 2

Consider a regression of  $M_{t,t+j}$  on  $R_{t,t+j}^m$  and  $R_{t,t+j}^{O\perp}$ . Since  $R_{t,t+j}^m$  and  $R_{t,t+j}^{O\perp}$  are, by construction, conditionally uncorrelated with each other, we have

$$M_{t,t+j} = \text{const.} + \frac{\text{cov}_t(R_{t,t+j}^m, M_{t,t+j})}{\text{var}_t[R_{t,t+j}^m]} R_{t,t+j}^m + \frac{\text{cov}_t(R_{t,t+j}^{O\perp}, M_{t,t+j})}{\text{var}_t(R_{t,t+j}^{O\perp})} R_{t,t+j}^{O\perp} + \text{resid.} \quad (21)$$

Additionally,  $R_{t,t+j}^m$  is conditionally uncorrelated with  $\hat{M}_{t,t+j}$  by construction, so that

$$\text{cov}(R_{t,t+j}^m, M_{t,t+j}) = \text{cov}(R_{t,t+j}^m, \bar{M}_{t,t+j}) \quad (22)$$

The same fact works for  $R_{t,t+j}^{O\perp}$  since, conditional on date- $t$  information,  $R_{t,t+j}^{O\perp}$  is a (nonlinear) function of  $R_{t,t+j}^m$ . We then can replace  $M_{t,t+j}$  on the left-hand side above with  $\bar{M}_{t,t+j}$ .

Under the assumption that  $R_{t,t+j}^m$  and  $R_{t,t+j}^{O\perp}$  are priced,

$$\text{cov}(R_{t,t+j}^m, M_{t,t+j}) = -E_t[R_{t,t+j}^m - 1] \quad (23)$$

$$\text{cov}(R_{t,t+j}^{O\perp}, M_{t,t+j}) = -E_t[R_{t,t+j}^{O\perp} - 1] \quad (24)$$

But since  $R_{t,t+j}^{O\perp}$  is uncorrelated with  $R_{t,t+j}^m$ , its (conditional) CAPM beta is zero by construction, and hence

$$E_t[R_{t,t+j}^{O\perp} - 1] = \alpha_{t,t+j}^O \quad (25)$$

The result then follows.

### A.2 Proposition 3

The proof follows similar lines to that for proposition 2. The only required adjustment is to note that

$$\text{cov}_t(R_{t,t+j}^S, M_{t,t+j}) = -E_t[R_{t,t+j}^S - 1] = -\alpha_{t,t+j}^S \quad (26)$$

and

$$\text{cov}_t(R_{t,t+j}^S, \bar{M}_{t,t+j}) = \text{cov}_t(R_{t,t+j}^S, M_{t,t+j}) - \text{cov}_t(\hat{M}_{t,t+j}, \hat{R}_{t,t+j}^S) \quad (27)$$

$$= -\left(\alpha_{t,t+j}^S + \text{cov}_t(\hat{M}_{t,t+j}, \hat{R}_{t,t+j}^S)\right) \quad (28)$$

## B Synthetic puts and replicating portfolio

### B.1 Standard delta-hedging

Consider an option with price  $P_t$ , and the underlying with price  $S_t$ . The payoff from holding the option from time 1 to  $T$  is:

$$\Pi_{1,T}^{option} = P_T - P_1 = \sum_{t=1}^{T-1} (P_{t+1} - P_t)$$

The Black-Scholes-Merton replication portfolio is a dynamic strategy that buys a time-varying number of shares of the underlying,  $\Delta_t$ , and invests a time-varying amount  $B_t$  in the risk free rate. These numbers are chosen so that, in the original setup of Black and Scholes, they guarantee an exact replication of the option value as it evolves over time: equivalently, they guarantee that at maturity the payoff of the option and the replicating portfolio are equal, while cash flows are zero in every period except the first and the last.

To achieve this,  $\Delta_t$  is chosen as the Black-Scholes delta, and  $B_t$  is chosen as the difference between the BS option value  $P_t$  and the cost of buying the underlying  $\Delta_t S_t$ :

$$B_t = P_t - \Delta_t S_t$$

This portfolio costs  $B_t + \Delta_t S_t$  to buy at time  $t$ , and has a value of  $B_t(1 + \frac{r_t}{365}) + \Delta_t S_{t+1}$  at time  $t + 1$ , where  $r_t$  is the annualized interest rate at time  $t$ . At time  $t + 1$ , the strategy requires buying the new portfolio  $B_{t+1} + \Delta_{t+1} S_{t+1}$ , and so on. In every period between the first and the last one, this strategy generates an *intermediate cash flow* of:

$$ICF_{t+1} = (B_t(1 + \frac{r_t}{365}) + \Delta_t S_{t+1}) - (B_{t+1} + \Delta_{t+1} S_{t+1})$$

and at the last period  $T$ , it generates a final cash flow of  $B_{T-1}(1 + \frac{r_{T-1}}{365}) + \Delta_{T-1} S_T$  which is equal to the option payoff. In the BS model, the portfolio is self-financing, so  $ICF_{t+1} = 0$ .

Therefore, the total P&L that this replication portfolio generates is:

$$\Pi_{1,T}^{hedge} = \underbrace{\sum_{t=1}^{T-2} \left[ (B_t(1 + \frac{r_t}{365}) + \Delta_t S_{t+1}) - (B_{t+1} + \Delta_{t+1} S_{t+1}) \right]}_{=0 \text{ under BS}} + \underbrace{\left[ B_{T-1}(1 + \frac{r_{T-1}}{365}) + \Delta_{T-1} S_T \right]}_{\text{final cash flow}} - \underbrace{[B_1 + \Delta_1 S_1]}_{\text{initial cost}} \quad (29)$$

This can be also rewritten in the following way:

$$\Pi_{1,T}^{hedge} = \sum_{t=1}^{T-1} \left[ (B_t(1 + \frac{r_t}{365}) + \Delta_t S_{t+1}) - (B_t + \Delta_t S_t) \right] = \sum_{t=1}^{T-1} \Delta_t (S_{t+1} - S_t) + \sum_{t=1}^{T-1} B_t \frac{r_t}{365} \quad (30)$$

which corresponds to the formula of Bakshi and Kapadia (2003) and Buchner and Kelly (2022) when setting  $B_t = P_t - \Delta_t S_t$ . When the assumptions of Black-Scholes do not hold, the rebalancing of the portfolio generates intermediate cash flows (i.e.,  $CF_t$  is not zero), which is accounted for by the formula above.

Finally, we examine *excess* returns of the delta-hedged portfolio. Both  $\Pi_{1,T}^{option}$  and  $\Pi_{1,T}^{hedge}$  are dollar payoffs that correspond to initial investments of  $P_1$  and  $B_1 + \Delta_1 S_1$ , respectively. If the objective is to compute delta-hedged returns, then one can compute:

$$\Pi_{1,T}^{option} - \Pi_{1,T}^{hedge} = \sum_{t=1}^{T-1} (P_{t+1} - P_t) - \sum_{t=1}^{T-1} \Delta_t (S_{t+1} - S_t) - \sum_{t=1}^{T-1} B_t \frac{r_t}{365} \quad (31)$$

and this is an excess return: the cost of buying the option is  $P_1$ , the income from shorting the hedge portfolio is  $B_1 + \Delta_1 S_1 = (P_1 - \Delta_1 S_1) + \Delta_1 S_1 = P_1$ . This is because the hedge portfolio is designed to borrow exactly an amount  $B_1$  that fully matches the price of the option  $P_1$  (which, in this case, is observed). The hedge strategy updates  $B_t$  over time by always setting  $B_t = P_t - \Delta_t S_t$ , using the observed price of the option  $P_t$  at each point in time. Therefore, the delta-hedging strategy described above (eq. 31) is an excess return whether  $P_t$  conforms or not to the BS prices. The only difference is, if the BS model is correct, then the  $ICF_t$  and the last period cash flow will all be zero.

## B.2 Synthetic options: P&L of zero-cost strategy

Next, we consider the case in which we do not observe the option price  $P_t$ . In that case, we cannot use it as an input for computing  $B_t$  and  $\Delta_t$ . We also cannot obviously compute  $\Pi_{1,T}^{option}$ . However, we note that equation (29) still describes the total P&L of *any* dynamic trading strategy that at each point in time buys  $\Delta_t$  units of the underlying and invests  $B_t$  in the risk free rate, whether or not those are chosen as per the BS model. Therefore, we can still compute  $\Pi_{1,T}^{hedge}$  for a choice of  $P_t$  and  $\Delta_t$ <sup>25</sup> (and hence  $B_t$ ). In particular, we determine the Black-Scholes price of an option,  $P_t$ , using as input the current underlying  $S_t$  and an estimated value for  $\sigma_t^2$ , and then build the hedging portfolio for that idealized option. The

---

<sup>25</sup>  $\Delta_t$  corresponds to  $\delta_{t,t+j}^S$  in the main text. To evaluate  $\Delta_t$ , we follow Hull and White (2017); specifically, we use equation (5) of the referenced paper and select  $a = -0.25, b = -0.4, c = -0.5$ .

fact that that option is not directly tradable is irrelevant, in the sense that the P&L we build as described above is an actual return of a portfolio that is just a dynamic portfolio of the market.

While the delta-hedged P&L,  $\Pi_{1,T}^{option} - \Pi_{1,T}^{hedge}$ , is, as described above, the P&L of a zero-cost portfolio,  $\Pi_{1,T}^{hedge}$  is not. So we next describe the P&L of hedge portfolios that yield  $\Pi_{1,T}^{option}$  and  $\Pi_{1,T}^{hedge}$  separately but are funded at the risk-free rate at inception. For the option (i.e. when we do observe  $P_t$ ), we have:

$$\Pi_{1,T}^{option,exc} = P_T - P_1 \left(1 + \frac{r_1}{365}\right)^T$$

Funding the hedge portfolio requires borrowing  $B_1 + \Delta_1 S_1 = P_1$ , where  $P_t$  is the theoretical BS price of an option. So the total P&L can be written as:

$$\Pi_{1,T}^{hedge,exc} = \underbrace{\sum_{t=1}^{T-2} \left[ (B_t(1 + \frac{r_t}{365}) + \Delta_t S_{t+1}) - (B_{t+1} + \Delta_{t+1} S_{t+1}) \right]}_{=0 \text{ under BS}} + \left[ B_{T-1} \left(1 + \frac{r_{T-1}}{365}\right) + \Delta_{T-1} S_T \right] - [B_1 + \Delta_1 S_1] \left(1 + \frac{r_1}{365}\right)^T$$

Note that in these formulas the intermediate cash flows are assumed not to be reinvested. One can also reinvest them to obtain:

$$\begin{aligned} \Pi_{1,T}^{hedge,exc} &= \sum_{t=1}^{T-2} (1 + r_{t+1})^{T-t-1} \left[ (B_t(1 + \frac{r_t}{365}) + \Delta_t S_{t+1}) - (B_{t+1} + \Delta_{t+1} S_{t+1}) \right] \\ &\quad + \left[ B_{T-1} \left(1 + \frac{r_{T-1}}{365}\right) + \Delta_{T-1} S_T \right] - [B_1 + \Delta_1 S_1] \left(1 + \frac{r_1}{365}\right)^T \end{aligned}$$

An alternative procedure is to get the same exposure  $\Delta_t$  every period, but entirely fund it at the risk-free rate every period. The P&L of this zero-cost portfolio is:

$$\tilde{\Pi}_{1,T}^{hedge,exc} = \sum_{t=1}^{T-1} \Delta_t (S_{t+1} - S_t (1 + \frac{r_t}{365})) = \sum_{t=1}^{T-1} \Delta_t (S_{t+1} - S_t) - \sum_{t=1}^{T-1} \Delta_t S_t \frac{r_t}{365}$$

Note that this relates closely to  $\Pi_{1,T}^{hedge,exc}$ , since

$$\Pi_{1,T}^{hedge,exc} = \sum_{t=1}^{T-1} \Delta_t (S_{t+1} - S_t) + \sum_{t=1}^{T-1} (P_t - \Delta_t S_t) \frac{r_t}{365} - P_1 \left[ \left(1 + \frac{r_1}{365}\right)^T - 1 \right]$$

where the latter term is the total interest paid on the original loan of  $P_1$ . So we can write:

$$\Pi_{1,T}^{hedge,exc} - \tilde{\Pi}_{1,T}^{hedge,exc} = \sum_{t=1}^{T-1} P_t \frac{r_t}{365} - P_1 \left[ \left(1 + \frac{r_1}{365}\right)^T - 1 \right]$$

The difference is effectively only coming from the different timing of the borrowing (every period vs. at the beginning of the month), and is unlikely to make any substantial difference empirically.

In fact, we can also consider funding the original hedging strategy,  $\Pi_{1,T}^{hedge}$  day by day instead of once at the very beginning. Modifying eq. (30):

$$\hat{\Pi}_{1,T}^{hedge} = \sum_{t=1}^{T-1} \left[ (B_t(1 + \frac{r_t}{365}) + \Delta_t S_{t+1}) - (B_t + \Delta_t S_t)(1 + \frac{r_t}{365}) \right] = \sum_{t=1}^{T-1} \Delta_t (S_{t+1} - S_t(1 + \frac{r_t}{365}))$$

So that:

$$\hat{\Pi}_{1,T}^{hedge,exc} = \tilde{\Pi}_{1,T}^{hedge,exc} \simeq \Pi_{1,T}^{hedge,exc}$$

### B.3 P&L and returns

The P&Ls described above ( $\tilde{\Pi}_{1,T}^{hedge,exc}$  and  $\Pi_{1,T}^{hedge,exc}$ ) correspond to strategies that have zero cost. Therefore, they also represent excess returns. Scaling that excess return by any time-1 quantity is also an excess return. Just like Buchner and Kelly (2022), we scale P&Ls by the the underlying at time 1:

$$R_{1,T}^{hedge,exc} = \frac{\Pi_{1,T}^{hedge,exc}}{S_1}$$

and

$$\tilde{R}_{1,T}^{hedge,exc} = \frac{\tilde{\Pi}_{1,T}^{hedge,exc}}{S_1}$$

Finally, we consider another related zero-cost trading strategy. Instead of scaling by  $S_1$ , we scale the position of the strategy that funds every day ( $\tilde{\Pi}_{1,T}^{hedge}$ ) by  $S_t$  every day:

$$\tilde{R}_{1,T}^{hedge,scaled} = \sum_{t=1}^{T-1} \Delta_t \frac{S_{t+1} - S_t(1 + \frac{r_t}{365})}{S_t}$$

Defining  $R_{t+1}^M = \frac{S_{t+1} - S_t}{S_t}$  we obtain:

$$\frac{S_{t+1} - S_t(1 + \frac{r_t}{365})}{S_t} = \frac{S_{t+1} - S_t}{S_t} - (1 + \frac{r_t}{365}) = R_{t+1}^M - R_{t+1}^f$$

and therefore

$$\tilde{R}_{1,T}^{hedge,scaled} = \sum_{t=1}^{T-1} \Delta_t (R_{t+1}^M - R_{t+1}^f)$$

A final point concerns dividends. While dividends make a small difference over short time

horizons, we can incorporate them easily in our analysis since we are not trying to hedge an actual option. In other words, we consider a synthetic option that aims to hedge a value  $S_t$  that tracks the value of an investment in the underlying that reinvests all the dividends. In that case,  $R^M$  is the one-day gross return (including dividends) of the market.

## B.4 Comparison of the different approaches

In this section, we compare our baseline excess returns ( $R_{1,T}^{hedge,exc}$ , with reinvested intermediate cash flow) to the one obtained by funding the position each day,  $\tilde{R}_{1,T}^{hedge,scaled}$ . The table below reports, for different combinations of maturity  $M$  and strike  $K$ , the correlation between  $R_{1,T}^{hedge,exc}$  and  $\tilde{R}_{1,T}^{hedge,scaled}$  and the information ratio of each of them.

## C Robust confidence bands for alpha

To combine statistical uncertainty with uncertainty from  $\text{cov}_t(\hat{R}_{t,t+j}^S, \hat{M}_{t,t+j})$ , we treat them as two independent sources of error. Specifically, suppose one starts with a diffuse prior for

$$\alpha_{t,t+j}^{S,adjusted} \equiv \alpha_{t,t+j}^S + \text{cov}_t(\hat{M}_{t,t+j}, \hat{R}_{t,t+j}^S) \quad (32)$$

Denote by  $\alpha_{t,t+j}^{S,estimated}$  the empirical estimate. Asymptotically,  $\alpha_{t,t+j}^{S,estimated} \sim N(\alpha_{t,t+j}^S, SE^2)$ , where  $SE$  is the standard error for the estimate. We also treat the second term as though it is drawn from the distribution,

$$\text{cov}_t(\hat{M}_{t,t+j}, \hat{R}_{t,t+j}^S) \sim N\left(0, \left(\frac{1}{2} \times 0.5 \times \text{std}_t(\hat{R}_{t,t+j}^S) \frac{E[R_{t,t+j}^m - 1]}{\text{std}(R_{t,t+j}^m)}\right)^2\right) \quad (33)$$

Recall from the main text that we take  $0.5 \times \text{std}_t(\hat{R}_{t,t+j}^S) \frac{E[R_{t,t+j}^m - 1]}{\text{std}(R_{t,t+j}^m)}$  as an upper bound for  $\text{cov}_t(\hat{M}_{t,t+j}, \hat{R}_{t,t+j}^S)$ . To incorporate that with the estimation uncertainty, we treat  $\text{cov}_t(\hat{M}_{t,t+j}, \hat{R}_{t,t+j}^S)$  as though it has a standard deviation of  $0.5 \times \text{std}_t(\hat{R}_{t,t+j}^S) \frac{E[R_{t,t+j}^m - 1]}{\text{std}(R_{t,t+j}^m)}$ , so that the upper bound is two standard deviations from the prior mean – i.e. at the edge of a  $\pm 2$  standard deviation interval.

Given those two assumptions along with the diffuse prior, we then have

$$\alpha_{t,t+j}^{S,adjusted} \sim N\left(\alpha_{t,t+j}^{S,estimated}, SE^2 + \left(\frac{1}{2} \times 0.5 \times \text{std}_t(\hat{R}_{t,t+j}^S) \frac{E[R_{t,t+j}^m - 1]}{\text{std}(R_{t,t+j}^m)}\right)^2\right) \quad (34)$$

## D Theoretical results for intermediary model

This section presents our version of the GPP model in more detail. The vast majority of the content is due to them; only only change is the addition of the trading friction and index-futures basis.

### D.1 Basic setup

There is a constant gross risk-free rate  $R_f$ . The underlying index has an exogenous excess return  $R_{t+1}^I$ . We consider a simplified version of the model where there is a single option traded that has some price  $P_t$ . Its excess return is then  $R_{t+1}^O = P_{t+1} - R_f P_t$ .

Dealers/intermediaries maximize discounted utility over consumption  $C_t$ ,

$$E_t \sum_{j=0}^{\infty} \rho^j (-\gamma^{-1}) \exp(-\gamma C_t) \quad (35)$$

subject to a transversality condition and budget constraint, which is

$$W_{t+1} = (W_t - C_t) R_f + q_t R_{t+1}^O + \theta_t R_{t+1}^F - \frac{\kappa_1}{2} \theta_t^2 \quad (36)$$

where  $W_t$  is wealth. The intermediaries optimize over  $q_t$ ,  $\theta_t$ , and  $C_t$  subject to the budget constraint and taking the returns as given.

As described in the text, the term  $\frac{\kappa_1}{2} \theta_t^2$  is a deviation from GPP, as is the distinction between  $R^F$  and  $R^I$ .

It is assumed that the futures contract on the underlying that the dealers trade is available in infinite supply. For the options, there is some exogenous demand from outside investors,  $d_t$ , and the market clearing condition is  $q_t + d_t = 0$ .

**Lemma 5** *In this model, assets are priced under a probability measure  $d$  which is equal to the measure  $P$  multiplied by the factor  $\frac{\exp(-k(W_{t+1} + G(d_{t+1}, X_{t+1})))}{E_t[\exp(-k(W_{t+1} + G(d_{t+1}, X_{t+1})))]}$ . In addition,*

$$\kappa_1 \theta_t = E_t^d [R_{t+1}^F] \quad (37)$$

$$P_t = R_f^{-1} E_t^d P_{t+1} \quad (38)$$

where  $P_t$  is the price of the option (equivalently,  $0 = E_t^d R_{t+1}^O$ ).



**Proof.** The value function and budget constraint satisfy

$$V_t = \max_{C_t, q_t, \theta_t} -\gamma^{-1} \exp(-\gamma C_t) + \rho E_t V_{t+1} \quad (39)$$

$$W_{t+1} = (W_t - C_t) R_f + q_t (P_{t+1} - R_f P_t) + \theta_t R_{t+1}^F - \frac{\kappa_1}{2} \theta_t^2 \quad (40)$$

Now guess that

$$V_t = -k^{-1} \exp(-k(W_t + G_t))$$

for some variable  $G_t$  that is exogenous to the dealers, and where

$$k = \gamma \frac{R_f - 1}{R_f} \quad (41)$$

We have

$$\frac{\partial}{\partial W_t} V_t = -k V_t \quad (42)$$

$$\text{and } \frac{dW_{t+1}}{dC_t} = -R_f \quad (43)$$

So then the FOC for consumption under this guess is

$$0 = \exp(-\gamma C_t) + k R_f \rho E_t V_{t+1} \quad (44)$$

Noting that

$$V_t = -\gamma^{-1} \exp(-\gamma C_t) + \rho E_t V_{t+1} \quad (45)$$

$$\rho E_t V_{t+1} = V_t + \gamma^{-1} \exp(-\gamma C_t) \quad (46)$$

We have

$$\exp(-\gamma C_t) = \exp(-k(W_t + G_t)) \quad (47)$$

Now consider the FOC with respect to  $\theta_t$ . First,

$$\frac{dW_{t+1}}{d\theta_t} = R_{t+1}^F - \kappa_1 \theta_t \quad (48)$$

And hence the FOC is

$$0 = \rho E_t \left[ \exp(-k(W_{t+1} + G_{t+1})) (R_{t+1}^F - \kappa_1 \theta_t) \right] \quad (49)$$

$$\kappa_1 \theta_t = E_t^d [R_{t+1}^F] \quad (50)$$

where  $E^d$  is the expectation under the risk-neutral measure, which is the physical measure distorted by the factor

$$\frac{\exp(-k(W_{t+1} + G_{t+1}))}{E_t[\exp(-k(W_{t+1} + G_{t+1}))]} \quad (51)$$

Next, for  $q_t$ ,

$$\frac{dW_{t+1}}{dq_t} = R_{t+1}^O \quad (52)$$

So then

$$0 = \rho E_t \left[ \exp(-k(W_{t+1} + G_{t+1})) R_{t+1}^O \right] \quad (53)$$

$$0 = R_f^{-1} E_t^d R_{t+1}^O \quad (54)$$

It is straightforward to get a recursion for  $G_t$  by following the derivation in GPP. ■

**Proposition 6** *The effect of options demand on prices is*

$$\frac{\partial P_t}{\partial q_t} = -\gamma (R_f - 1) E_t^d \left[ \left( R_{t+1}^O - \hat{\beta}_t R_{t+1}^F \right) P_{t+1} \right] \quad (55)$$

where

$$\hat{\beta}_t \equiv \beta_t^F \frac{E_t^d \left[ (R_{t+1}^F)^2 \right]}{E_t^d \left[ (R_{t+1}^F)^2 \right] + k^{-1} R_f \kappa_1} \quad (56)$$

$$\beta_t^F \equiv \frac{\text{cov}_t^d (R_{t+1}^F, R_{t+1}^O)}{\text{var}_t^d (R_{t+1}^F)} \quad (57)$$

**Proof.** Based on the analysis from the previous proof, the pricing kernel can be written as

$$m_{t+1}^d \equiv \frac{\exp(-k(\theta_t R_{t+1}^F + q_t R_{t+1}^O + G_{t+1}))}{R_f E_t \exp(-k(\theta_t R_{t+1}^F + q_t R_{t+1}^O + G_{t+1}))} \quad (58)$$

Differentiate  $m_{t+1}^d$  with respect to  $q_t$  to get

$$\begin{aligned}
\frac{\partial m_{t+1}^d}{\partial q_t} &= \frac{-k \left( R_{t+1}^O + R_{t+1}^F \frac{\partial \theta_t}{\partial q_t} \right) \exp \left( -k \left( \theta_t R_{t+1}^F + q_t R_{t+1}^O + G_{t+1} \right) \right)}{R_f E_t \exp \left( -k \left( \theta_t R_{t+1}^F + q_t R_{t+1}^O + G_{t+1} \right) \right)} \\
&\quad - \frac{\exp \left( -k \left( \theta_t R_{t+1}^F + q_t R_{t+1}^O + G_{t+1} \right) \right)}{\left( R_f E_t \exp \left( -k \left( \theta_t R_{t+1}^F + q_t R_{t+1}^O + G_{t+1} \right) \right) \right)^2} E_t \left[ -k R_f P_{t+1} \exp \left( -k \left( \theta_t R_{t+1}^F + q_t R_{t+1}^O + G_{t+1} \right) \right) \right] \\
&= -k \left( R_{t+1}^O + R_{t+1}^F \frac{\partial \theta_t}{\partial q_t} \right) m_{t+1}^d - \frac{\exp \left( -k \left( \theta_t R_{t+1}^F + q_t R_{t+1}^O + G_{t+1} \right) \right)}{R_f E_t \exp \left( -k \left( \theta_t R_{t+1}^F + q_t R_{t+1}^O + G_{t+1} \right) \right)} E_t \left[ -k R_f R_{t+1}^O \frac{\exp \left( -k \left( \theta_t R_{t+1}^F + q_t R_{t+1}^O + G_{t+1} \right) \right)}{R_f E_t \exp \left( -k \left( \theta_t R_{t+1}^F + q_t R_{t+1}^O + G_{t+1} \right) \right)} \right] \\
&= -k \left( R_{t+1}^O + R_{t+1}^F \frac{\partial \theta_t}{\partial q_t} \right) m_{t+1}^d
\end{aligned} \tag{5}$$

Next, we differentiate the first-order condition for  $\theta_t$  with respect to  $q_t$ ,

$$\kappa_1 \frac{\partial \theta_t}{\partial q_t} = E_t \left[ \frac{\partial}{\partial q_t} m_{t+1}^d R_{t+1}^F \right] \tag{63}$$

$$= E_t \left[ -k \left( R_{t+1}^O + R_{t+1}^F \frac{\partial \theta_t}{\partial q_t} \right) R_{t+1}^F m_{t+1}^d \right] \tag{64}$$

Solving for the derivative yields

$$k^{-1} \kappa_1 \frac{\partial \theta_t}{\partial q_t} = E_t \left[ -R_{t+1}^O R_{t+1}^F m_{t+1}^d - R_{t+1}^F \frac{\partial \theta_t}{\partial q_t} R_{t+1}^F m_{t+1}^d \right] \tag{65}$$

$$k^{-1} \kappa_1 \frac{\partial \theta_t}{\partial q_t} + \frac{\partial \theta_t}{\partial q_t} R_f^{-1} E_t^d \left[ \left( R_{t+1}^F \right)^2 \right] = R_f^{-1} E_t^d \left[ -R_{t+1}^O R_{t+1}^F \right] \tag{66}$$

$$\frac{\partial \theta_t}{\partial q_t} = - \frac{E_t^d \left[ R_{t+1}^O R_{t+1}^F \right]}{E_t^d \left[ \left( R_{t+1}^F \right)^2 \right] + k^{-1} R_f \kappa_1} \equiv -\beta_t^F \frac{E_t^d \left[ \left( R_{t+1}^F \right)^2 \right]}{E_t^d \left[ \left( R_{t+1}^F \right)^2 \right] + k^{-1} R_f \kappa_1} \tag{67}$$

where

$$\beta_t^F \equiv \frac{E_t^d \left[ R_{t+1}^O R_{t+1}^F \right]}{E_t^d \left[ \left( R_{t+1}^F \right)^2 \right]} = \frac{\text{cov}_t^d \left( R_{t+1}^F, R_{t+1}^O \right)}{\text{var}_t^d \left( R_{t+1}^F \right)} \tag{68}$$

The price sensitivity comes from differentiating the pricing equation for the option

$$\begin{aligned}
\frac{\partial P_t}{\partial q_t} &= E_t \left[ \frac{\partial m_{t+1}^d}{\partial q_t} P_{t+1} \right] \\
&= -k E_t \left[ \left( R_{t+1}^O + R_{t+1}^F \frac{\partial \theta_t}{\partial q_t} \right) m_{t+1}^d P_{t+1} \right] \\
&= -k E_t \left[ \left( R_{t+1}^O - \hat{\beta}_t R_{t+1}^F \right) m_{t+1}^d P_{t+1} \right] \\
&= -\gamma (R_f - 1) E_t^d \left[ \left( R_{t+1}^O - \hat{\beta}_t R_{t+1}^F \right) R_{t+1}^O \right]
\end{aligned}$$

where the last line uses the fact that  $E_t^d [R_{t+1}^O] = E_t^d [R_{t+1}^F] = 0$  since they are excess returns that are fairly priced under the pricing measure  $d$ . ■

## D.2 Proof of proposition 4

The proof involves simply analyzing the expectation in 6 above. We have

$$E_t^d \left[ \left( R_{t+1}^O - \hat{\beta}_t R_{t+1}^F \right) R_{t+1}^O \right] = E_t^d \left[ \left( R_{t+1}^O - \beta_t^F R_{t+1}^F - \left( \hat{\beta}_t - \beta_t^F \right) R_{t+1}^F \right) R_{t+1}^O \right] \quad (69)$$

$$= E_t^d \left[ \left( \varepsilon_{t+1}^F - \left( \hat{\beta}_t - \beta_t^F \right) R_{t+1}^F \right) \left( \beta_t^F R_{t+1}^F + \varepsilon_{t+1}^F \right) \right] \quad (70)$$

$$= \text{var}_t^d \left[ \varepsilon_{t+1}^F \right] - \left( \hat{\beta}_t - \beta_t^F \right) \beta_t^F E_t^d \left[ \left( R_{t+1}^F \right)^2 \right] \quad (71)$$

$$= \text{var}_t^d \left[ \varepsilon_{t+1}^F \right] - \left( \frac{-k^{-1} R_f \kappa_1}{E_t^d \left[ \left( R_{t+1}^F \right)^2 \right] + k^{-1} R_f \kappa_1} \right) \left( \beta_t^F \right)^2 E_t^d \left[ \left( R_{t+1}^F \right)^2 \right]$$

Next, we want to further decompose  $\text{var}_t^d \left[ \varepsilon_{t+1}^F \right]$ . We have

$$R_{t+1}^F = R_{t+1}^I + z_{t+1} \quad (73)$$

$$\beta_t^F = \beta_t^I \frac{\sigma_{I,t}^2}{\sigma_{I,t}^2 + \sigma_{z,t}^2} \quad (74)$$

where  $\sigma_{I,t}^2 = \text{var}_t^d (R_{t+1}^I)$ . We can write

$$R_{t+1}^O = \beta_t^I R_{t+1}^I + \varepsilon_{t+1}^I \quad (75)$$

where  $\beta_t^I$  is the ( $d$ -measure) regression coefficient. Then

$$\varepsilon_{t+1}^F = R_{t+1}^O - \beta_t^F R_{t+1}^F \quad (76)$$

$$= \beta_t^I R_{t+1}^I + \varepsilon_{t+1}^I - \beta_t^I \frac{\sigma_{I,t}^2}{\sigma_{I,t}^2 + \sigma_{z,t}^2} (R_{t+1}^I + z_{t+1}) \quad (77)$$

$$= \beta_t^I \frac{\sigma_{z,t}^2}{\sigma_{I,t}^2 + \sigma_{z,t}^2} R_{t+1}^I + \varepsilon_{t+1}^I - \beta_t^I \frac{\sigma_{I,t}^2}{\sigma_{I,t}^2 + \sigma_{z,t}^2} z_{t+1} \quad (78)$$

$$\text{var}_t^d [\varepsilon_{t+1}^F] = (\beta_t^I)^2 \frac{\sigma_{z,t}^2 \sigma_{I,t}^2}{\sigma_{I,t}^2 + \sigma_{z,t}^2} + \sigma_{\varepsilon,t}^2 \quad (79)$$

where  $\sigma_{\varepsilon,t}^2 \equiv \text{var}_t^d [\varepsilon_{t+1}^I]$ .

Up to first order in  $\kappa_1$  and  $\sigma_z^2$ ,

$$\frac{\partial P_t}{\partial q_t} = -\gamma (R_f - 1) \left( \underbrace{(\beta_t^I)^2 \sigma_{z,t}^2}_{\text{Basis risk}} + \underbrace{\sigma_{\varepsilon,t}^2}_{\text{Unhedgable risk}} \right) + \underbrace{\kappa_1 R_f^2 (\beta_t^I)^2}_{\text{Imperfect hedging}} \quad (80)$$

## E Additional figures

Figure A.1: Synthetic put returns for various strikes

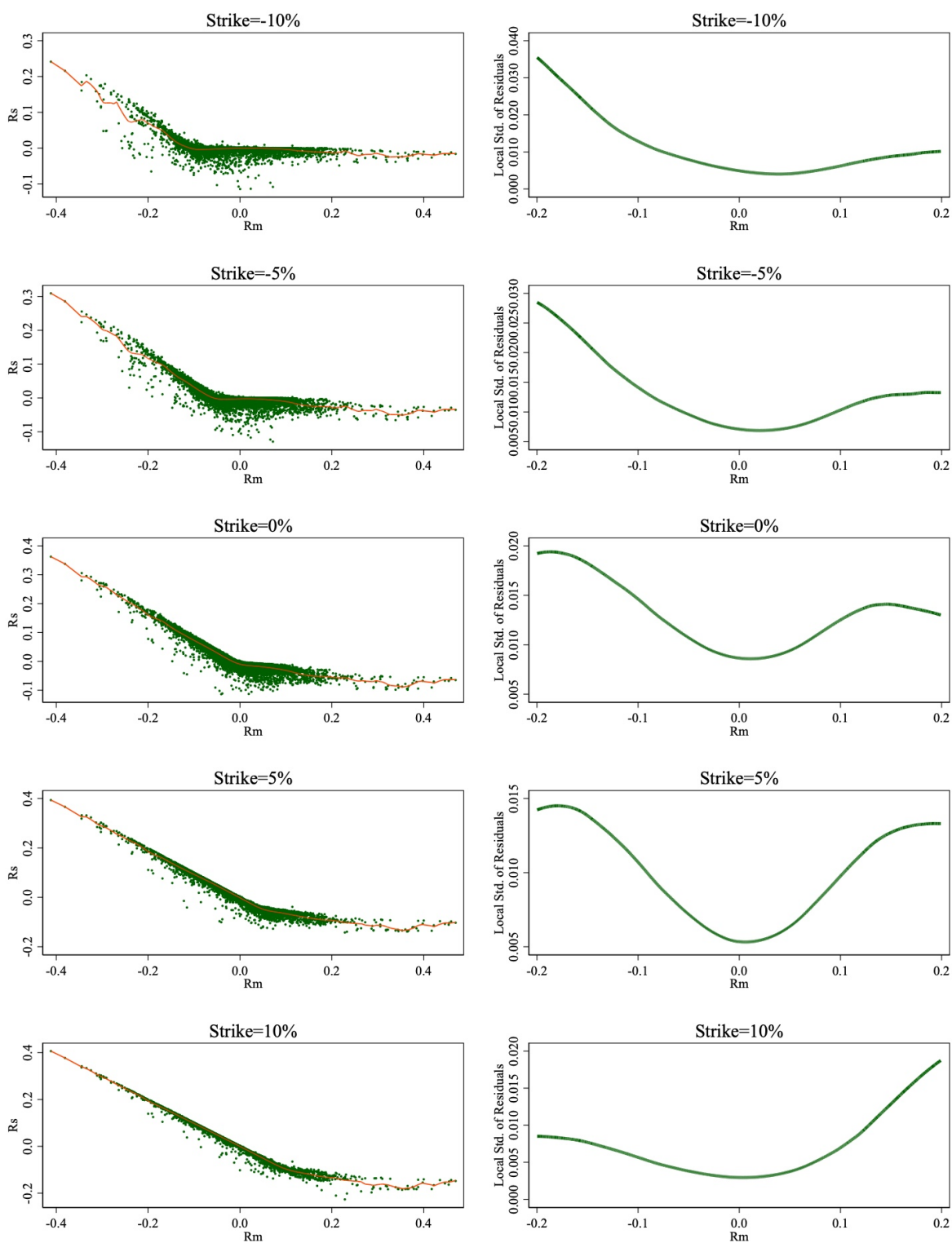
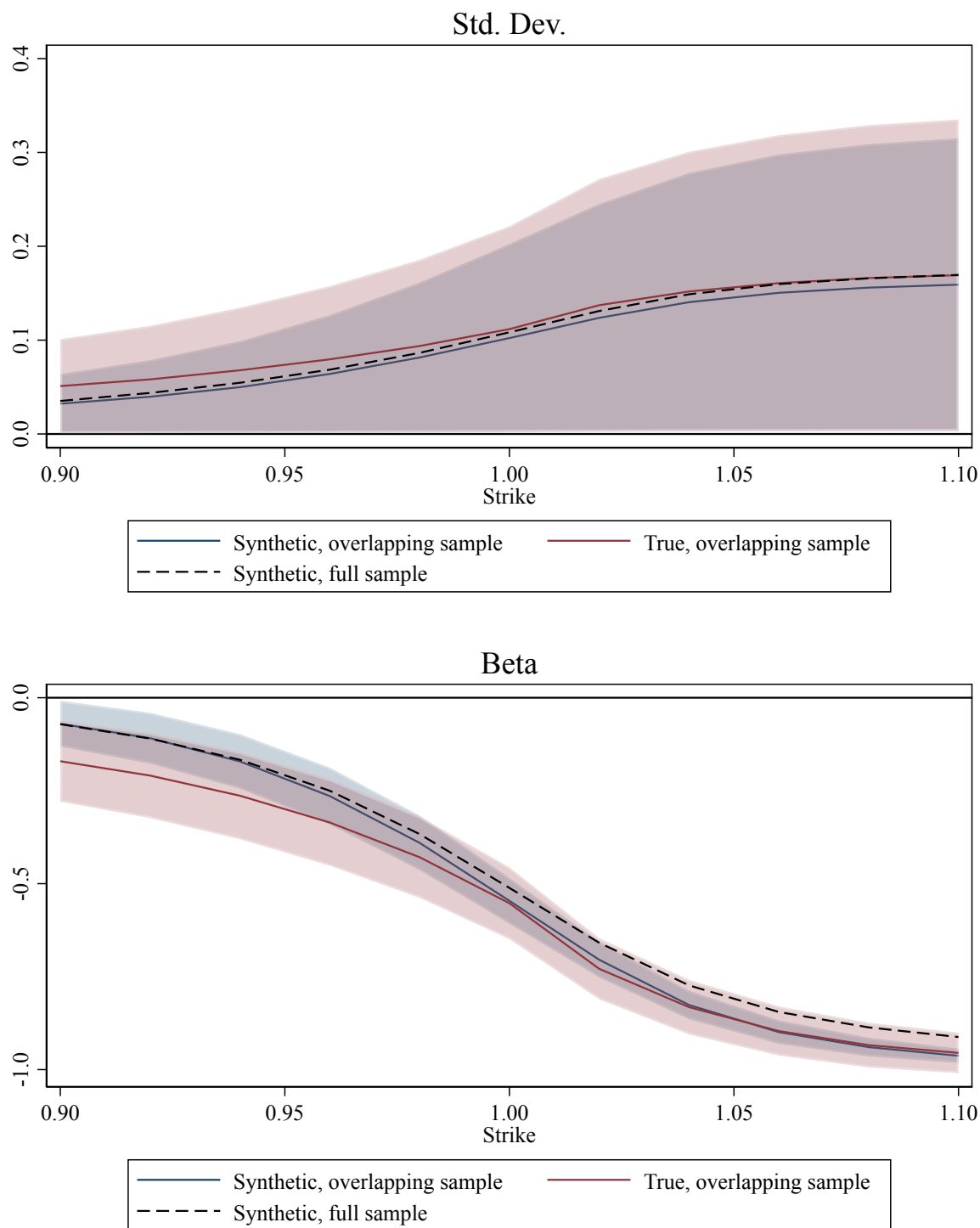
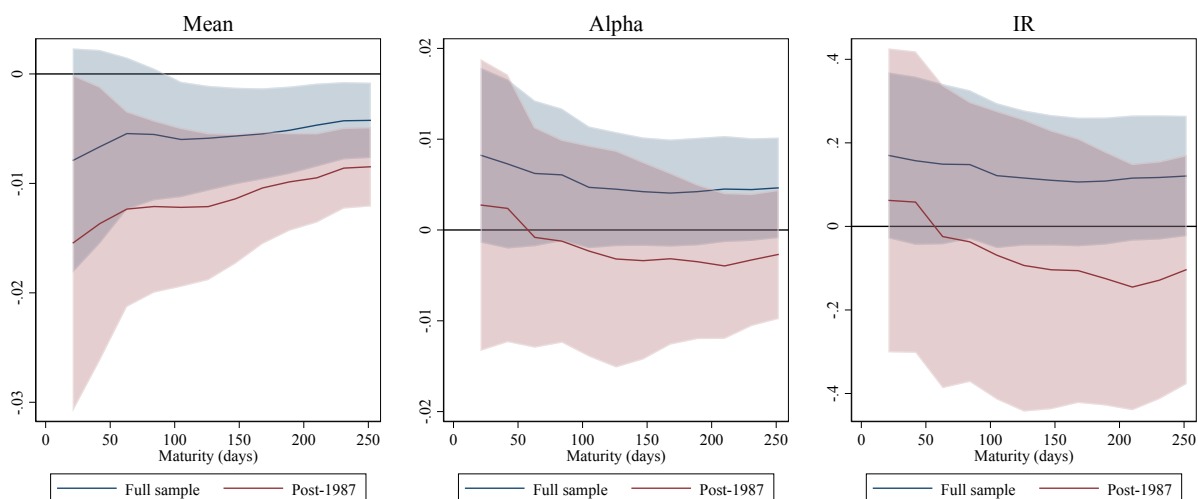


Figure A.2: Risk measures for synthetic and traded options.



**Note:** Standard deviations and CAPM betas for synthetic and traded option returns. The shaded regions are 95% confidence intervals (no confidence interval is shown for the dotted line)

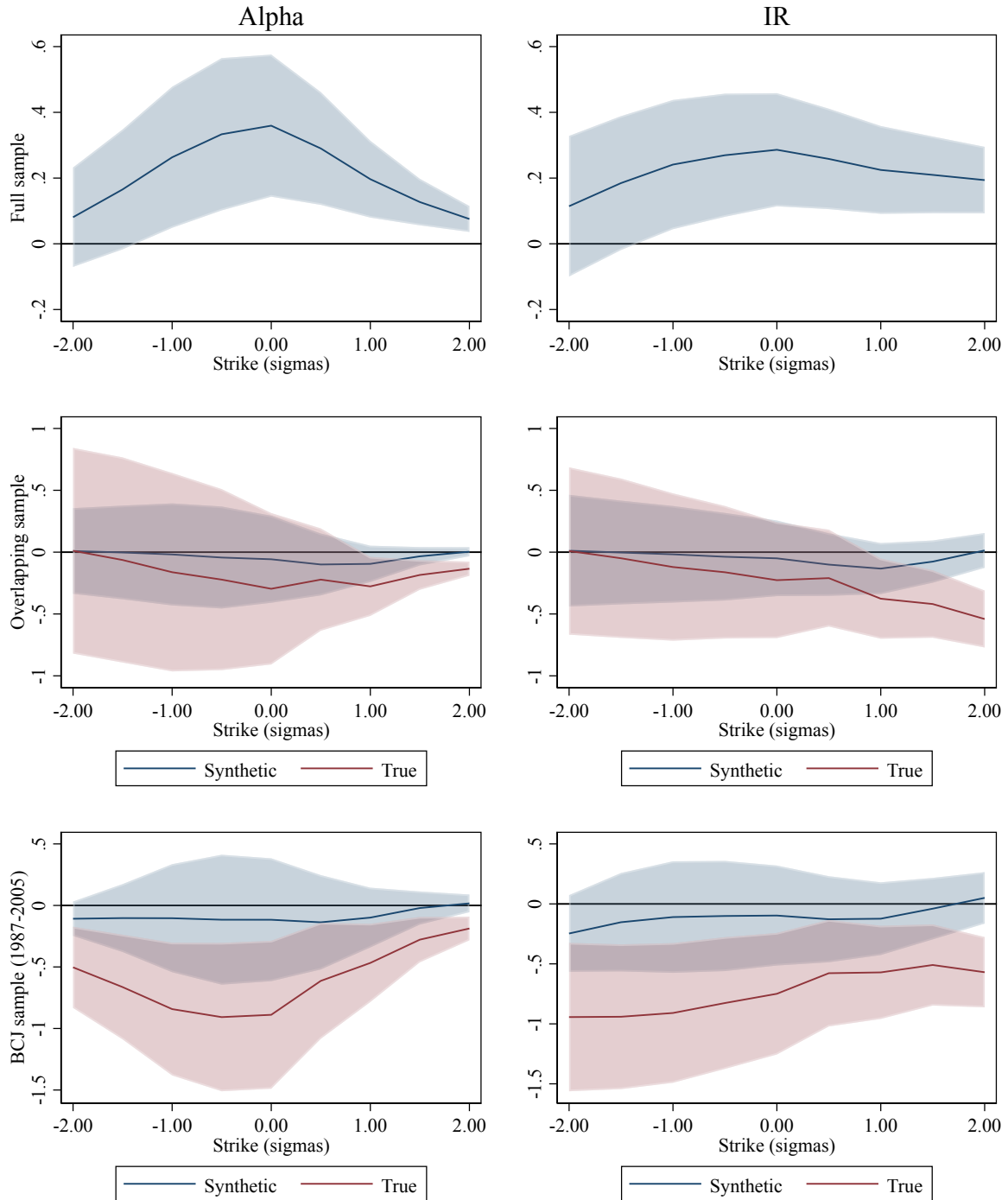
Figure A.3: Synthetic option returns across maturities



**Note:** Average return, CAPM alpha, and information ratio for synthetic options across maturities. In all cases, the strike is set to be equivalent to -5% at the monthly maturity (scaling with the square root of the maturity). Note that the two series plotted are not traded versus synthetic options but rather synthetic options in two different samples (the longer maturities are much less liquid for the traded options, especially early in the sample, and present a number of issues in implementation).

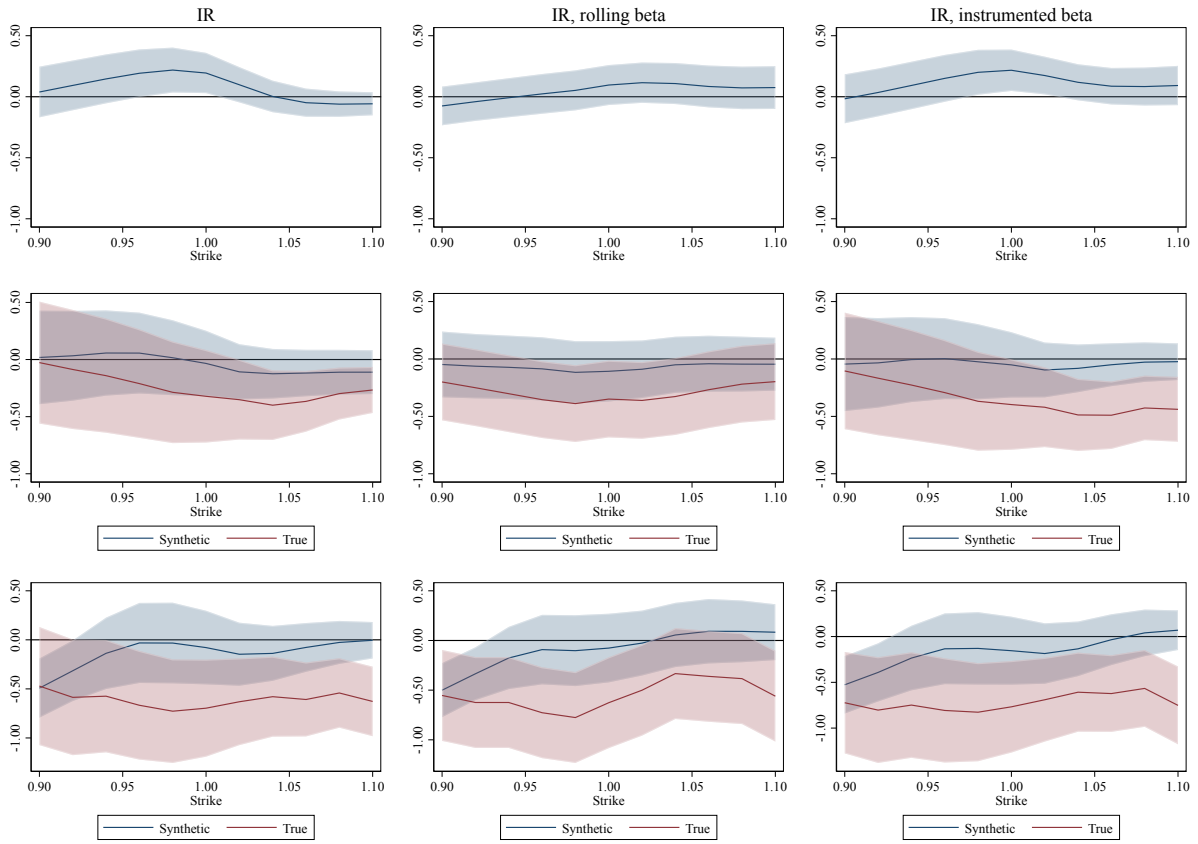


Figure A.4: Option returns for moneyness in volatility units



**Note:** Average return, CAPM alpha, and information ratio for synthetic and traded options. Strikes here are selected in units of volatility instead of as a fixed percentage of the price of the underlying.

Figure A.5: Information ratios under various specifications for beta



**Note:** Information ratios under various specifications for the betas. The left-hand column uses the full-sample beta, which is the benchmark specification. The middle column uses betas estimated from a rolling three-month window. Finally, the right-hand column instruments for the conditional beta, as described in the text.

Figure A.6: Synthetic versus traded option returns

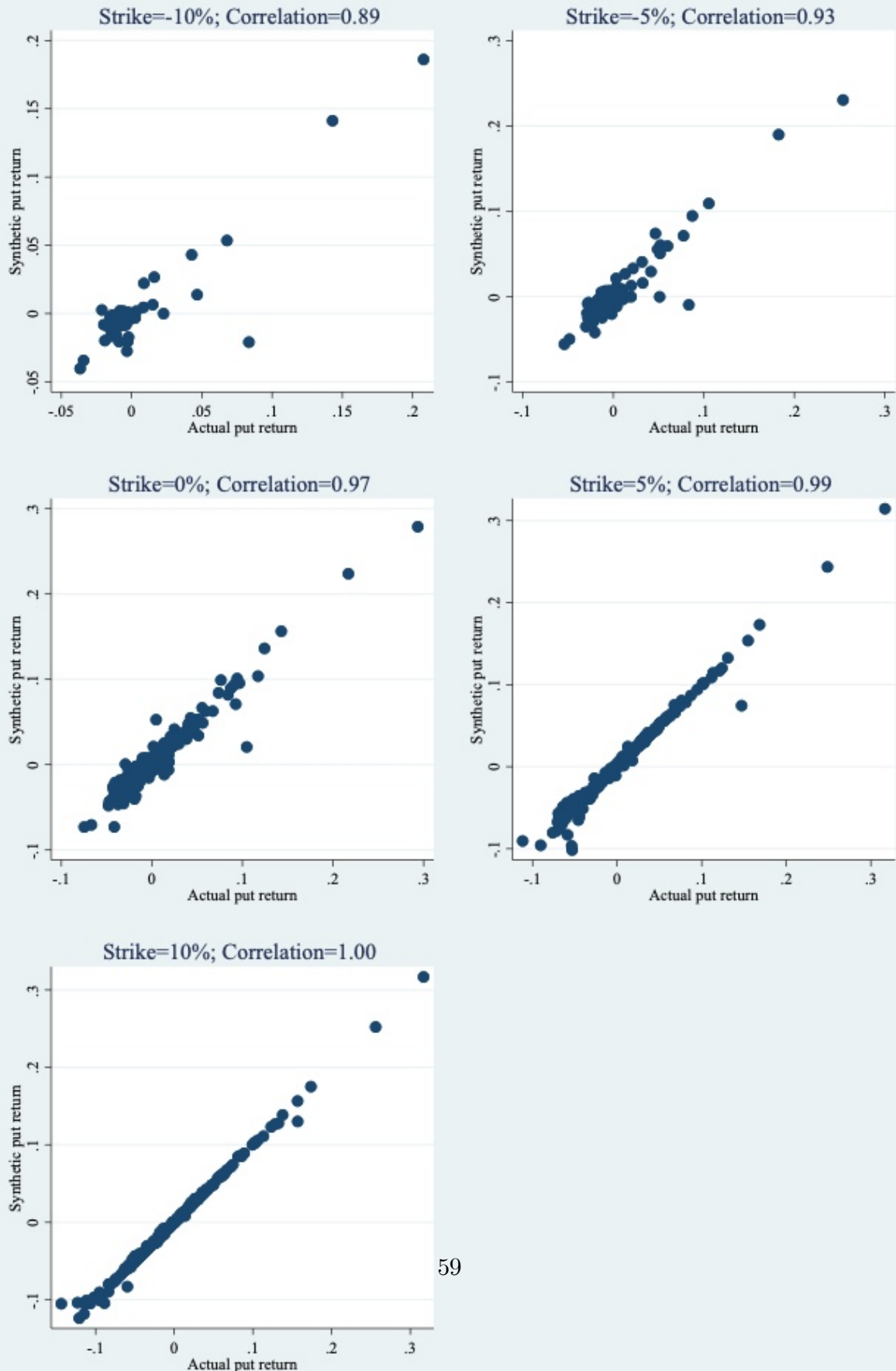
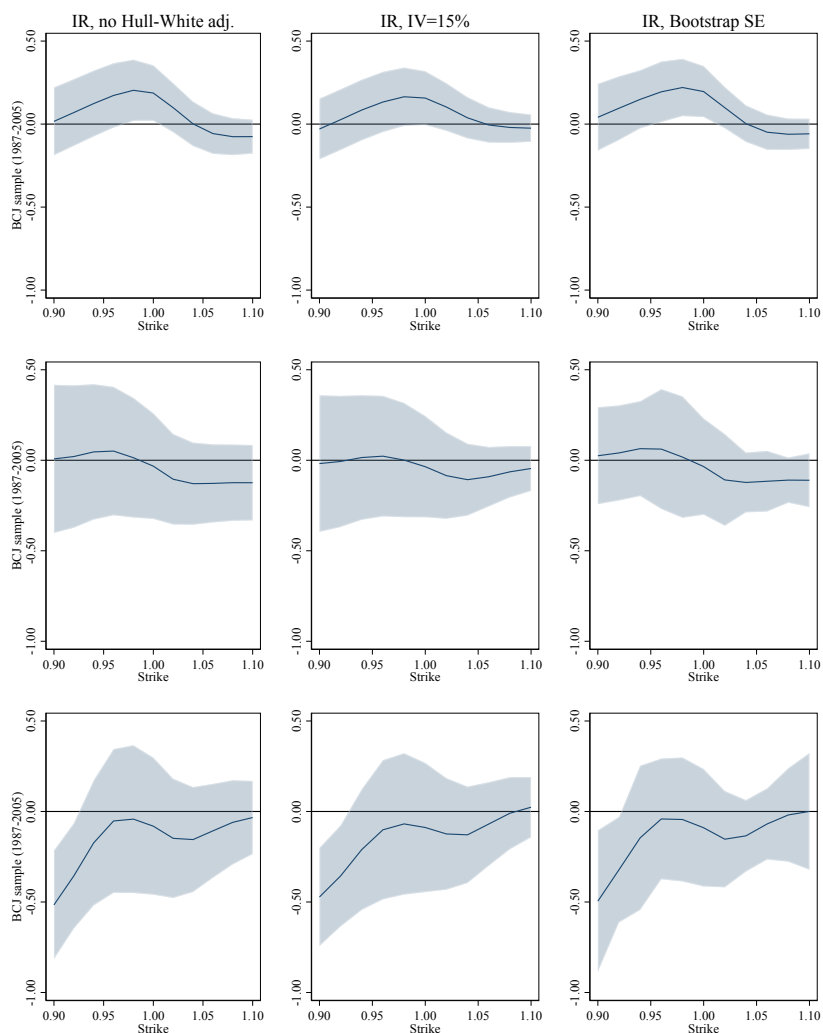


Figure A.7: Robustness for information ratios



**Note:** The figure reports the estimated SMU obtained by using options of all strikes jointly. The green line uses synthetic puts (sample 1926-2021), and the red line uses actual puts (sample 1987-2021).