

Portfolio Selection with Options^{*}

Semyon Malamud[†]

February 5, 2014

^{*}Financial Support of INQUIRE Europe is gratefully acknowledged.

[†]SFI EPFL, Quartier UNIL Dorigny, 1015 Lausanne. e-mail: semyon.malamud@epfl.ch. Tel: +41 79 599 10 28

Abstract

I introduce dynamic option trading and non-linear views into the classical portfolio selection problem. The optimal dynamic option portfolio is characterized explicitly in terms of its expected sensitivities (Greeks) and the role of the mean-variance efficient portfolio is played by the “*Greek-efficient*” portfolio. This is the portfolio that has the optimal sensitivities to chosen risk factors. I test these portfolios empirically and find that options significantly improve the risk-return profile due to predictability of powers (and other non-linear functions) of returns which allows for *optimal management of non-linear views*. To test the effects of higher moments on portfolio choice, I compute (both analytically and numerically) the Greek efficient portfolios for a CRRA investor and find that accounting for higher moments may have ambiguous effects on the optimal tail risk. In fact, even Greek efficient portfolios for a mean-variance investor already offer a highly attractive skewness-kurtosis profile. In the presence of transactions costs that depend on an option’s moneyness and maturity, optimal state-contingent option portfolios are characterized in terms of state prices Greeks as well as a new object, the *transactions costs Greeks*.

1 Introduction

Modern portfolio choice theory is dominated by the paradigm of mean-variance efficiency, introduced by Markowitz (1952): given the investor’s views/beliefs about risks and expected returns, the optimal asset allocation is determined by the optimal risk-return tradeoff; in the simplest, linear-quadratic one period case, the optimal tradeoff is characterized by mean variance efficiency. Various extensions have been proposed since then, accounting for dynamic trading, fluctuating risk-return tradeoffs, unhedgeable risks, and general, non-quadratic preferences.

Introducing options into the portfolio choice problem makes the applicability of the standard approach highly non-trivial. First of all, the number of assets increases dramatically: for some liquid instruments, such as, e.g., S&P 500, the number of publicly traded options can easily go above 40, and this number increases essentially to infinity if one accounts for over-the-counter derivatives. One therefore quickly runs into the problems of estimating high-dimensional covariance matrices, with many assets being highly correlated. Furthermore, selecting an optimal portfolio of options on the same underlying based on their covariance matrix does not make much sense, as they all provide exposure to very similar risks. Indeed, options are designed precisely to capture the non-linear structure of probability distributions and allow an investor to express views on the underlying that go beyond the simple mean-variance tradeoff. Finance practitioners are well aware of this; one would never hear an option trader talk about the covariance matrix of options or mean-variance efficiency. Instead, option traders think in terms of Greeks; that is, sensitivities, or, derivatives, of option prices with respect to various risk factors. If an option trader wants to express his views, he optimally decides which sensitivities he would like to eliminate, so that the resulting option portfolio provides optimal exposure only to the desired Greeks. In analogy with the mean-variance efficiency, one

could call such a portfolio “*Greek-efficient*”. Since the number of all possible Greeks³ is infinite, option traders typically select the desired order of Greek efficiency, and all Greeks (sensitivities) higher than this order are simply ignored due to their (supposedly) insignificant impact on the portfolio performance.

Though Greek-efficiency is a very intuitive concept and is in one way or another used by almost any option trader, it has never been formalized analytically. Without such a formalization, option traders often have to make discretionary decisions, based on their intuition and experience.⁴ This approach may work well for options on a single underlying, but is almost infeasible if one wants to trade multiple options on multiple underlyings. What is needed is a theory of Greek-efficiency, that should be a direct analog of the theory of mean-variance efficiency, taking into account the possibility of making state-contingent bets.

The goal of this paper is to develop such a theory. I develop a techniques for approximating the optimal option portfolio for an investor who can trade options on several of the underlying instruments (e.g., stocks, bonds, exchange rates, volatility, commodities). By construction, the approximation is designed a match a desired degree of Greek efficiency: the maximal order of Greeks that matter for the investor. Though options allow the investor to make bets contingent on any realization of each of the underlying instruments, I assume that there are no options for making “joint bets”; i.e., bets on the joint realization of several instruments.⁵ I provide an explicit characterization of Greek efficient optimal portfolios, for any given order of Greek-efficiency, covering both mean-variance and CRRA preferences, as well as quadratic transaction costs.

The high added value of Greek efficient portfolio comes from the predictability of powers and, more generally, non-linear functions of returns. In particular, options, and, more precisely, Greek efficient portfolios allow an investor to perform *optimal views management*. That is, select state contingent claims that exploit the views on the distribution of returns in an optimal fashion. I test my theoretical results on three liquid stock indices, S&P 500, Nasdaq 100, and Russell 2000. Exploting the (in-sample) predictability of powers of returns, I construct various Greek efficient portfolios and find a very strong performance out-of-sample, even after transactions costs. Naturally, Greek efficient returns are highly non-Gaussian and have very heavy tails. In particular, the kurtosis of Greek efficient returns can be very high (of the order of 30). However, at the same time Greek efficient returns are also highly positively skewed (with skewness above 4). This

³That is, mixed partial derivatives with respect to all risk factors.

⁴For example, it is a common wisdom that a straddle position (a call and a put with the same strike) provides a good delta-neutral exposure to volatility, and that is why many option traders use it.

⁵This is a very important form of market incompleteness that I observe in modern financial markets. For example, there are liquid options traded on S&P 500 and options on VIX, but there are no liquid options traded on the joint realization of S&P and its volatility. Even for pairs of large stock indices, options on their relative prices are typically illiquid and only trade OTC. For large stocks, basket options on their weighted average prices are often available, but, typically, for only a single collection of weights, which still makes it impossible to make state-contingent joint bets.

naturally raises the question of whether an investor with preferences for higher moments would like to significantly alter the risk-return profile of the Greek efficient portfolio. To answer this question, I develop new analytical techniques for studying efficient portfolios accounting for skewness and kurtosis and derive explicit formulae for the such portfolios. I then test this formula on the same stock index data and find that the Greek efficient portfolio for a CRRA investor indeed outperforms that for a mean-variance investor. However, the CRRA-efficient portfolio composition does not significantly differ from that of a mean-variance investor due to the highly attractive skewness-kurtosis profile of the latter portfolio. Interestingly enough, increasing the order of Greek efficiency by including claims on higher powers of returns significantly reduces kurtosis of Greek efficient returns. This means that using kurtosis as a measure of “non-Gaussianity” of returns on a trading strategy may be misleading: more non-linearity may actually lead to a lower kurtosis.

Having investigated the portfolio choice problem in the frictionless market, I introduce quadratic transaction costs into the model using the approach of Garleanu and Pedersen (2012). Transaction costs are extremely important for option trading, particularly because they fluctuate stochastically together with the moneyness and maturity of an option, and may be quite high for out-of-the-money (OTM) options. For example, transactions costs may vary from 5-10% for at-the-money (ATM) options to 40-50% for deeply OTM options. This means that an investor who bought a relatively liquid ATM option faces liquidity risk if the option suddenly goes out of the money and becomes illiquid. This naturally leads to non-trivial adjustments in the optimal portfolio, driven by the desire to hedge future fluctuations in transaction costs. It turns out that the approximate optimal portfolio can be characterised explicitly in terms of the Greeks of the state prices, as well as a new object that I call “*transaction costs Greeks*”. I test the formula for optimal portfolios with transactions costs using option data on three stock indices⁶ and find that it significantly improves the performance. This is driven by two important mechanisms: (1) stock index option returns can be very well approximated using *transactions costs adjusted Black-Scholes Greeks*; (2) options with longer maturity have lower transactions costs. Item (1) implies that we can relatively easily attain the desired Greek exposure with longer maturity options. Item (2) implies that the noise that is present in the approximation from (1) is partially offset by the lower transaction costs.

Furthermore, I also extend the analysis of Garleanu and Pedersen (2012) to account for arbitrary investor preferences by directly incorporating target portfolios into the optimization problem.

The paper is organized as follows. Section 2 discusses related literature. Section 3 describes model setup for the frictionless mean-variance case. Section 4 studies the frictionless problem for a CRRA investor. Section 5 tests the portfolios of Sections 3-4 empirically. Section 6 solves the dynamic mean-variance problem with quadratic transaction costs, and Section 7 develops the general results for target portfolios.

⁶S&P 500, Nasdaq 100, and Russell 2000.

2 Literature Review

Though the idea of using option-implied state prices for portfolio choice is very natural and intuitive, there is little academic literature on this topic. The only paper I am aware of that directly follows this approach is Ait-Sahalia and Brandt (2007).⁷ In their model, time is continuous, markets are dynamically complete, and hence all options are redundant. Then, they study theoretically how a consumer could optimally exploit differences in state prices across the states of the world. By contrast, in my model markets are incomplete and time is discrete. The assumed form of incompleteness (namely, the impossibility of fully state contingent bets on joint realizations of multiple instruments) closely resembles that in the real-world financial markets.

In the benchmark case that does not account for transactions costs, I assume that investors are myopic and consider both the mean-variance optimization and a more complicated portfolio choice problem of a CRRA investor. See, Ait-Sahalia and Brandt (2001), Campbell and Viceira (2002), Brandt, Goyal, Santa-Clara and Stroud (2005), Brandt and Santa-Clara (2006), Bansal and Kiku (2007), and Brandt, Santa-Clara and Valkanov (2009) for a similar approach. Due to market incompleteness and the possibility of bets contingent on the realization of a given trading instruments, even the one period Markowitz problem becomes highly non-trivial, and only approximate solutions for optimal option portfolios can be computed. I show that, given the degree of Greek-efficiency that the investor wants to achieve, the Greeks of the approximate optimal option portfolio can be computed by a simple formula that is similar to the Markowitz solution for the mean-variance efficient portfolio. The only difference is that the role of the covariance matrix is played by the matrix of all joint moments of the existing (both traded and non-traded) risk factors.⁸

This paper is also related to the general literature on portfolio choice in dynamic environments with stochastically varying investment opportunity sets. Several papers provide fully explicit solutions for optimal portfolios. See, e.g., Merton (1971), Kim and Omberg (1996), Campbell and Viceira (1999), Wachter (2002), Brennan and Xia (2002), Cvitanic, Martellini and Zapatero (2006), and Liu (2007). None of these papers studies the role of options in portfolio allocation. In general, obtaining fully-explicit, closed form solutions to dynamic portfolio choice problems with stochastic investment opportunities is a highly non-trivial problem, and one would need to resort to numerical methods, such

⁷Some papers study how one can empirically extract information about the joint distribution of securities' returns and use this information to improve portfolio selection. See, e.g., DeMiguel et al. (2013).

⁸In the main text, I only consider myopic mean-variance objective function for the investor. In the appendix, I extend the results to the multi-period mean-variance optimization, using the time-consistent version of the dynamic mean-variance optimization problem, developed in Basak and Chabakauri (2010). Note that dynamic mean-variance portfolio choice problem is time inconsistent. Before Basak and Chabakauri (2010), all papers assumed the so-called pre-commitment optimal portfolio that is chosen at time zero. See, Duffie and Richardson (1991), Bajeux-Besnainou and Portait (1998), Leippold, Trojani and Vanini (2004), Brandt and Santa-Clara (2006), Cochrane (2008), and Brandt (2009).

as those proposed by Detemple, Garcia and Rindisbacher (2003), Cvitanic, Goukasian and Zapatero (2003), and Brandt, Goyal, Santa-Clara and Stroud (2005).

Main theoretical results of my paper provide explicit analytic approximations to the optimal option portfolio. The techniques used to derive this approximation are based on Taylor expansions, which makes my paper indirectly linked to the large literature on Taylor approximations in portfolio choice. See, e.g., Samuelson (1970), Loistl (1976), Kroll, Levy, and Markowitz (1984), Markowitz (1991), Hlawitschka (1994), Brandt, Goyal, Santa-Clara and Stroud (2005), Garlappi and Skoulakis (2010). All these papers use Taylor expansions to approximate the unknown value function and then use this approximation to solve for the optimal portfolio. None of these papers study optimal option portfolios. My approach is very different. Due to the availability of fully state contingent bets, investor's problem is infinitely dimensional and therefore, even in the one-period mean-variance case, approximate methods have to be developed. Effectively, the investor would like to optimally choose all possible Greeks of the option portfolio, and there is an infinite number of those. Therefore, I use Taylor expansion of the state prices (risk-neutral probabilities) in order to reduce the dimensionality of the problem and make it finite-dimensional, with the dimension equal to the number of Greeks the investor would like to optimize over. I.e., the order of "Greek efficiency".

Several papers study the problem of dynamic portfolio choice with options. Liu and Pan (2003) derive an analytical expression for the optimal portfolio for a CRRA investor trading one stock and one OTM put option. Their analysis is restrictive because it is based on a parametric model that needs to be properly estimated. By contrast, my results are independent on the number of assets that are being traded, and are completely non-parametric.

Eraker (2007) studies mean-variance portfolio allocation with a single stock and three assets: ATM straddles, OTM puts, and OTM calls. It is a-priori not clear, why this particular choice of option structures is optimal. By contrast, the concept of Greek efficiency developed in my paper provides a theoretical justification of the type of option structures that should be used in portfolio allocation with options.

Plyakha and Vilkov (2008) use the parametric portfolio optimization approach of Brandt, Santa-Clara and Valkanov (2009) and show that the shape of the implied volatility surface contains information about future option returns. As in Liu and Pan (2003) and Eraker (2007), they use specific types of option structures for forming their portfolios and find that this information can be exploited to improve portfolio performance. However, they also find that transaction costs essentially destroy all the gains from using options if rebalancing is done at monthly frequency. By contrast, I find that Greek efficient portfolio provide very large gains for an investor, even after transaction costs. The reason is that the Greek efficient portfolios are designed to optimally exploit significant predictability of powers of returns.

Faias and Santa-Clara (2011) use a (purely numerical) simulation approach to solving portfolio optimization problems for a myopic CRRA investor. Namely, using historical data on S&P 500 returns, they estimate their distribution. Simulating from this distribution allows them to forecast the distribution of option payoffs and then numerically

compute the optimal portfolio for the investor. My approach is different because it is based on analytical expressions for optimal portfolio and exploits non-linear predictability.

Finally, my paper is also related to the literature on portfolio choice with quadratic transactions costs. See, Heaton and Lucas (1996), Grinold (2006), Garleanu and Pedersen (2012,2013) and Collin-Dufresne et al. (2013). Garleanu and Pedersen (2012) were the first to develop a tractable model that allows for multiple assets, predictable returns and transactions costs. Assuming a constant covariance structure for the returns and constant transactions costs, Garleanu and Pedersen (2013) show that the optimal portfolio is a linear combination of the existing portfolio and an aim portfolio, which is a weighted average of the current Markowitz portfolio and the expected Markowitz portfolios on all future dates. In my model, I consider two ramifications of their basic approach:

(1) In the first one, I assume same form of the value function as in Garleanu and Pedersen (2013). However, I deviate from their assumptions in several important ways. First, as in the case without transactions costs, I assume that a complete set of single-instrument options is available. Second, I allow for an arbitrary no-arbitrage dynamics of option prices, and hence the covariance structure is stochastic and time-varying.⁹ Third, I assume that option transactions costs are also stochastic and time varying: Namely, transactions costs are functions of an option's moneyness and maturity.¹⁰ The cost of such a generality is that, as in the case of costless trading, I have to rely on approximations and, by contrast to Garleanu and Pedersen (2013), there is no simple and elegant characterization of optimal portfolios. Nevertheless, Greek efficient portfolios can be characterized up to solve an explicit linear system and several general insights can be derived. In particular, I show that the time variation in the transactions costs is accounted for in the optimal portfolio via "transactions costs Greeks"; that is, Greeks of specific, explicitly given option portfolios whose payoff depends on the future transactions costs.

(2) In the second one, I consider a more general problem that does not implicitly depend on the investor's utility function. Namely, I assume that investors goal is to trade as close as possible to a target portfolio, but accounting for transactions costs. The only input needed in this problem is the target portfolio, which could for example be the optimal portfolio for a CRRA investor in a frictionless market. It turns out that in this case the optimal option portfolio can be characterized fully explicitly due to a maturity rotation effect specific in portfolio choice with options: as time goes by, maturity decreases to zero, and the option stops influencing future portfolio decisions.

⁹Collin-Dufresne et al. (2013) also allow for stochastic covariance structure, but do not consider options trading.

¹⁰See, George and Longstaff (1993) and Cao and Wei (2010) for empirical evidence. In particular, transactions costs for out of the money and longer maturity options can be really huge, making it extremely costly to rebalance the positions.

3 Model Setup

Time is discrete, $t = 0, 1, \dots, T$. There is a one period risk-free bond with interest rate $1+r$, and N risky assets in the economy on which derivatives can be traded, with random returns X_i realized at time 1. Investor believes that the joint density of the returns is given by $p(X_1, \dots, X_N)$, and the unconditional densities are given by $p_i(X_i)$, $i = 1, \dots, N$. Everywhere in the sequel, $E[\cdot]$ denotes the expectation under the investor's beliefs. If the market beliefs are different from those of the investor, I will always use $E^M[\cdot]$ to denote the expectations under the market beliefs. I also use X_i to denote the expectation of X_i .

I assume that options can be traded on each of the X_i , but there are no tradable liquid options, contingent on the joint realization of multiple risks X_i . This is a very important type of market incompleteness that is present in the real world financial markets. For example, even though options on pairs of stocks exist, they are not as liquid as single name options. Furthermore, basket options (options on baskets of several stocks) are typically traded for only a single collection of weights. As another example, one can trade options on S&P 500, as well as options on its volatility (VIX). However, no liquid options exist whose payoff is contingent of the joint realization of S&P 500 and its volatility.

The state of the economy is characterized by a multi-dimensional Markov state variable X_t with a transition density $p(X_t, X_{t+1})$. The first N coordinates of this process correspond to log prices of tradable instruments. At every instant of time, the investor can trade the instruments directly, as well as liquid European call (and/or put) options of maturities

$$\theta_i \in \Theta_{it} \equiv \{\theta_{i1}, \dots, \theta_{iT}\},$$

contingent on the any of the tradable instruments $X_{i+\theta_i,t}$, $\theta_i \in \Theta_{it}$, $i = 1, \dots, N$, with arbitrary strikes. Note that the set of available maturities is allowed to change over time: in reality, new options are issued infrequently, which leads to non-stationarity in the set of available options. For simplicity, I will always assume that $\theta_i > 1$.¹¹ That is, investors always have to liquidate the options before maturity.¹²

The price of an option with strike K and maturity θ on the underlying X_i is denoted by $C_{i,t+\theta}(X_t, K)$. Note that the price of the option is allowed to depend on all of the underlying state variables in X_t , and not just on $X_{i,t}$. Using these options, an investor can express general state-contingent views on the realization of any of the tradable instruments. The Breeden-Litzenberger formula implies that the investor can directly

¹¹In the appendix I allow for $\theta_i = 1$ and show that my results still hold, but the expressions become slightly more complicated.

¹²In practice, most option traders liquidate the option before maturity. Of course if time step is sufficiently small, there is little difference between option with maturity $\theta = 1$ and $\theta = 2$. Furthermore, the maturity step θ for traded options is typically quite large (e.g., one month for S&P 500 and VIX options). This means that making explicit bets on the exact time $t + 1$ -realization of risk factors is impossible, and what matters are the option returns, that are, in turn, determined by the options' Greeks.

observe the state prices

$$\eta_{(i,\theta)}(X_t, K) = \frac{\partial^2}{\partial K^2} C_{i,t+\theta}(X_{i,t+\theta}, K).$$

The time- t price of a maturity- θ claim $f_{(i,\theta)}(X_{i,t+\theta})$ is given by $\int \eta_{(i,\theta)}(X_t, K) f_{(i,\theta)}(K) dK$. Any such claim $f_{(i,\theta)}(X)$ can be replicated by a portfolio of plain vanilla calls and puts, using the Carr-Madan (1998) formula

$$\begin{aligned} f_{(i,\theta)}(X_{i,t+\theta}) &= f(e^{\bar{X}}) + f'(\bar{X})(e^{X_{i,t+\theta}} - e^{\bar{X}}) \\ &+ \int_{-\infty}^{\bar{X}} (K - e^{X_{i,t+\theta}})^+ f_i''(K) dK + \int_{\bar{X}}^{+\infty} (e^{X_{i,t+\theta}} - K)^+ f_i''(K) dK, \end{aligned} \quad (1)$$

where \bar{X} is some reference level.¹³ Practitioners often use the term “option structure” to denote such a general nonlinear claim f_i . I will also frequently use this term in the sequel. Note that I do not need to explicitly introduce trading in the underlying itself: I have (assuming no dividends) that

$$e^{X_{i,t}} = e^{q_i \theta} \int \eta_{(i,\theta)}(X_t, X_{i,t+\theta}) e^{X_{i,t+\theta}} dX_{i,t+\theta},$$

where q_i is the (continuously compounded) dividend yield. Hence, the underlying is itself an option structure.

The wealth of the investor evolves according to

$$\begin{aligned} w_{t+1} &= e^r w_t + \sum_i \sum_{\theta \in \Theta_{it}} \int f_{(i,\theta)}(X_t, y) (\eta_{i,\theta-1}(X_{t+1}, y) - e^r \eta_{(i,\theta)}(X_t, y)) dy \\ &\equiv e^r w_t + \mathbf{W}_{t+1}(\mathbf{f}_t) - e^r \beta_t(\mathbf{f}_t), \end{aligned} \quad (2)$$

where

$$\beta_t(\mathbf{f}_t) \equiv \sum_{(i,\theta)} \int f_{(i,\theta)}(X_t, y) \eta_{(i,\theta)}(X_t, y) dy$$

is the price of the option portfolio at time t , and π_{it} is the number of shares of asset i that investor holds at time t .

Note that the projected state prices $\eta_{(i,\theta)}$ cannot be in general chosen independently of each other. In particular, they should all satisfy

$$e^{-r\theta} = \int \eta_{(i,\theta)}(X_t, X_{i,t+\theta}) dX_{i,t+\theta}.$$

More generally, to assure absence of arbitrage, I will always assume that there exists a universal (unobservable) one period state price density that I will denote by $\Xi(X_t, X_{t+1})$. Since markets are incomplete, such a “global” state price density is not unique.

¹³One typically uses either $e^{\bar{X}} = e^{X_{i,t}}$, or the futures value $e^{\bar{X}} = e^r / q e^{X_{i,t}}$ where q is the dividend yield.

Assumption 3.1 (No arbitrage) *There exists a positive, square integrable function $\Xi(X_t, X_{t+1})$ such that*

$$\eta_{(i,\theta)}(X_t, y) = \int \Xi(X_t, X_{t+1}) \eta_{(i,\theta-1)}(X_{t+1}, y) dX_{t+1} \quad (3)$$

for all $\theta \in \Theta_{it}$, $i = 1, \dots, N$.

In order to state the results, I will need several definitions. Let $\mathbf{k} = (k_1, \dots, k_N) \in \mathbb{Z}_+^N$ be a *multi-index*. I will use the standard notation $|\mathbf{k}| = \sum_i |k_i|$, $\mathbf{k}! = k_1! \dots k_N!$ and $\binom{\mathbf{k}+1}{\mathbf{k}} = \frac{(\mathbf{k}+1)!}{\mathbf{k}!}$ and

$$\frac{\partial^{\mathbf{k}}}{\partial X^{\mathbf{k}}} = \frac{\partial^{|\mathbf{k}|}}{\partial X_1^{k_1} \dots \partial X_N^{k_N}}, \quad X^{\mathbf{k}} = X_1^{k_1} \dots X_N^{k_N}.$$

These are higher order mixed partial derivatives with respect to various risk factors. In this notation, we can write down the Taylor expansion

$$\eta_{(i,\theta-1)}(X_{t+1}, y) \approx \sum_{|\mathbf{k}|=0}^K \eta_{(i,\theta-1)}^{(\mathbf{k})}(\bar{X}_t, y) (X_{t+1} - \bar{X}_t)^{\mathbf{k}}$$

where \bar{X}_t is any reference level of X_{t+1} at time t ¹⁴ and

$$\eta_{(i,\theta-1)}^{(\mathbf{k})}(X_t, y) = \frac{1}{\mathbf{k}!} \frac{\partial^{\mathbf{k}}}{\partial X_{t+1}^{\mathbf{k}}} \eta_{(i,\theta-1)}(X_{t+1}, y) \Big|_{X_{t+1}=\bar{X}_t}. \quad (4)$$

In order to justify this approximation, I will need the following assumption.

Assumption 3.2 *For any (i, θ) , the function $\eta_{(i,\theta-1)}(X_{t+1}, y)$ can be written as a convergent power series*

$$\eta_{(i,\theta-1)}(X_{t+1}, y) = \sum_{|\mathbf{k}|=0}^{\infty} \eta_{(i,\theta-1)}^{(\mathbf{k})}(\bar{X}_t, y) (X_{t+1} - \bar{X}_t)^{\mathbf{k}}. \quad (5)$$

The quantities $\eta_{(i,\theta-1)}^{(\mathbf{k})}(X_t, y)$ have a very clear economic meaning: by the Breeden-Litzenberger formula, we can express them as

$$\eta_{(i,\theta-1)}^{(\mathbf{k})}(\bar{X}_t, y) = \frac{1}{\mathbf{k}!} \frac{\partial^2}{\partial y^2} \left(\frac{\partial^{\mathbf{k}}}{\partial X_{t+1}^{\mathbf{k}}} C_{(i,\theta-1)}(X_{t+1}, y) \right). \quad (6)$$

Thus, $\eta_{(i,\theta-1)}^{(\mathbf{k})}(X_t, y)$ can be expressed in terms of the sensitivities (Greeks) of the observable Call option prices. Furthermore, given an option structure $f_{(i,\theta)}(y)$, the next-period

¹⁴E.g., one could pick $\bar{X}_t = X_t$ or $\bar{X}_t = E_t[X_{t+1}]$.

price of this structure is given by $P(X_{t+1}; f_{(i,\theta)}) = \int \eta_{(i,\theta-1)}(X_{t+1}, y) f_{(i,\theta)}(y) dy$ and therefore the Greeks of this option structure are given by

$$\frac{1}{\mathbf{k}!} \frac{\partial^{\mathbf{k}}}{\partial X_{t+1}^{\mathbf{k}}} P(X_{t+1}; f_{(i,\theta)}) = \int \eta_{(i,\theta-1)}^{(\mathbf{k})}(X_{t+1}, y) f_{(i,\theta)}(y) dy.$$

Substituting expansion (5) into formula (3) in Assumption 3.1, I arrive at the following interesting result.

Lemma 3.3 *There exist functions $\mathcal{Z}_{(\mathbf{k})}(X_t)$ independent of (i, θ) such that*

$$\eta_{(i,\theta)}(X_t, y) = \sum_{|\mathbf{k}|=0}^{\infty} \mathcal{Z}_{(\mathbf{k})}(X_t) \eta_{(i,\theta-1)}^{(\mathbf{k})}(X_t, y) \quad (7)$$

In fact, $\mathcal{Z}_{(\mathbf{k})}(X_t, X_{t+1})$ is related to state prices Ξ via¹⁵

$$\mathcal{Z}_{(\mathbf{k})}(X_t) = \int \Xi(X_t, X_{t+1}) (X_{t+1} - \bar{X}_t)^{\mathbf{k}} dX_{t+1}. \quad (8)$$

Lemma 3.3 is very useful because it allows us to approximate horizon θ state price by the Greeks of horizon $(\theta - 1)$ state price. In reality, given the $\eta_{(i,\theta-1)}^{(\mathbf{k})}(X_t, y)$, coefficients $\mathcal{Z}_{(\mathbf{k})}(X_t)$ can be estimated directly by finding the best approximation to $\eta_{(i,\theta)}(X_t, y)$ by linear combinations of $\eta_{(i,\theta-1)}^{(\mathbf{k})}(X_t, y)$.

Let us now consider the same one period mean-variance optimization problem as above, but with multiple option maturities. Note that we can always normalize $E_t[\mathbf{W}_{t+1}(\mathbf{f}_t)]$ by adding a constant to any of $f_{(i,\theta)}$. The following is true.

Lemma 3.4 *Suppose that there exists an optimal option structure. Under the normalization $E_t[\mathbf{W}_{t+1}(\mathbf{f}_t)] = 0$, it satisfies the first order conditions that have the following form:*

$$\begin{aligned} & E_t[\eta_{(j,\theta_j-1)}(X_{t+1}, y)] - e^r \eta_{(j,\theta_j)}(X_t, y) \\ & = \gamma \int p(X_t, X_{t+1}) \eta_{(j,\theta_j-1)}(X_{t+1}, y) \mathbf{W}_{t+1}(\mathbf{f}_t) dX_{t+1}, \quad \theta_j \in \Theta_{j,t}, \quad j = 1, \dots, N. \end{aligned} \quad (9)$$

As we can see from Lemma 3.4, the possibility of full state-contingent bets leads to first order conditions that form a highly non-trivial system of integral equations.¹⁶ However, using expansion (5), we can approximate the first order conditions to an arbitrary degree of precision by a simple finite-dimensional system of linear equations. Universality result

¹⁵Note that $\mathcal{Z}_{(0)}(X_t, X_{t+1}) = e^{-r}$ always.

¹⁶The only simple case in which it can be explicitly solved is when the stochastic discount factor $\Xi(X_t, X_{t+1})/p$ is replicable, in which case $\mathbf{W}_{t+1}(\mathbf{f}_t) = \gamma^{-1}(1 - e^r \Xi(X_t, X_{t+1})/p)$.

(Lemma 3.3) implies that, in fact, the whole system (9) can be significantly simplified. Namely, the problem reduces to finding the Greeks of the optimal portfolio. Define

$$\Gamma_{(\mathbf{k})}(X_t, \mathbf{f}_t) \equiv \frac{\partial^{(\mathbf{k})}}{\partial X_{t+1}^{\mathbf{k}}} \mathbf{W}_{t+1}(\mathbf{f}_t)|_{X_{t+1}=\bar{X}_t} = \sum_{(i,\theta)} \int \eta_{(i,\theta-1)}^{(\mathbf{k})}(\bar{X}_t, y) f_{(i,\theta)}(X_t, y) dy. \quad (10)$$

The optimal choice of $\Gamma_{(\mathbf{k})}$ is precisely what characterizes a Greek-efficient portfolio.

I will now need to make some additional assumptions about the nature of market incompleteness. Namely, some of the higher order risks (measured by higher order Greeks) may simply be impossible to hedge due to market incompleteness. This will happen if, for some \mathbf{k} , I have $\eta_{(i,\theta)}^{(\mathbf{k})} = 0$ for all (i, θ) . For example, this will happen if the state prices for stock $i \neq j$ do not depend on the realization of the price of stock j , in which case the mixed derivatives with respect to log stock prices will all be identically zero. Furthermore, even if an exposure to a particular Greek is spanned, there will generally be different equivalent ways of generating exposure to this Greek. For example, suppose that the state prices for a stock i depend on its own price and the (observable) stock volatility so that $\eta_{(i,\theta)}(X_t, y) = \frac{1}{\sigma_t} \eta_{(i,\theta)}((y - X_{i,t})/v_{i,t})$. Let k_{X_i} and k_{v_i} denote the components of \mathbf{k} corresponding to the derivatives with respect to X_i and v_i . Then, only exposure a direct calculation implies that $\eta_{(i,\theta)}(X_t, y)^{(\mathbf{k})} = 0$ if \mathbf{k} involves derivatives with respect to other variables than X_t , σ_t , whereas, for any \mathbf{k} with only k_{X_i} and k_{v_i} non-zero, I have

$$\eta_{(i,\theta)}^{(\mathbf{k})}(X_t, y) = \frac{\partial^{k_{X_i} + k_{v_i}}}{\partial X_i^{k_{X_i}} \partial v_i^{k_{v_i}}} \frac{1}{\sigma_t} \eta_{(i,\theta)}((y - X_{i,t})/v_{i,t}) = \sum_{k \leq |\mathbf{k}|} B_k(X_t) \eta_{(i,\theta)}^{(k)}((y - \bar{X}_{i,t})/v_{i,t}).$$

Consequently, the span of $\eta_{(i,\theta)}^{(k_{X_i} + k_{v_i})}(X_t, y)$, $k_{X_i} + k_{v_i} \leq k$ has dimension k . If claims on volatility $v_{i,t}$ are tradable¹⁷ and the corresponding conditional state prices are given

$$\eta_{(v_i,\theta)}(X_t, y) = \eta_{(v_i,\theta)}(X_{i,t}, v_{i,t}, y),$$

the span of the corresponding derivatives will be larger, and also needs to be taken into account. To cover all degeneracies of this sort, I will make the following assumption.

Assumption 3.5 *For any (i, θ) and any l , there exist a set $\mathbf{k}_{l,1}, \dots, \mathbf{k}_{l,m(i,\theta)}$ with $|\mathbf{k}_{l,j}| = l$ such that the Greeks*

$$\eta_{(i,\theta)}^{\mathbf{b},K}(\bar{X}_t, y) \equiv \left(\eta_{(i,\theta)}^{(\mathbf{k}_{l,j})}(\bar{X}_t, y) \right)_{1 \leq l \leq K, 1 \leq j \leq m(i,\theta)}$$

are linearly independent and form a basis for the span of $\eta_{(i,\theta)}^{(\mathbf{k})}(\bar{X}_t, y)$, $|\mathbf{k}| \leq K$.

¹⁷This is indeed true for most large stock indices now (e.g., for S& P500, EURO STOXX 50, FTSE 100, DAX, SMI and the Nikkei).

By linear independence, we can always find a portfolio of option structures, $\mathbf{f}_t = (f_{(i,\theta)}(X_t, y))_{(i,\theta)}$ that attains any Greek vector

$$\mathbf{\Gamma}^{\mathbf{b}} \equiv (\mathbf{\Gamma}_{(1,\mathbf{k}_l,j)})_{(i,\theta), 1 \leq l \leq K, 1 \leq j \leq m_{(i,\theta)}}.$$

All other Greeks of order L and lower are then expressed as linear combinations of these “basis Greeks”. Let $\mathbf{A}_{t,K}$ be the linear matrix that links all Greeks with the basis Greeks. That is,

$$(\mathbf{\Gamma}_{(1,\mathbf{k})})_{|\mathbf{k}| \leq K} = \mathbf{A}_{t,K} \mathbf{\Gamma}^{\mathbf{b}}.$$

Effectively, the hedgeable Greeks, $\mathbf{\Gamma}^{\mathbf{b}}$, of the optimal option portfolio, will play the role of portfolio weights in the classical Markowitz problem. The role of the variance-covariance matrix will be played by the matrix of conditional higher order moments. Namely, I define

$$\mathbf{M}(X_t, \mathbf{k}) \equiv E_t[(X_{t+1} - \bar{X}_t)^{\mathbf{k}}] \quad (11)$$

and then define the matrix $\mathbf{\Sigma}_t = (\mathbf{M}(X_t, \mathbf{k} + \mathbf{l}))_{|\mathbf{k}|, |\mathbf{l}| \geq 0}$. For simplicity, I will always assume that this matrix is non-degenerate (and hence strictly positive definite).¹⁸

Lemma 3.6 *The matrix $\mathbf{\Sigma}_t$ is nonnegative definite. It is degenerate if and only if there exists a polynomial relationship between the prices of the assets.*

Let

$$\mu_{\mathbf{k}}(X_t) \equiv \mathbf{M}(X_t, \mathbf{k}) - e^r \mathcal{Z}_{(\mathbf{k})}(X_t)$$

be the risk premium of the moment $(X_{t+1} - \bar{X}_t)^{\mathbf{k}}$ (i.e., the difference between its physical and risk neutral expectation¹⁹ and $\mu = (\mu_{\mathbf{k}})_{\mathbf{k} \in H}$ be the vector of risk premia with H being the set of multi-indices for which Greeks are non-zero.²⁰) Combining Assumption 3.2 with Lemma 3.3, I arrive at the following expansion for the left-hand side of (9):

$$E_t[\eta_{(j,\theta_j-1)}(X_{t+1}, y)] - e^r \eta_{(j,\theta_j)}(X_t, y) \approx \sum_{\mathbf{k}} \mu_{\mathbf{k}}(X_t) \eta_{(i,\theta-1)}^{(\mathbf{k})}(X_t, y). \quad (12)$$

Substituting (12) into (9) and using Assumption 3.2 and formula (10), I arrive at the following result.²¹

¹⁸If one of the traded assets is a basket of some other traded assets, the matrix is obviously degenerate. My results can be directly extended to cover such situations.

¹⁹Lemma 3.3 implies that $e^r \mathcal{Z}_{(\mathbf{k})}(X_t)$ can be interpreted as a market price of the claim $(X_{t+1} - \bar{X}_t)^{\mathbf{k}}$. However, since markets are incomplete, such a security will in many cases be unspanned, and hence its price is not uniquely defined.

²⁰That is, only those \mathbf{k} for which $\eta_{(i,\theta)}^{(\mathbf{k})}$ is non-zero, at least for some (i, θ) .

²¹Theorem 3.7 is an approximate characterization and depends on the chosen order of Greek efficiency. However, as in Garlappi and Skoulakis (2010), only a few terms should be sufficient to obtain a good degree of precision.

Theorem 3.7 (Efficient Greeks) *The Greeks $\Gamma = (\Gamma_{(\mathbf{k})}(X_t, \mathbf{f}_t))$ of the optimal portfolio satisfy*

$$\Gamma^{\mathbf{b}} = \gamma^{-1} (\mathbf{A}_{t,K}^T \Sigma_t \mathbf{A}_{t,K})^{-1} \mathbf{A}_{t,K}^T \mu. \quad (13)$$

Formula (13) is a direct analog of the Markowitz formula. The only difference is that the set of available securities involves claims on higher moments $(X_{t+1} - \bar{X}_t)^{\mathbf{k}}$, and not just the plain returns on the available underlyings. To implement the Greek-efficient optimal portfolio empirically, I need find an option portfolio that does have the desired Greeks. To this end, new can proceed as follows: First, pick a collection of liquid, plain vanilla European Call options with available maturities and strikes K_i^1, \dots, K_i^m . For each of these options, compute the Greeks using the underlying model, and then find a linear combination of them that has the desired Greeks.

4 Beyond Mean-Variance

In this section, I assume that the agent is maximizing a CRRA utility of next period's wealth. Instead of formulating the problem in terms of options, I directly work with the tradable claims and assume that the agent is choosing exposures $\{\Gamma_{(\mathbf{k})}\}_{\mathbf{k}}$ to a particular (possibly mixed) power $(X_{t+1} - X_t)^{\mathbf{k}}$. I will assume that the claims on moments are directly replicable with existing options, and hence their prices $S(X_t, \mathbf{k})$ are directly observable. Furthermore, I will directly work with the "demeaned" claims $(X_{t+1} - X_t)^{\mathbf{k}} - e^r M(X_t, \mathbf{k})$. Then, the problem can be reformulated as

$$\max_{\{\Gamma_{(\mathbf{k})}\}} \frac{1}{1-\gamma} E_t \left[\left(\Gamma_0 e^r + \sum_{|\mathbf{k}|>0} \Gamma_{(\mathbf{k})} ((X_{t+1} - X_t)^{\mathbf{k}} - M(X_t, \mathbf{k})) \right)^{1-\gamma} \right] \quad (14)$$

under the budget constraint

$$\Gamma_0 + \sum_{|\mathbf{k}|>0} \Gamma_{(\mathbf{k})} (S(X_t, \mathbf{k}) - e^{-r} M(X_t, \mathbf{k})) = 1$$

where I have normalized initial wealth to one. The first order conditions take the form

$$\begin{aligned} E_t \left[\left(\Gamma_0 e^r + \sum_{|\mathbf{k}|>0} \Gamma_{(\mathbf{k})} ((X_{t+1} - X_t)^{\mathbf{k}} - e^r M(X_t, \mathbf{k})) \right)^{-\gamma} ((X_{t+1} - X_t)^{\mathbf{l}} - M(X_t, \mathbf{l})) \right] \\ = \lambda (e^r S(X_t, \mathbf{k}) - M(X_t, \mathbf{k})), \end{aligned} \quad (15)$$

where λ is the Lagrange multiplier for the budget constraint. Assuming that the time period is sufficiently small and denoting $A_0 = \Gamma_0 e^r$, we can use the Taylor expansion

$$\begin{aligned}
& \left(1 + A_0^{-1} \sum_{|\mathbf{k}|>0} \Gamma_{(\mathbf{k})} ((X_{t+1} - X_t)^{\mathbf{k}} - e^r M(X_t, \mathbf{k})) \right)^{-\gamma} \\
& \approx 1 - \gamma A_0^{-1} \sum_{|\mathbf{k}|>0} \Gamma_{(\mathbf{k})} ((X_{t+1} - X_t)^{\mathbf{k}} - M(X_t, \mathbf{k})) \\
& + \frac{\gamma(\gamma+1)}{2} A_0^{-2} \left(\sum_{|\mathbf{k}|>0} \Gamma_{(\mathbf{k})} ((X_{t+1} - X_t)^{\mathbf{k}} - M(X_t, \mathbf{k})) \right)^2 \\
& - \frac{\gamma(\gamma+1)(\gamma+2)}{6} A_0^{-3} \left(\sum_{|\mathbf{k}|>0} \Gamma_{(\mathbf{k})} ((X_{t+1} - X_t)^{\mathbf{k}} - M(X_t, \mathbf{k})) \right)^3.
\end{aligned} \tag{16}$$

The rows in this approximation have a clear interpretation. The first row is the standard risk-return tradeoff, the second row reflects preferences for skewness, and the third row preferences for kurtosis. If I only keep the first row, the first order condition takes the form

$$\gamma A_0^{-\gamma-1} \text{Cov} \left(\sum_{|\mathbf{k}|>0} \Gamma_{(\mathbf{k})} (X_{t+1} - X_t)^{\mathbf{k}}, (X_{t+1} - X_t)^{\mathbf{1}} \right) = \lambda (M(X_t, \mathbf{1}) - e^r S(X_t, \mathbf{1})), \quad |\mathbf{1}| > 0,$$

and hence the efficient Greeks coincide with those computed above for the case of mean-variance optimization, but with an endogenous risk aversion $\gamma A_0^{-\gamma-1} \lambda^{-1}$. By contrast, introducing preferences for skewness and kurtosis leads to corrections in the optimal portfolio. Namely, we can rewrite the first order condition as

$$\begin{aligned}
& \gamma A_0^{-1-\gamma} \sum_{|\mathbf{k}|>0} \Gamma_{(\mathbf{k})} M^c(X_t, \mathbf{k}, \mathbf{1}) \\
& - \frac{\gamma(\gamma+1)}{2} A_0^{-2-\gamma} \sum_{|\mathbf{k}_1|, |\mathbf{k}_2|>0} \Gamma_{(\mathbf{k}_1)} \Gamma_{(\mathbf{k}_2)} M^c(X_t, \mathbf{k}_1, \mathbf{k}_2, \mathbf{1}) \\
& + \frac{\gamma(\gamma+1)(\gamma+2)}{6} A_0^{-3-\gamma} \sum_{|\mathbf{k}_1|, |\mathbf{k}_2|, |\mathbf{k}_3|>0} \Gamma_{(\mathbf{k}_1)} \Gamma_{(\mathbf{k}_2)} \Gamma_{(\mathbf{k}_3)} M^c(X_t, \mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{1}) \\
& = \lambda (M(X_t, \mathbf{1}) - e^r S(X_t, \mathbf{1})), \quad |\mathbf{1}| > 0,
\end{aligned} \tag{17}$$

where I have defined

$$M^c(X_t, \mathbf{k}_1, \dots, \mathbf{k}_m) \equiv E_t [((X_{t+1} - X_t)^{\mathbf{k}_1} - M(X_t, \mathbf{k}_1)) \cdots ((X_{t+1} - X_t)^{\mathbf{k}_m} - M(X_t, \mathbf{k}_m))]$$

to be the central moment of the joint distribution of asset returns. Note that the approximate first order condition (17) is non-parametric in the sense that it does not depend

on the particular parametric form of the joint distribution of returns, which in reality might be quite non-trivial to estimate. Rather, it only depends on higher-order moments that can be estimated directly from the data, assuming the order of the moments is not too high.

Though system (17) looks complicated, it can actually be solved “almost explicitly”. Let us first consider the simpler case when there is only one tradable claim involved. Then, (17) reduces to a single cubic equation for $\Gamma_{(\mathbf{k})}$,

$$\begin{aligned} & \gamma A_0^{-1-\gamma} \Gamma_{(\mathbf{k})} E_t[Y_{t+1,(\mathbf{k})}^2] - \frac{\gamma(\gamma+1)}{2} A_0^{-2-\gamma} \Gamma_{(\mathbf{k})}^2 E_t[Y_{t+1,(\mathbf{k})}^3] \\ & + \frac{\gamma(\gamma+1)(\gamma+2)}{6} A_0^{-3-\gamma} \Gamma_{(\mathbf{k})}^3 E_t[Y_{t+1,(\mathbf{k})}^4] = \lambda(M(X_t, \mathbf{k}) - e^r S(X_t, \mathbf{k})) \end{aligned} \quad (18)$$

with

$$Y_{t+1,(\mathbf{k})} = (X_{t+1} - X_t)^{\mathbf{k}} - M(X_t, \mathbf{k}).$$

A direct calculation implies that this equation has a unique real root, given by²²

$$\Gamma_{t,(\mathbf{k})}^{(0)} = g(a_{t,(\mathbf{k})}, b_{t,(\mathbf{k})}, c_{t,(\mathbf{k})}, d_{t,(\mathbf{k})})$$

and where the coefficients are given by

$$\begin{aligned} a_{t,(\mathbf{k})} &= \frac{\gamma(\gamma+1)(\gamma+2)}{6} A_0^{-3-\gamma} E_t[Y_{t+1,(\mathbf{k})}^4], \quad b_{t,(\mathbf{k})} = -\frac{\gamma(\gamma+1)}{2} A_0^{-2-\gamma} E_t[Y_{t+1,(\mathbf{k})}^3] \\ c_{t,(\mathbf{k})} &= \gamma A_0^{-1-\gamma} E_t[Y_{t+1,(\mathbf{k})}^2], \quad d_{t,(\mathbf{k})} = -\lambda(M(X_t, \mathbf{k}) - e^r S(X_t, \mathbf{k})). \end{aligned} \quad (19)$$

Now, suppose that the covariance, co-skewness and co-kurtosis terms are non-zero, but are sufficiently small relative to the “diagonal” variance, skewness, and kurtosis terms. Fix the vector of all coordinates of Γ_t except for $\Gamma_{(1)}$ and denote this vector by Γ_{-1} . Then, we can rewrite (17) as

$$\begin{aligned} & (a_{t,(1)} + \tilde{a}_{t,(1)}(\Gamma_{-1}))(\Gamma_{(1)})^3 + (b_{t,(1)} + \tilde{b}_{t,(1)}(\Gamma_{-1}))(\Gamma_{(1)})^2 \\ & + (c_{t,(1)} + \tilde{c}_{t,(1)}(\Gamma_{-1}))(\Gamma_{(1)}) + (d_{t,(1)} + \tilde{d}_{t,(1)}(\Gamma_{-1})) = 0 \end{aligned} \quad (20)$$

for some polynomial functions $\tilde{a}_{t,(1)}, \tilde{b}_{t,(1)}, \tilde{c}_{t,(1)}, \tilde{d}_{t,(1)}$. Then, we can use the same Taylor expansion techniques as above, and arrive at the following result.

Proposition 4.1 *Let $\mathcal{D}_j \equiv \text{diag} \left(\frac{\partial}{\partial_j} g(a_{t,(\mathbf{k})}, b_{t,(\mathbf{k})}, c_{t,(\mathbf{k})}, d_{t,(\mathbf{k})}) \right)_{|\mathbf{k}|>0}$. Then, the vector of efficient Greeks is approximately given by*

$$\Gamma_t \approx (\text{Id} - \mathcal{D}_d (\Sigma_t - \text{diag}(\Sigma_t)))^{-1} \left(\Gamma_t^{(0)} + \mathcal{D}_a \tilde{a}_{t,(\mathbf{k})} + \mathcal{D}_b \tilde{b}_{t,(\mathbf{k})} + \mathcal{D}_c \tilde{c}_{t,(\mathbf{k})} + \mathcal{D}_d \tilde{d}_{t,(\mathbf{k})} \right). \quad (21)$$

²²In fact, $g(a, b, c, d) = -\frac{1}{3a} (b + C + \Delta_0/C)$ with $C = \frac{1}{2^{1/3}} (\Delta_1 + (\Delta_1^2 - 4\Delta_0^3)^{1/2})^{1/3}$ and $\Delta_0 = b^2 - 3ac$, $\Delta_1 = 2b^3 - 9abc + 27a^2d$.

The approximation used in Proposition 4.1 is based on the assumption of small off-diagonal terms in covariances, co-skewness and co-kurtosis. However, extensive numerical experiments suggest that the approximation is highly accurate and works even for correlations of the order of 70%. It is also important to note that, in general, system (17) might be extremely difficult to solve even numerically. If correlations are high (above 50%), it becomes highly unstable and most algorithms either diverge or converge to saddle point solutions. But contrast, Proposition 4.1 provides closed form and easy to implement expressions for optimal portfolios. In the next section, I test Proposition 4.1 on real data and show that it may lead to significant gains for an investor who cares about higher moments.

5 Empirical Tests

The key element that is needed for implementation of the Greek efficient portfolios are strong (i.e., those with high predictive power) views on the underlying returns and/or their higher powers. Here, the very strong predictability of higher powers (and, particularly, even-order powers) is especially useful. In this section I test the theoretical results from the previous sections on several portfolio selection problems.

5.1 Views on powers of returns

Let $K = 2$, and let us compute the second order efficient Greeks. That is, I will allow investors to optimally select exposure to asset returns and their squares. For convenience, I will work directly with the stock prices, without taking the logs.²³ Furthermore, instead of generating exposure to higher moments indirectly, using longer maturity options, I follow a simpler root. Namely, I implement the strategy at monthly frequency, using only options with one month maturity. This way, we can construct claims whose payoff is exactly equal to squared returns.²⁴ The exact details of the procedure are described in the appendix.²⁵ In this case, coefficients $\mathcal{Z}_{k_i}(S_{i,t})$, $k_i = 1, 2$ are directly observable as the prices of (replicable) claims on $(S_{i,t+1} - \bar{S}_{i,t})^{k_i}$, which can be recovered directly from

²³Since for relatively short horizons, the differences between returns and log returns are small, the implementation is not too sensitive to this.

²⁴Below I also test strategies that use longer maturity options.

²⁵In reality, this is of course not possible because the number of strikes for liquid options is finite. However, as I show, the errors are sufficiently small and do not significantly influence the profitability of the strategy.

traded option prices using the Carr-Madan formula²⁶

$$\begin{aligned}\mathcal{Z}_{(i,1)}(S_{i,t}) &= e^{-q}S_{i,t} - e^{-r}\bar{S}_{i,t} \\ \mathcal{Z}_{(i,k)}(S_{i,t}) &= k(k-1) \left(\int_0^{\bar{S}_{i,t}} K^{k-2} P_{i,t+1}(S_{i,t}, K) dK + \int_{\bar{S}_{i,t}}^{+\infty} K^{k-2} C_{i,t+1}(S_{i,t}, K) dK \right).\end{aligned}\tag{22}$$

For simplicity, I will always select $\bar{S}_{i,t} = S_{i,t}$. The full matrix \mathbf{M} involves first physical moments of the risk factors. These moments are non-trivial to estimate, but I use a rolling covariance matrix, modified by a simple shrinkage procedure.²⁷ In this subsection, I consider the case when investor only has views on the second moments, and uses a CAPM-based estimate of the stock returns returns.²⁸ Then, I compute the vector of risk premia $\mu = (\mu^1, \mu^2)^T$ (i.e., the differences between expected returns and their risk-neutral expectations); here μ^1 are the risk premia for returns, and μ^2 the vector of risk premia for squared returns. Similarly, I partition the vector of efficient Greeks $\mathbf{\Gamma} = (\mathbf{\Gamma}_1, \mathbf{\Gamma}_2)$. Denote $\mathbf{M}(X_t, k) = (\mathbf{M}(S_{it}, k))_{i=1, \dots, N}$ and let

$$\Sigma_{k_1, k_2} = ((\bar{S}_{i,t})^{k_1} (\bar{S}_{j,t})^{k_2} E_t[(S_{i,t+1}/\bar{S}_{i,t} - 1)^{k_1} (S_{j,t+1}/\bar{S}_{j,t} - 1)^{k_2}])_{i,j=1}^N, \quad k_1, k_2 = 1, 2.$$

Then, the moment matrix can be partitioned as

$$\Sigma_t = \begin{pmatrix} \Sigma_{1,1} & \Sigma_{1,2} \\ \Sigma_{2,1} & \Sigma_{2,2} \end{pmatrix}$$

and the optimal Greeks are given by $\mathbf{\Gamma} = \gamma^{-1} \Sigma_t^{-1} \mu$.

I test this methodology on the S&P 500, Nasdaq and Russell 2000 returns. I use option data on these three indices available from the OptionMetrics. I only consider options with one month maturity, with the delivery on Saturday immediately following the third Friday of the expiration month. Trading is done at monthly frequency, always on the first business day after the option delivery date.

²⁶In the case when prices of the desired moments are non-observable, one could use the expansion (7) and find the coefficients that minimize the approximation error in

$$\eta_{(i,\theta)}(S_{it}, y) \approx \eta_{(i,\theta-1)}(S_{it}, y) + \sum_{k_i=1}^2 \mathcal{Z}_{(i,k_i)}(S_{i,t}) \eta_{(i,\theta-1)}^{(k_i)}(S_{it}, y).$$

²⁷Various shrinkage procedures have been developed for estimating covariance matrices. See, e.g., Ledoit and Wolf (2004). I use a very simple shrinkage procedure, whereby I multiple all off-diagonal elements of the matrix by a fixed shrinkage factor below one. I use a shrinkage 0.9, but all values in the range of [0.6, 0.9] give comparable results. The main effect in the removal of potential degeneracies that might lead to extreme portfolio positions.

²⁸This means that I am completely ignoring predictability in stock returns. In reality, out of sample predictability is non-trivial to use. See, Goyal and Welch (2008) and Ferreira and Santa-Clara (2011). Below, I use option-based dividend strips to construct out of sample forecasts (i.e., form views) of stock returns and show that my results are robust to the inclusion of such views.

Since implementation of the strategy requires availability of a sufficiently large set of liquid options with sufficient dense strikes, I only consider the period 2004- August 2013.²⁹ Portfolio rebalancing is done at monthly frequency, and I use a 100 business day rolling moments as an estimate for Σ_t . To isolate the effects of diversification and risk premia, *I always compute all returns assuming that the market interest rate is zero.*

To form views about future squared returns, I use the standard log-claim based forward looking measure of realized volatility. Namely, I compute the price of the portfolio of OTM calls and puts with maturity of one month, weighted by strike⁻². To account for the leverage effect, I regress realized squared returns on this volatility measure as well as on realized one month returns, and then use the regression coefficients to form the forecast of future realized squared returns based on contemporaneous volatility measure and the last realized one month return. As I am always using the regression coefficients that are based only on past data, my procedure is free from look-ahead bias and is implementable in real time. Namely, let $L_{i,t}$ is the price of the one-month maturity log claim on the stock index i .³⁰ Then, for every date T in the out-of-sample period 2005-2013, I compute the (a_0, a_1, a_2, a_3) coefficients from a simple linear regression

$$(R_{i,t,t+25})^2 = a_0 + a_1 L_{i,t} + a_2 R_{t-25,t} + a_3 L_{1,t} + \varepsilon_t, \quad t < T - 25,$$

using all the data from the beginning 2004 to T .³¹ Here, $R_{i,t,t+25}$ is the 25-day return on stock index i , without including dividends.³² Then, I use these coefficients to build the forecast of future one month (25 business days) returns. For the options, I always use only OTM options, and also only those options whose moneyness (defined as strike/price) is between 0.8 and 1.2. In addition, I discard options with zero trading volume, options with zero bid prices, and options with the bid-ask spread higher than 50%.

I then compute the risk premium μ as the difference between the forecast of the squared return and the average of the bid and the ask prices of the option claim with squared return as the payoff. Then, I compute the optimal portfolio, and use the actual bid and ask prices to compute the realized returns of the portfolio.³³

I consider six different portfolio strategies:

- (1) *Equally weighted index returns.* This is a simple strategy that always invests the same fixed fraction of the wealth into each of the three indices. This simple 1/N

²⁹Before 2004, data Russell 2000 options is too noisy.

³⁰If one month maturity is not available, I use linear interpolation with the available maturities.

³¹I use the S&P 500 volatility measure given by $L_{1,t}$ as an additional forecast variable for index volatilities (except for $i = 1$). One could of course directly use VIX. However, I purposely do not use the VIX data to illustrate than the methodology is self-contained and can be applied to any instrument on which liquid options are traded, even if there are no volatility claims for that particular instrument.

³²That is, quotient of prices minus one. Including dividends has a minor impact on the results.

³³Another common approach is to construct simple mean-variance efficient portfolios, using the simple in-sample mean as the view on the future returns. I have tried this and there are no gains relative to the simpler equally weighted portfolio of all three stock indices.

strategy serves as a benchmark for my analysis. It corresponds to the mean-variance efficient market portfolio for an investor who believes in CAPM with the $1/N$ portfolio as the proxy for the market portfolio.

- (2) *Index portfolio, with views on index returns.* Due the leverage effect, views on squared returns (and hence on volatility) also have forecasting power for stock returns themselves. Assume that the benchmark (no views) expected returns are given by CAPM, whereas the (no views) expected squared returns are given by their in-sample means. Assuming that returns and their squares are jointly Gaussian, we can use the Gaussian conditioning formulae³⁴ to express views on stock returns using the views on the squared returns.³⁵ Then, the portfolio is just a simple mean-variance efficient portfolio based on the realized covariance matrix and the updated expected returns that incorporate views. Note that this portfolio strategy is also linked to options and cannot be implemented based solely on index prices, because the active views on the squared returns are based on the forward looking volatility measure.
- (3) *Greek MV-efficient portfolio.* This are the Greek-efficient portfolio based on the “Markowitz Greeks” and using all six instruments: three stocks indices and three “claims on the squared returns”. The latter are synthetic instruments, but they have well defined bid and ask prices, and hence the corresponding trading gains can be directly computed. Note that since I only use moneyness in $[0.8, 1.2]$ and the set of available strikes is limited, the “squared returns” claims do not perfectly replicate the squared return. When computing the trading gains, I always use the actual realized payoff of the corresponding option portfolio, as well as actual bid and ask prices.³⁶
- (4) *Greek MV-efficient portfolio, with views on index returns.* This is the same as in (3), but with expected stock returns given by the conditional expectation from (2), given the views on the squared returns.
- (5) *Greek CRRA-efficient portfolio.* This is the Greek-efficient portfolio for a CRRA investor, based on the formula of Proposition 4.1. I use the relative risk aversion parameter $\gamma = 5$ for all computation. The covariance, co-skewness and co-kurtosis matrices are computed as the simple realized 100-day covariance, co-skewness and co-kurtosis matrices respectively. However, in order to avoid degeneracies, I also

³⁴In this setting, this is equivalent to the original Black and Litterman (1990) approach.

³⁵Of course, returns and squared returns are not jointly Gaussian, and their actual joint distribution may be highly non-trivial. However, I still follow the simple Gaussian approach because it is straightforward to implement and captures the first order effects. Incorporating the actual distribution is an interesting and important topic for future research.

³⁶I have also tested the performance of the strategy assuming that the actual squared claim is perfectly replicable. Empirical results suggest that the discrepancies are small and barely have any influence on the strategy performance.

shrink the off-diagonal elements of the co-skewness and co-kurtosis matrix by a factor of 0.3.³⁷

- (6) *Greek CRRA-efficient portfolios, with views on index returns.* This is the Greek CRRA-efficient portfolio from (5), but with with expected stock returns given by the conditional expectation from (2), given the views on the squared returns.

Obviously, option-based returns involve leverage and are therefore defined to be more volatile. Therefore, for all plots in this paper, I normalize the (monthly) returns by their ex-post realized standard deviation of the whole period and then multiply them by $0.2/\sqrt{12}$. This way we can compare profitability of the two strategies with identical ex-post annualized volatility of 20%.

The following table summarises basic characteristics of the six strategies.³⁸

Portfolio Type	Sharpe Ratio	Skewness	Kurtosis
Equally weighted index returns	0.53	-1.2	5.8
Index portfolio, with views on index returns	0.61	-1.7	8.2
Greek MV-efficient	0.87	3.9	27.2
Greek MV-efficient, with views on index returns	0.9	4	27.2
Greek CRRA-efficient	0.92	4.3	28.3
Greek CRRA-efficient, with views on index returns	0.95	4.3	30.8

First, even the equally weighted index returns exhibit some non-normality: even though the Sharpe Ratio is reasonable over the sample period, returns are negatively skewed and have a positive excess kurtosis of 2.8. Incorporating forecasts of squared returns into the views on expected index returns increases Sharpe ratio by about 15%. However, surprisingly, at the same time it also significantly increases the kurtosis and produces more pronounced negative skewness. As one can see from Figure 2, there are periods when the views portfolio involves significant leverage and rebalancing, which explains more pronounced non-normality of the returns.

Insert Figures 1 and 2 about here.

Introducing claims on squared returns changes the performance of the strategy dramatically. As one can see from Figure 3, the Greek MV-efficient portfolio significantly outperforms the equally weighted portfolio. While it naturally leads to a significantly higher Sharpe ratio (about 64% higher), it also offers a highly attractive skewness-kurtosis profile: returns are highly positively skewed, and this compensates the investor for a significantly higher kurtosis. Figures 4 and 5 show that the Greek-efficient portfolio is highly non-trivial, involves significant rebalancing, and constantly exploits relative movements of all the three indices and their squared returns. In particular, the portfolio is neither long nor short volatility.

³⁷The results are robust to chaining this factor, as long as it stays sufficiently far away from one.

³⁸Sharpe Ratios are annualized. Kurtosis is computed without subtracting 3.

Insert Figures 3, 4 and 5 about here.

The very high kurtosis of the Greek MV-efficient portfolio naturally raises the question of whether an investor with preferences for higher moments would like to significantly modify the risk-return profile. Figure 6 shows that, in fact, a CRRA investor with a risk aversion of 5 does improve the risk-return profile slightly (Sharpe Ratio increases by about 6%). Interestingly enough, the Greek CRRA-efficient actually has a higher kurtosis, but this is the price the investor is willing to pay to achieve a 10% higher positive skewness.

Finally, Figure 7 shows the effects of incorporating views into Greek-efficient portfolios. As one can see, the Greek CRRA-efficient portfolio still slightly outperforms the Greek MV-efficient portfolio but, surprisingly, at the cost of a lightly higher kurtosis.

Insert Figures 6 and 7 about here.

6 Dynamic Trading with Transactions Costs

Trading options can be quite costly. Short maturity options that are close to being at the money are usually relatively cheap to trade, however for longer maturity and/or out-of-the-money options, transactions costs can be very large, making rebalancing extremely costly if not impossible. See, George and Longstaff (1993) and Cao and Wei (2010).

This dependence of the transactions costs on maturity and moneyness is a very important feature of option trading that distinguishes options from stocks and other simple securities. Indeed, for a sufficiently liquid stock, transactions costs barely move over time, at least over a reasonable time horizon. By contrast, a position in a longer-term option suffers from the liquidity risk: even if the option is close to being at the money today, it may easily become out-of-the-money tomorrow, making it very costly to liquidate the position.

Following Garleanu and Pedersen (2012), I will assume that transaction costs are quadratic in the size of the trade. Modeling transaction costs with options is a non-trivial problem because the same state-contingent claim can be replicated with different option portfolios.³⁹ My goal is to capture the moneyness effect implying that it is expensive to make bets on the realizations of the underlying that are far away from the expected return. To strip down this effect, I will assume that a bet on the value $f_{(i,\theta)}(y)$ costs⁴⁰

$$0.5(1 - \rho)\Lambda_{(i,\theta)}(X_{i,t}, y)f_{(i,\theta)}^2(y)dy.$$

³⁹For example, due to Put-Call parity, a put option can be also replicated by buying a call option and shorting the stock.

⁴⁰This means that transaction costs are specified directly for “Arrow securities”. A closest proxy for such a security is a butterfly spread with a small difference in moneyness. This is a relatively simple option structure, and we can think of $\Lambda(y)$ as the cost of trading such a structure.

Consequently, trading the whole option portfolio costs

$$K_t = \frac{1}{2} \sum_i \sum_\theta \int \Lambda_{(i,\theta)}(X_{i,t}, y) (f_{(i,\theta+1),t-1}(X_{t-1}, y) - f_{(i,\theta),t}(X_t, y))^2 dy \quad (23)$$

at time t . As in Garleanu and Pedersen (2013), I will assume that the investor's objective function is given by

$$V_t(\mathbf{f}_{t-1}, X_t) = \max_f E_t \left[\sum_{\tau=t}^{T-1} (1-\rho)^{\tau-t} \left((w_{\tau+1} - e^r w_\tau) - \frac{1}{2} \gamma \text{Var}_\tau[w_{\tau+1}] - K_\tau \right) \right] \quad (24)$$

where T is the planning horizon and K_t is the transactions cost at time t .⁴¹

Standard dynamic programming principle implies that the value function satisfies

$$\begin{aligned} V_t(\mathbf{f}_{t-1}, X_t) &= E_t[\mathbf{W}_{t+1}(\mathbf{f}_t)] - e^r \beta_t(\mathbf{f}_t) - \frac{1}{2} \gamma \text{Var}_t[\mathbf{W}_{t+1}(\mathbf{f}_t)] \\ &\quad - \frac{1}{2} \sum_{(i,\theta)} \int \Lambda_{(i,\theta)}(X_{i,t}, y) (f_{(i,\theta+1),t-1}(X_{t-1}, y) - f_{(i,\theta),t}(X_t, y))^2 dy \\ &\quad - \frac{1}{2} \sum_i \int \Lambda_{(i,\theta_{i,\max})}(X_{i,t}, y) (f_{(i,\theta_{i,\max}),t}(X_t, y))^2 dy + (1-\rho) V_{t+1}(X_{t+1}, \mathbf{f}_t) \end{aligned} \quad (25)$$

Here, the additional term involving θ_{\max} corresponds to newly acquired options of maximal maturity that only became available for trading at time t .

As above, I will consider the approximate problem for an investor who only cares about a finite number of Greeks. In this case, one can approximate the option returns by linear combinations of Greeks times the corresponding mixed powers of asset returns. However, there is a very important difference between the cases with and without transaction costs. Namely,

- without transaction costs, the exact form of the option structure that generates efficient Greeks is completely irrelevant. Only the Greeks themselves matter;
- by contrast, with transaction costs, the exact form of the option structure is important because it influences the cost at which the optimal Greek exposure can be acquired.

This simple but very important observation naturally leads us to expect that there are “trading-cost optimal option structures” that efficiently balance trading costs and the exposure to risk factors. These option structures should of course depend both on the functional form of $\Lambda_{(i,\theta)}(X_{i,t}, y)$ and on the state prices.

⁴¹For simplicity, I assume that there are no liquidation costs at time T . The analysis can be easily modified to include quadratic liquidation costs at the planning horizon.

To get the basic idea, let first solve the myopic problem for an investor who is endowed with an option portfolio \mathbf{f}_{t-1} at time t , and trades optimally once. The optimization problem for the investor is then

$$\max_{\mathbf{f}_t} \left\{ E_t[\mathbf{W}_{t+1}(\mathbf{f}_t)] - e^r \beta_t(\mathbf{f}_t) - \frac{1}{2} \gamma \text{Var}_t[\mathbf{W}_{t+1}(\mathbf{f}_t)] - \frac{1}{2} \sum_{(i,\theta)} \int \Lambda_{(i,\theta)}(X_{i,t}, y) (f_{(i,\theta+1),t-1}(X_{t-1}, y) - f_{(i,\theta),t}(X_t, y))^2 dy \right\}. \quad (26)$$

Approximating the returns on the option structure by the corresponding sums of (Greeks) \times (powers) of returns, we immediately arrive at the following result.

Lemma 6.1 *Let $\Sigma_t \equiv (\text{Cov}_t((X_{t+1} - X_t)^{\mathbf{k}}, (X_{t+1} - X_t)^{\mathbf{l}}))$ and $\Gamma(X_t, \mathbf{f}_t) = (\Gamma_{(\mathbf{k})})_{\mathbf{k}}$. Then,*

$$E_t[\mathbf{W}_{t+1}(\mathbf{f}_t)] - e^r \beta_t(\mathbf{f}_t) \approx \mu(X_t)^T \Gamma(X_t, \mathbf{f}_t) \quad (27)$$

and

$$\text{Var}_t[(\mathbf{W}_{t+1}(\mathbf{f}_t))] \approx \Gamma(X_t, \mathbf{f}_t)^T \Sigma_t \Gamma(X_t, \mathbf{f}_t). \quad (28)$$

Now substituting the approximate expressions of Lemma 6.1 into (26), writing down the first order condition with respect to $f_{(i,\theta),t}(X_t, y)$, and using the expansion (12), we get

$$\begin{aligned} & \Lambda_{(j,\theta_j)}(X_t, y) (f_{t,(j,\theta_j)}(X_t, y) - f_{(i,\theta+1),t-1}(X_{t-1}, y)) \\ &= \sum_{\mathbf{k}} \mu_{\mathbf{k}}(X_t) \eta_{(j,\theta_j-1)}^{(\mathbf{k})}(X_t, y) - \gamma \sum_{\mathbf{k}} \eta_{(j,\theta_j-1)}^{(\mathbf{k})}(X_t, y) (\Sigma_t \Gamma(X_t, \mathbf{f}_t))_{\mathbf{k}} \end{aligned} \quad (29)$$

It immediately follows that, for all $\theta_j > 1$, the optimal option structure is given by the “old” structure $f_{(i,\theta+1),t-1}(X_{t-1}, y)$ plus a linear combination of the “optimal claims”

$$\Phi_{(j,\theta_j)}^{(\mathbf{k})}(y, X_t) \equiv \frac{\eta_{(j,\theta_j-1)}^{(\mathbf{k})}(X_t, y)}{\Lambda_{(j,\theta_j)}(X_t, y)}, \quad \theta_j > 1.$$

In particular, if $f_{(i,\theta+1),t-1}(X_{t-1}, y)$ can itself be well approximated by a combination of the optimal claims, we get that the optimal option structure itself should be a combination of these claims. We summarize these findings in the following result.

Lemma 6.2 *Suppose that $f_{(i,\theta+1),t-1}(X_{t-1}, y)$ is a linear combination of $\Phi_{(j,\theta_j)}^{(\mathbf{k})}(y, X_t)$, $|\mathbf{k}| \geq 0$. Then, so is the optimal option structure $f_{t,(j,\theta_j)}(X_t, y)$.*

The assumption of Lemma 6.2 is not restrictive at all, as is shown by the following result.

Lemma 6.3 *Suppose that $\eta_{(j,\theta_j-1)}(X_t, t) = \eta_{(j,\theta_j-1)}(X_{-j,t}, y - X_{j,t})$ has Fourier transform that is almost everywhere non-zero.*

Suppose also that all derivatives $\eta_{(j,\theta_j-1)}^{(\mathbf{k})}(X_t, y)$, $|\mathbf{k}| \geq 0$, are square integrable. Then, the assumption of Lemma 6.2 holds.

It remains to study the first order conditions for one period options (i.e., $\theta_j = 1$). In this case, I assume that the optimal option structure is given by a linear combination of powers of returns.⁴² That is,

$$f_{t,(i,1)}(X_t, y) \approx \sum a_{t,(i,1)}^{(k)}(X_t)(y - X_{i,t})^k,$$

for some coefficients $a_{t,(i,1)}^{(k)}(X_t)$. Then, the first order condition with respect to $a_{t,(i,1)}^{(k)}(X_t)$ takes the form

$$\begin{aligned} & (\Sigma_{t,i}^\Lambda a_{t,(i,1)}(X_t))_k - \int \Lambda_{(i,1)}(X_t, y)(y - X_{i,t})^k f_{(i,2),t-1}(X_{t-1}, y) dy \\ &= \mu_k(X_t) - \gamma \sum_k a_{t,(i,1)}^{(k)}(X_t) (\Sigma_t \Gamma(X_t, \mathbf{f}_t))_k \end{aligned} \quad (30)$$

where I have defined the “transaction costs-adjusted moment matrix”

$$\Sigma_{t,i}^\Lambda \equiv \left(\int \Lambda_{(i,1)}(X_t, y)(y - X_{i,t})^{k+l} \right)_{k,l \geq 0}.$$

To homogenise the notation, we will also denote

$$\Phi_{(j,1)}^{(k)}(y, X_t) \equiv (y - X_{j,t})^k.$$

From now on, we will make the following assumption.

Assumption 6.4 *For any (j, θ_j) , there exists a matrix $\Omega^{(j,\theta_j)}(X_{t-1}, X_t) = (\Omega_{\mathbf{k},\mathbf{l}}^{(j,\theta_j)}(X_{t-1}, X_t))_{\mathbf{k},\mathbf{l}}$ such that*

$$\Phi_{(j,\theta_j+1)}^{(k)}(y, X_{t-1}) = \sum_{\mathbf{l}} \Omega_{\mathbf{l},\mathbf{k}}^{(j,\theta_j)}(X_{t-1}, X_t) \Phi_{(j,\theta_j)}^{(\mathbf{l})}(y, X_t).$$

We will call $\Omega^{(j,\theta_j)}(X_{t-1}, X_t)$ the transition matrix. It plays a very important role in the portfolio allocation problem: due to the changing maturity of the options, investor’s portfolio effectively consists of different securities every period. This is the *rotation effect*. The matrices Ω formalize this effect.

It turns out that, under the assumption made, the optimal option portfolios can be characterized in terms of the option Greeks. The main difference from the case without transaction costs is that not only total Greeks $\Gamma_{(\mathbf{k})}$ of the whole portfolio matter, but also individual Greeks for any given asset-maturity pair (j, θ_j) play an important role in the structure of the optimal portfolio.

Let $\mathbf{B}^{(j,\theta_j)}(X_t)$ be the matrix with elements

$$\mathbf{B}_{\mathbf{l},\mathbf{k}}^{(j,\theta_j)}(t, X_t) \equiv \int \frac{\eta_{(j,\theta_j-1)}^{(\mathbf{l})}(X_t, y) \eta_{(j,\theta_j-1)}^{(\mathbf{k})}(X_t, y)}{\Lambda_{(j,\theta_j)}(t, X_t, y)} dy. \quad (31)$$

⁴²This can always be achieved by taking the corresponding Taylor polynomial approximation.

These are the *transaction costs Greeks*. I.e., the Greeks of optimal option structures.

Let also $\kappa_{(j,\theta_j)}(\mathbf{f}_t) = (\kappa_{(j,\theta_j)}^{(\mathbf{k})}(\mathbf{f}_t))_{\mathbf{k}}$ be the vector of Greeks for the option structure for the pair (j, θ_j) . If

$$f_{t,(j,\theta_j)}(X_t, y) = \sum_{\mathbf{k}} a_{t,(j,\theta_j)}^{\mathbf{k}}(X_t) \Phi_{(j,\theta_j)}^{(\mathbf{k})}(y, X_t) \quad (32)$$

then the option Greeks can be characterized as

$$\kappa_{(j,\theta_j)}(\mathbf{f}_t) = \mathbf{B}^{(j,\theta_j)}(X_t) \mathbf{a}_{t,(j,\theta_j)}. \quad (33)$$

We will also need some more notation. Let $\mathbf{B}_t = \text{diag}(\mathbf{B}^{(j,\theta_j)}(t, X_t))$ and $\Psi = \text{diag}(\Psi_{(j,\theta_j)})_{(j,\theta_j)}$ be the block-diagonal matrix with

$$\Psi_{(j,\theta_j)} = \Omega^{(j,\theta_j)}(X_{t-1}, X_t) (\mathbf{B}^{(j,\theta_j)}(t, X_t))^{-1}.$$

Let also \mathcal{I} be the matrix operating in the space of Greek vectors κ and performing summation ver all pairs (j, θ_j) , for a fixed \mathbf{k} . That is, $(\mathcal{I}\kappa(\mathbf{f}))_{\mathbf{k}} = \Gamma_{\mathbf{k}}(\mathbf{f})$. Let also \mathcal{S} be the *maturity shift matrix* such that $(\mathcal{S}\kappa(\mathbf{f}))_{(j,\theta_j)}^{(\mathbf{k})} = \kappa_{(j,\theta_j-1)}^{(\mathbf{k})}(\mathbf{f}_t)$, and where I am using the convention $\kappa_{(j,0)} = 0$.

We can now state the main result of this section.

Theorem 6.5 *There exists a function $L(t, X_t)$, a function $\mathcal{A}(t, X_t)$ taking values in the set of positive definite matrices, and a vector-valued function $\mathcal{C}(X_t)$ such that the value function $V_t(X_t, \mathbf{f}_{t-1})$ is given by*

$$V_t(X_t, \mathbf{f}_{t-1}) = L(X_t) - 0.5 \kappa(\mathbf{f}_{t-1})^T \mathcal{A}(t, X_t) \kappa(\mathbf{f}_{t-1}) - \kappa(\mathbf{f}_{t-1})^T \mathcal{C}(X_t).$$

In particular, the value function only depends on the Greeks $\kappa(\mathbf{f}_{t-1})$ of the old position. The optimal Greeks of the time- t optimal portfolio, $\kappa(\mathbf{f}_t)$, are given by

$$\kappa(\mathbf{f}_t) = \left(\gamma \mathcal{I}^T \Sigma_t \mathcal{I} + \mathbf{B}_t^{-1} + (1-\rho) E_t[\mathcal{A}_{t+1}(X_{t+1})] \right)^{-1} \left(\mathcal{I}^T \mu - (1-\rho) E_t[\mathcal{C}_{t+1}(X_{t+1})] + \Psi \kappa(\mathbf{f}_{t-1}) \right)$$

and the optimal option structure is given by (32), with \mathbf{a} defined via (33). The coefficient matrices are defined recursively backwards in time via

$$\mathcal{A}_t(X_{t-1}, X_t) = \mathcal{S}^T \Psi^T \left(\mathbf{B}_t - \left(\gamma \mathcal{I}^T \Sigma_t \mathcal{I} + \mathbf{B}_t^{-1} + (1-\rho) E_t[\mathcal{A}_{t+1}(X_t, X_{t+1})] \right)^{-1} \right) \Psi \mathcal{S} \quad (34)$$

and

$$\begin{aligned} \mathcal{C}_t(X_{t-1}, X_t) &= \mathcal{S}^T \Psi^T \left(\gamma \mathcal{I}^T \Sigma_t \mathcal{I} + \mathbf{B}_t^{-1} + (1-\rho) E_t[\mathcal{A}_{t+1}(X_t, X_{t+1})] \right)^{-1} \\ &\times \left(-\mathcal{I}^T \mu + (1-\rho) E_t[\mathcal{C}_{t+1}(X_t, X_{t+1})] \right). \end{aligned} \quad (35)$$

In the empirical implementation of the above formulae, computing the conditional expectation $E_t[\mathcal{A}_{t+1}(X_t, X_{t+1})]$ presents the greatest challenge because it involves expectations of complicated algebraic expressions involving the \mathbf{B}_t and Σ_t matrices. However, to capture the main effects, we can replace $E_t[\mathcal{A}_{t+1}(X_t, X_{t+1})] \approx \mathcal{A}_{t+1}(X_t, X_t)$ and, similarly, $E_t[\mathcal{C}_{t+1}(X_t, X_{t+1})] \approx \mathcal{C}_{t+1}(X_t, X_t)$. Then, there are no more expectation involved, and all quantities can be easily computed. Assuming the planning horizon is infinite, the matrix $\mathcal{A}(X_t, X_{t+1})$ becomes stationary and can be computed as follows: first, solve the fixed point

$$\mathcal{A}_t(X_t, X_t) = \mathcal{S}^T \Psi^T(X_t, X_t) \left(\mathbf{B}_t - \left(\gamma \mathcal{I}^T \Sigma_t \mathcal{I} + \mathbf{B}_t^{-1} + (1 - \rho) \mathcal{A}(X_t, X_t) \right)^{-1} \right) \Psi(X_t, X_t) \mathcal{S}$$

for $\mathcal{A}_t(X_t, X_t)$ and then define

$$\mathcal{A}_t(X_{t-1}, X_t) = \mathcal{S}^T \Psi^T(X_{t-1}, X_t) \left(\mathbf{B}_t - \left(\gamma \mathcal{I}^T \Sigma_t \mathcal{I} + \mathbf{B}_t^{-1} + (1 - \rho) \mathcal{A}_t(X_t, X_t) \right)^{-1} \right) \Psi(X_{t-1}, X_t) \mathcal{S}.$$

Then, exactly the same procedure can be used to determine $\mathcal{C}_t(X_{t-1}, X_t)$. We follow this approach in the empirical implementation.

7 Managing Transaction Costs: Beyond Mean-Variance

The quadratic Ansatz used to compute the value function in the previous section can only be used if the investor's objective function is of the mean-variance form. However, the techniques developed in the previous section can be used in many other settings. The idea is as follows. Suppose that I have a frictionless (with no transaction costs) model in which I know the functional form of the optimal option structure. Denote this target structure by $\mathbf{f}^{\text{target}}(X_t, y)$. This optimal structure may arise either from some non-quadratic portfolio optimization problem, or from some optimal hedging problem. The exact origin is not important. Obviously, after the introduction of transaction costs, the target portfolio is not anymore optimal. However, an investor should still be willing to match the desired exposure as closely as possible. Therefore, we assume that the investor's objective is to minimize

$$\begin{aligned} & \sum_{t=0}^T (1 - \rho)^t \sum_{(i, \theta)} \left(E \left[\alpha \int \Lambda_{(i, \theta)}(X_t, y) (f_{(i, \theta)}^{\text{target}}(X_t, y) - f_{(i, \theta)}(X_t, y))^2 dy \right. \right. \\ & \left. \left. + (1 - \alpha) \int \Lambda_{(i, \theta)}(X_t, y) (f_{(i, \theta+1)}(X_{t-1}, y) - f_{(i, \theta)}(X_t, y))^2 dy \right] \right) \end{aligned} \quad (36)$$

over all possible option structures ($f_{(i, \theta)}$). That is, the investor is trying to minimize the distance between the target portfolio and the selected portfolio (adjusted for transaction costs) and is at the same time trying to minimize the incurred transaction costs.⁴³ The

⁴³The inclusion of transaction costs into the first term is done for two reasons: first, it slightly simplifies the expressions; second, it punishes deviations for OTM options, which are more important. All the results directly extend to the case without Λ in the first term.

relative importance of the two objectives is controlled by the weight α . Standard dynamic programming arguments imply the following result.

Proposition 7.1 *There exist functions $A_{(i,\theta)}(t, X_t, y)$, $B_{(i,\theta)}(t, X_t, y)$ such that the value function of this optimization problem is given by*

$$V(t, X_t, \mathbf{f}_{t-1}) = L(t, X_t) + \sum_{(i,\theta), \theta > 1} \int (A_{(i,\theta)}(t, X_t, y)(f_{(i,\theta)}(X_{t-1}, y))^2 + 2B_{(i,\theta)}(t, X_t, y)f_{(i,\theta)}(X_{t-1}, y))dy. \quad (37)$$

The optimal option structure is then given by

$$f_{(i,1)}(X_t, y) = (\alpha f_{(i,1)}^{\text{target}}(X_t, y) + (1 - \alpha)f_{(i,2)}(X_{t-1}, y))$$

whereas for $\theta > 1$ we have

$$f_{(i,\theta)}(X_t, y) = (\alpha f_{(i,\theta)}^{\text{target}}(X_t, y) + (1 - \alpha)f_{(i,\theta+1)}(X_{t-1}, y))\delta_{(i,\theta)}(t, X_t, y) - \xi_{(i,\theta)}(t, X_t, y) \quad (38)$$

The functions $A_{(i,\theta)}(t, X_t, y)$, $B_{(i,\theta)}(t, X_t, y)$ are defined recursively via

$$A_{(i,\theta+1)}(t, X_t, y) = \Lambda_{(i,\theta)}(X_t, y) (\alpha((1 - \alpha)\delta_{(i,\theta)}(t, X_t, y))^2 + (1 - \alpha)((1 - \alpha)\delta_{(i,\theta)}(t, X_t, y) - 1)^2) + (1 - \rho)E_t[A_{(i,\theta)}(t + 1, X_{t+1}, y)]((1 - \alpha)\delta_{(i,\theta)}(t, X_t, y))^2 \quad (39)$$

and

$$\begin{aligned} B_{(i,\theta+1)}(t, X_t, y) &= \alpha\Lambda_{(i,\theta)}(X_t, y)((\alpha\delta_{(i,\theta)}(t, X_t, y) - 1)f_{(i,\theta)}^{\text{target}}(X_t, y) - \xi_{(i,\theta)}(t, X_t, y))(1 - \alpha)\delta_{(i,\theta)}(t, X_t, y) \\ &+ (1 - \alpha)\Lambda_{(i,\theta)}(X_t, y)((\alpha\delta_{(i,\theta)}(t, X_t, y))f_{(i,\theta)}^{\text{target}}(X_t, y) - \xi_{(i,\theta)}(t, X_t, y))((1 - \alpha)\delta_{(i,\theta)}(t, X_t, y) - 1) \\ &+ (1 - \rho)E_t[A_{(i,\theta)}(t + 1, X_{t+1}, y)]((\alpha\delta_{(i,\theta)}(t, X_t, y))f_{(i,\theta)}^{\text{target}}(X_t, y) - \xi_{(i,\theta)}(t, X_t, y))(1 - \alpha)\delta_{(i,\theta)}(t, X_t, y) \\ &+ (1 - \rho)B_{(i,\theta)}(t + 1, X_{t+1}, y)(1 - \alpha)\delta_{(i,\theta)}(t, X_t, y), \end{aligned} \quad (40)$$

with the boundary conditions $A_{(i,\theta)}(T, X_T, y) = B_{(i,\theta)}(T, X_T, y)$ and where we have defined

$$\begin{aligned} \delta_{(i,\theta)}(t, X_t, y) &\equiv \frac{\Lambda_{(i,\theta)}(X_t, y)}{\Lambda_{(i,\theta)}(X_t, y) + (1 - \rho)E_t[A_{(i,\theta)}(t + 1, X_{t+1}, y)]} \\ \xi_{(i,\theta)}(t, X_t, y) &\equiv \frac{(1 - \rho)E_t[B_{(i,\theta)}(t + 1, X_{t+1}, y)]}{\Lambda_{(i,\theta)}(X_t, y) + (1 - \rho)E_t[A_{(i,\theta)}(t + 1, X_{t+1}, y)]}. \end{aligned} \quad (41)$$

Quite remarkably, the rotation effect makes this system explicitly solvable. Indeed, as the maturity on the right-hand side of (42) is always smaller than that on the left-hand side, we can continue the process until we arrive to maturity 1 for which $A_{(i,1)}(t, X_t, y) = B_{(i,1)}(t, X_t, y) = 0$. In the next section, we test Proposition 7.1 empirically, using a setting in which θ only takes two values, one and two months. In this case, Proposition 7.1 is particularly simple as we only need to deal with the options of maturity two, and we get

$$A_{(i,2)}(t, X_t, y) = \alpha(1 - \alpha)\Lambda_{(i,1)}(X_t, y) \quad (42)$$

and

$$B_{(i,2)}(t, X_t, y) = -\alpha(1 - \alpha)\Lambda_{(i,1)}(X_t, y)f_{(i,1)}^{\text{target}}(X_t, y) \quad (43)$$

and hence

$$\begin{aligned} f_{(i,2)}(X_t, y) &= \alpha f_{(i,2)}^{\text{target}}(X_t, y) \frac{\Lambda_{(i,2)}(X_t, y)}{\Lambda_{(i,2)}(X_t, y) + (1 - \rho)E_t[\alpha(1 - \alpha)\Lambda_{(i,1)}(X_{t+1}, y)]} \\ &+ \frac{(1 - \rho)\alpha(1 - \alpha)E_t[\Lambda_{(i,1)}(X_{t+1}, y)f_{(i,1)}^{\text{target}}(X_{t+1}, y)]}{\Lambda_{(i,\theta)}(X_t, y) + (1 - \rho)E_t[\alpha(1 - \alpha)\Lambda_{(i,1)}(X_{t+1}, y)]} \end{aligned} \quad (44)$$

In the empirical implementation of this formula, when computing the expectation

$$E_t[\Lambda_{(i,1)}(X_{t+1}, y)f_{(i,1)}^{\text{target}}(X_{t+1}, y)],$$

one can use the constructed forecast of powers of returns up to power three, then the squared forecast of the squared return as a forecast for the fourth power, and ignore the fifth power of returns.⁴⁴

In the empirical implementation, I first use Proposition 4.1 to compute the optimal Greeks for a CRRA investor, and then use the optimal option structures $\Phi_{(i,\theta)}^{(\mathbf{k})}$ from the previous section to implement these Greeks. This defines the target portfolio. Finally, I use the value $\alpha = 0.9$ in the empirical tests.

8 Empirical Tests

Two crucial ingredients are necessary for a proper implementation of the theoretical strategy developed above:

- (1) We need to construct the option liquidity $\Lambda_{(i,\theta)}(X_t, y)$ and then construct the corresponding optimal claims $\frac{1}{\Lambda_{(i,\theta)}(X_t, y)}\eta_{(i,\theta)}^{(k)}(X_t, y)$ and relate them to the claims on powers of returns;
- (2) We need a good model for the option Greeks for optimal option claims. That is, we need a simple and robust way of approximating option returns via the sum of products of Greeks times the corresponding powers of the underlying;
- (3) We need strong and robust views on all of the powers to which the investor wishes to get exposure. While these views are not so important for the basic assets themselves (e.g., just assuming CAPM, as we did in the previous sections, may already do the job), strong views on higher powers are extremely important. The reason is that the higher the power, the higher is the embedded leverage that the corresponding option structure provides. Thus, the risk of enormous losses is also very high.

⁴⁴In fact, the effect of powers higher than three are marginal and can be neglected.

I will now to implement (1)- (3), and then use it to test the strategy. Namely, I consider a periodic structure as in Proposition ???: as above, trading happens with monthly frequency, but, at every date, I assume that agents invest in options with one- and two-months maturity. This naturally creates a link between option positions through the *rotation effect*: two-month options become one-month next period, and, because existing this position is costly, it influences the choice of the new two-month maturity options, which will in turn influence future portfolio choice, etc.

I start with item (1). To this end, on every date, on which there are options with one- and two-months maturity, I compute consider bid and ask prices of the OTM puts and calls, and compute the bid-ask spread $L(K)$ in percentage terms (i.e., $\text{ask}/\text{bid} - 1$) as a function of strike K . For each of the two maturities, this gives a curves as a function of strike. To capture the shape of this curve in a parsimonious fashion, I regress $L(K)$ on K/price and $(K/\text{price})^2$. This gives an approximation of option liquidity of the form $a_0^i(\theta) + a_1^i(\theta)(K/\text{price} - 1) + a_2^i(\theta)(K/\text{price} - 1)^2$ and the coefficients a_0^i, a_1^i, a_2^i respectively capture the level, asymmetry and concavity of the option liquidity for stock index i . Here, $\theta = 1, 2$ is the option maturity. The coefficients $a_k(\theta)$, $k = 0, 1, 2$, exhibit significant variation over time and across the three stock indices. Investigating their time series behaviour is an important topic that goes beyond the scope of this paper.⁴⁵ However, in this paper I take a simpler approach. I average the values of $a_k(m)$ over time and use the corresponding quadratic function $\Lambda_{(i,\theta)}(X_t, y) \approx \bar{a}_0^i(\theta) + \bar{a}_1^i(\theta)(y/X_t - 1) + \bar{a}_2^i(\theta)(y/X_t - 1)^2$ as the input in the portfolio optimization problem.

Figures 8 and 9 show the true option liquidity function and the corresponding quadratic approximation. Two important effects are worth noting. First, ask/bid is increasing as we go deeper out-of-the-money. Second, quite surprisingly, bid-ask spread is significantly lower for longer maturity options. This leads to the following trade-off. On the one hand, shorter maturity options give an investor a possibility to gain a pure exposure to the desired Greeks over the horizon equal to the maturity of the option. On the other hand, longer maturity options give the same exposure, but with noise (due to the effects of other, possibly unhedgeable, risk factors), and are at the same time cheaper to trade. An clever strategy should exploit this tradeoff in an efficient way.

Insert Figures 8 and 9 about here.

Note that the optimal option structures derived in the previous section depend only on the transaction costs of longer maturity options. For short maturity options, investor directly trades the desired claims on powers of returns, and hence we only need transaction costs for these instruments to compute the optimal strategy. Figure 10 shows the approximate 2 months option liquidities for all three indices.

Insert Figure 10 about here.

Now, in order to proceed with step (2), we need to specify a model for the state prices. One could proceed either by calibrating a parametric model (such as the Heston model)

⁴⁵See, Malamud and Zhang (2014) for a theoretical and empirical analysis of these issues.

to the observed option prices, or by estimating state prices non-parametrically and using their dynamics to obtain the “actual” option Greeks. Both approaches (especially the second one) are highly non-trivial to implement empirically. Here, I will follow a much simpler (but also more robust) approach and assume that the Greeks are derived from the Black-Scholes formula, but with a small modification: I use *Black-Scholes Greeks, adjusted for transaction costs*. It turns out that, for all three indices, these adjusted Black-Scholes Greeks have a very high explanatory power for changes in the prices of the desired option structures.

I proceed as follows. At every trading date and for each of the three indices, I look at the options with two months maturity, I compute average of the two largest OTM implied volatilities: 20% moneyness OTM put and 10-12% moneyness OTM call, and denote it by $\sigma_{i,t}$.⁴⁶ Then, defining the Black-Scholes state prices *one month ahead*⁴⁷

$$\eta_{(i,1)}(y, X_{i,t}, \sigma_{i,t}) = e^{-r} \frac{1}{\sqrt{2\pi\sigma_{i,t}^2}} e^{-0.5(\log y - \log(X_{i,t}) - (r-q-0.5\sigma_{i,t}^2))^2 / (2\sigma_{i,t}^2)} \frac{1}{y}$$

and the corresponding state prices Greeks

$$\begin{aligned} \eta_{(i,1)}^{(0)}(y, X_{i,t}, \sigma_{i,t}) &\equiv \eta_{(i,\theta)}(y, X_{i,t}, \sigma_{i,t}), \quad \eta_{(i,\theta)}^{(1,0)}(y, X_{i,t}, \sigma_{i,t}) \equiv \frac{\partial}{\partial X_{i,t}} \eta_{(i,\theta)}(y, X_{i,t}, \sigma_{i,t}) \\ \eta_{(i,\theta)}^{(2,0)}(y, X_{i,t}, \sigma_{i,t}) &\equiv \frac{\partial^2}{\partial X_{i,t}^2} \eta_{(i,\theta)}(y, X_{i,t}, \sigma_{i,t}), \quad \eta_{(i,\theta)}^{(0,1)}(y, X_{i,t}, \sigma_{i,t}) \equiv \frac{\partial}{\partial \sigma_{i,t}} \eta_{(i,\theta)}(y, X_{i,t}, \sigma_{i,t}). \end{aligned} \quad (45)$$

Then, we use the Carr-Madan formula to build two-months claims on⁴⁸

$$\tilde{\Phi}_{(i,\theta)}^{(\mathbf{k})}(y, X_{i,t}, \sigma_{i,t}) \equiv \Phi_{(i,\theta)}^{(\mathbf{k})}(y, X_{i,t}, \sigma_{i,t}) - \Phi_{(i,\theta)}^{(\mathbf{k})}(X_{i,t}, X_{i,t}, \sigma_{i,t}) - (y - X_{i,t}) \frac{\partial}{\partial y} \Phi_{(i,\theta)}^{(\mathbf{k})}(X_{i,t}, X_{i,t}, \sigma_{i,t}),$$

where $\Phi_{(i,\theta)}^{(\mathbf{k})}(y, X_{i,t}, \sigma_{i,t}) = \eta_{(i,\theta)}^{(\mathbf{k})}(y, X_{i,t}, \sigma_{i,t}) / \Lambda_{(i,2)}(y, X_{i,t})$.⁴⁹ Then, I compute the “candidate Greeks”

$$\Gamma_{(1)}(\Phi_{(i,\theta)}^{(\mathbf{k})}(y, X_{i,t}, \sigma_{i,t})) \equiv \int_{\mathbb{R}} \eta_{(i,\theta)}^{(1)}(y, X_{i,t}, \sigma_{i,t}) \frac{1}{\Lambda_{(i,2)}(y, X_{i,t})} \Phi_{(i,\theta)}^{(\mathbf{k})}(y, X_{i,t}, \sigma_{i,t}).$$

⁴⁶Similar results are obtained with the ATM implied volatility, but it often underestimates the “true” volatility level.

⁴⁷In agreement with the formulas above, we need here the “future” state prices, hence the one month horizon. In the formulas below, I consciously suppress the dependence on interest rate r and dividend yield q .

⁴⁸We purposely remove the pure X -exposure from the claim, so that the claim can be replicated purely by the OTM options, without using the underlying itself.

⁴⁹In the list of optimal option structures, we will not be considering the claim that has the state prices vega, $\frac{\partial}{\partial \sigma_{i,t}} \eta_{(i,\theta)}(y, X_{i,t}, \sigma_{i,t})$: due to the degeneracy of the Black-Scholes model, $\frac{\partial^2}{\partial y^2} \frac{\partial}{\partial \sigma_{i,t}} \eta_{(i,\theta)}(y, X_{i,t}, \sigma_{i,t})$ and $\frac{\partial^2}{\partial y^2} \frac{\partial^2}{\partial X_{i,t}^2} \eta_{(i,\theta)}(y, X_{i,t}, \sigma_{i,t})$ are proportional.

Note that these are *Black-Scholes Greeks, adjusted for transaction costs*. That is, states that are deeper out-of-the money obtain a smaller weight in computing the Greeks.

The structure of the claims $\tilde{\Phi}_{(i,\theta)}^{(\mathbf{k})}(y, X_{i,t}, \sigma_{i,t})$ is quite stable over time and also across the three stock indices, and Figure 11 shows this structure, with claims properly rescaled to be of comparable size. As we can see, all three optimal claims can be very closely approximated by a combination of a constant and three powers of returns. This means that (i) we need to trade claims on cubic returns; (2) the transition matrix $\Omega(X_{t-1}, X_t)$ can be easily computed, and the corresponding representation is accurate.

Suppose now that today is time period t and consider now the time $t + 1$ -price $P_{t+1}(\Phi_{(i,\theta)}^{(\mathbf{k})}(y, X_{i,t}, \sigma_{i,t}))$ of the contingent claim with maturity of one month⁵⁰ (as of time $t + 1$).⁵¹ If the Black-Scholes Greeks are indeed capturing the true exposure, the approximation

$$P_{t+1}(\Phi_{(i,\theta)}^{(\mathbf{k})}(y, X_{i,t}, \sigma_{i,t})) \approx a + b \left(\Gamma_{(0)}(\Phi_{(i,\theta)}^{(\mathbf{k})}(y, X_{i,t}, \sigma_{i,t})) + (X_{i,t+1} - X_{i,t})\Gamma_{(1,0)}(\Phi_{(i,\theta)}^{(\mathbf{k})}(y, X_{i,t}, \sigma_{i,t})) + 0.5(X_{i,t+1} - X_{i,t})^2\Gamma_{(2,0)}(\Phi_{(i,\theta)}^{(\mathbf{k})}(y, X_{i,t}, \sigma_{i,t})) \right), \quad (46)$$

should have a very high R^2 , and the coefficients a and b can be estimated from a linear in-sample regression. Indeed, this regression leads to a very high R^2 , for all three stock indices, and all three $\tilde{\Phi}$ claims: all 9 combinations (stock index, claim), the R^2 is above 75%. Furthermore, the coefficient a is insignificant and coefficient b is very stable with respect to the choice of the regression window. Figures 12-14 shows how closely the Greek-based approximation follows the true realized prices. The largest component of the unexplained part is driven by changes in volatility. One could extract the true realized volatility from the regression errors, then develop a forecasting model for this volatility, and use it to improve the strategy performance. We leave this important topic for future research.⁵²

In order to be able to apply the general methodology, it remains to complete step (3). To this end, we need to develop a model for forecasting cubic returns. This is non-trivial because, in contrast to squared returns, cubic returns provide directional views and hence we need to have a reasonable forecast for returns themselves. In addition, a forecast of cubic returns need to take into account the very high leverage that the corresponding claims provide, as well as their relationship with volatility. Suppose that we have a time series of forecasts $(\alpha_{i,t})$ for expected returns on stock index i at time $t + 1$. I will use

⁵⁰I compute it as the mean of bid and ask prices.

⁵¹Note that, since the “actual” volatility is not observable, we are not including it here. If this true volatility $\tilde{\sigma}_{i,t+1}$ were observable, we could consider adding the term $(\tilde{\sigma}_{i,t+1} - \sigma_{i,t})\Gamma_{(1,0)}(\Phi_{(i,\theta)}^{(\mathbf{k})}(y, X_{i,t}, \sigma_{i,t}))$.

⁵²There is also imperfect replication noise due to discreteness and finiteness of the range of available strikes.

the following model. Let also $B_{it} < A_{it}$ be the bid and ask prices for the squared return claim, and $\pi_{i,t} = 0.5(A_{it} + B_{it})$. I will use the following linear model

$$(R_{i,t,t+25})^3 = a_0 + a_1 B_{it} + a_2 A_{it} + a_3 \alpha_{it} \pi_{it} + a_4 \text{sign}(\alpha_{it}) \pi_{it}^{3/2} + \varepsilon_t.$$

Here, $R_{i,t,t+25}$ is the 25-day return on stock index i , without including dividends.⁵³

The intuition for each of the terms is as follows: coefficients a_1 and a_2 capture the leverage effect and the exposure of the cubic claim to volatility. Coefficients a_3 and a_4 capture the directional exposure to returns, whose magnitude is determined by the market expectation of squared returns.⁵⁴ It remains to find a good model α_{it} for forecasting stock index returns. Here, we will exploit two facts: (1) dividend yields forecast stock returns; (2) stock option prices contain information about future dividends. Observation (2) is a well-known fact: using put-call parity, one can extract from option prices the present value of the dividends to be paid over the life span of the option. See, van Binsbergen, Brandt and Kojien (2012). I will denote by $\text{Strip}_t^{(i,\theta)}$ the value of of the dividend strip over the period of θ months. Now, doing the in-sample test over just one year of data is not anymore sufficient: the period is too short. We use the period 2004-2007 (about 1000 days) to run the predictive regression and construct the coefficients of the linear model that we then use to build out of sample forecasts for the period 2008-2013 (about 1400 days). It is important to note that Nasdaq 100 is a special index in my instrument set: being too small, dividend data is irregular and noisy. In particular, we find that Nasdaq dividend yields and dividend strips only add noise to the predictive regressions. Therefore, we do not use Nasdaq dividend data in my return forecasts. We run the following regressions on daily data,

$$R_{j,t,t+25} = a_0 + a_1 q_{1,t} + a_3 q_{3,t} + a_4 \text{Strip}_t^{(i,1)} + a_5 \text{Strip}_t^{(i,\theta_1)} + a_4 \text{Strip}_t^{(i,\theta_2)} + a_5 \text{Strip}_t^{(i,\theta_3)}$$

for all three indices $j = 1, 2, 3$. Here, we can pick the index i of the dividend strip to be either S&P 500 or Russell 2000. We use the latter throughout this section.⁵⁵

Now, having build the forecasts, we can do study the same strategies as we did above. Note that we are now trading third order Greeks (i.e., exposure to third powers of the stock indices), both directly though claims with one month maturity, and indirectly, through the two month claims and their exposure. Therefore, we add “3rd order” to the names of the corresponding portfolios. The next Table presents basic characteristics of the different third-order Greek efficient portfolios.

⁵³That is, quotient of prices minus one. Including dividends has a minor impact on the results.

⁵⁴This regression is robust to small modification. For example, a simpler linear model

$$(R_{i,t,t+25})^3 = a_0 + a_1 (B_{it} + A_{it}) + a_3 \alpha_{it} \pi_{it} + \varepsilon_t$$

gives comparable results with a slightly lower Sharpe ratio.

⁵⁵The results with S&P 500 are similar. Dividend strips are always normalized by the maturity, and maturities $\theta_j, j = 1, 2, 3$, correspond to the nearest 3 maturities of the options maturing on the third Saturday of each month.

Portfolio Type	Sharpe Ratio	Skewness	Kurtosis
Greek efficient, managing TrCosts	1.17	3.1	16.3
Greek efficient, managing TrCosts, CRRA	1.25	1.9	10.4
Greek efficient, managing TrCosts, CRRA, with views	1.25	2.21	10.8
Greek-efficient, managing TrCosts, with views	1.23	3.26	17.3

Here, “with views” means that, in addition to the forecasts of expected returns, we also incorporate the Gaussian updating based on squared returns’ forecasts. The first observation is that the improvement in the mean-variance efficiency due to using third order claims and management of transaction costs is significant: Sharpe ratio goes up by about 30%. Intuitively, we would also expect that the inclusion of claims on higher powers of returns implied even more non-linear profile of returns. As a result, we would anticipate even more non-Gaussianity in returns. However, quite surprisingly, exactly the opposite happens: kurtosis goes down from 27 to about 16. And, at the same time, skewness also drops significantly, from about four to 3. Managing higher order moments (with CRRA preferences) leads to an even lower kurtosis, but the skewness also drops, and the increase in the Sharpe ratio is quite small. In particular, these results mean that using kurtosis as a measure of “non-Gaussianity” may be misleading: more non-linearity may actually lead to a lower kurtosis. Incorporating additional views does not bring anything this time.

Figure 15 shows the cumulative performance of the four strategies. As above, we normalize the returns by their ex-post realized volatility to achieve a 20% annualized volatility. The out-of-sample performance of the strategy is quite strong, and it clearly outperforms the simpler second order Greek efficient strategy.

Insert Figure 15 about here.

9 Bibliography

- Ait-Sahalia, Y., and M. W. Brandt. 2001. Variable Selection for Portfolio Choice. *Journal of Finance* 56: 1297-1351.
- Ait-Sahalia, Y., and M. W. Brandt. 2007. Consumption and Portfolio Choice with Option Implied State Prices. *Working paper*.
- Ait-Sahalia, Y. and A. Lo. 1998. Nonparametric Estimation of State-Price Densities Implicit in Financial Asset Prices. *Journal of Finance* 53: 499-547.
- Bajeux-Besnainou, I., and R. Portait. 1998. Dynamic Asset Allocation in a Mean-Variance Framework. *Management Science*: 44, 79-95.
- Bansal, R., M. Dahlquist, and C.R. Harvey. 2004. Dynamic Trading Strategies and Portfolio Choice. *Working Paper, Duke University*.

- Bansal, R., and D. Kiku. 2011. Cointegration and Long-Run Asset Allocation. *Journal of Business and Economic Statistics* 29(1): 161-173. Duke University.
- Barberis, N. 2000. Investing for the Long Run When Returns Are Predictable. *Journal of Finance* 55: 225-264.
- Basak, S., and G. Chabakauri. 2010. Dynamic Mean-Variance Asset Allocation. *Review of Financial Studies* 23: 2970-3016.
- Basak, S., and G. Chabakauri. 2012. Dynamic Hedging in Incomplete Markets: A Simple Solution. *Review of Financial Studies* 25: 1845-1896.
- Black, F., and R. Litterman. 1990. Asset Allocation: Combining Investor Views with market equilibrium. *Goldman Sachs Fixed Income Research*.
- Black, F. and M. Scholes. 1973. The Pricing of Options and Corporate Liabilities. *Journal of Political Economy* 81(3): 637-654.
- Brandt, M.W. 1999. Estimating Portfolio and Consumption Choice: A Conditional Euler Equations Approach. *Journal of Finance* 54: 1609-1645.
- Brandt M.W., Goyal A., Santa-Clara P., and J.R. Stroud. 2005. A Simulation Approach to Dynamic Portfolio Choice with an Application to Learning About Return Predictability. *Review of Financial Studies* 18(3): 831-873.
- Brandt, M.W., and P. Santa-Clara. 2006. Dynamic Portfolio Selection by Augmenting the Asset Space. *Journal of Finance* 61: 2187-2218.
- Brandt, M. W., P. Santa-Clara, and R. Valkanov. 2009. Parametric Portfolio Policies: Exploiting Characteristics in the Cross Section of Equity Returns. *Review of Financial Studies* 22: 3411-3447.
- van Binsbergen, J., M. Brandt, and R. Koijen. 2012. "On the Timing and Pricing of Dividends." *American Economic Review* 102(4): 1596-1618.
- Breeden, D. T. and R. H. Litzenberger. 1978. Prices of State-Contingent Claims Implicit in Option Prices. *Journal of Business* (51)4: 621-651.
- Brennan, M. J., and Y. Xia. 2002. Dynamic Asset Allocation under Inflation. *Journal of Finance* 57: 1201-1238.
- Campbell, J.Y., and L.M. Viceira. 1999. Consumption and Portfolio Decisions when Expected Returns are Time Varying. *Quarterly Journal of Economics* 114: 433-495.
- Campbell, J.Y., and L.M. Viceira. 2002. Strategic Asset Allocation: Portfolio Choice for Long-Term Investors. Oxford University Press.

- Cao, M., and J. Wei. 2010. Option Market Liquidity: Commonality and Other Characteristics. *Journal of Financial Markets* 13: 20-48.
- Carr, P., and D. Madan. 1998. Towards a theory of volatility trading; in R.A. Jarrow, ed.: *Volatility: New estimation Techniques for Pricing Derivatives* (RISK Publications, London).
- Chacko, G., and L.M. Viceira. 2005. Dynamic Consumption and Portfolio Choice with Stochastic Volatility in Incomplete Markets. *Review of Financial Studies* 18: 1369-1402.
- Cochrane, J.H. 2008. A Mean Variance Benchmark for Intertemporal Portfolio Theory. *Working Paper, University of Chicago*.
- Collin-Dufresne, P., K. Daniel, C. Moallemi, and M. Saglam. 2013. Strategic Asset Allocation with Predictable Returns and Transaction Costs. *Working Paper*.
- Constantinides, G. M., J. C. Jackwerth, and A. Z. Savov. 2009. The Puzzle of Index Option Returns. *Working Paper*.
- Cox, J.C., and C.-F. Huang. 1989. Optimal Consumption and Portfolio Policies when Asset Prices follow a Diffusion Process. *Journal of Economic Theory* 39: 33-83.
- Cvitanic, J. and I. Karatzas. 1992. Convex Duality in Constrained Portfolio Optimization, *Annals of Applied Probability* 2(4): 767-818.
- Cvitanic, J., L. Goukasian, and F. Zapatero. 2003. Monte Carlo Computation of Optimal Portfolios in Complete Markets. *Journal of Economic Dynamics and Control* 27: 971-986.
- Cvitanic, J., A. Lazrak, L. Martellini, and F. Zapatero. 2006. Dynamic Portfolio Choice with Uncertainty and the Economic Value of Analysts Recommendations. *Review of Financial Studies* 19: 1113-1156.
- DeMiguel, V., Y. Plyakha, R. Uppal, and G. Vilkov. 2013. Improving Portfolio Selection Using Option-Implied Volatility and Skewness. *Forthcoming in Journal of Financial and Quantitative Analysis*.
- Detemple, J.B., R. Garcia, and M. Rindisbacher. 2003. A Monte Carlo Method for Optimal Portfolios. *Journal of Finance* 58: 401-446.
- Driessen, J., and P. Maenhout. 2007. An Empirical Portfolio Perspective on Option Pricing Anomalies. *Review of Finance* 11: 561-603.
- Duffie, D., and H. Richardson. 1991. Mean-Variance Hedging in Continuous Time. *Annals of Probability* 1: 1-15.

- Eraker, B. 2007. The Performance of Model Based Option Trading Strategies. *Working Paper*.
- Faias, J. A., and P. Santa-Clara. 2011. Optimal Option Portfolio Strategies. *Working paper*.
- Ferreira, M., and P. Santa-Clara. 2011. Forecasting Stock Market Returns: The Sum of the Parts is More Than the Whole. *Journal of Financial Economics* 100: 514-537.
- Garlappi, L. and G. Skoulakis. 2010. Solving Consumption and Portfolio Choice Problems: The State Variable Decomposition Method. *Review of Financial Studies* 23(9): 3346-3400.
- Garleanu, N., and L. H. Pedersen. 2012. Dynamic Trading with Predictable Returns and Transaction Costs. *Journal of Finance, forthcoming*.
- Garleanu, N., and L. H. Pedersen. 2013. Dynamic Portfolio Choice with Frictions. *Working paper*.
- George, T. J. and F. Longstaff. 1993. Bid-Ask Spreads and Trading Activity in the S&P 100 Options Market. *Journal of Financial And Quantitative Analysis* 28(3): 381-397.
- Goyal, A. and Saretto A. 2009. Cross-Section of Option Returns and Volatility. *Journal of Financial Economics* 94(2): 310-326.
- Goyal A., and I. Welch. 2008. A Comprehensive Look at the Empirical Performance of Equity Premium Prediction. *Review of Financial Studies* 21(4): 1455-1508.
- Grinold, R. 2006. A Dynamic model of portfolio management. *Journal of Investment Management* 4: 5-22.
- Heaton, J., and D. Lucas. 1996. Evaluating the Effects of Incomplete Markets on Risk Sharing and Asset Pricing. *Journal of Political Economy* 104, 443-487.
- Heston, S.L. 1993. A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options. *Review of Financial Studies* 6: 327-343.
- Hlawitschka, W. 1994. The Empirical Nature of Taylor-Series Approximations to Expected Utility. *American Economic Review* 84: 713-719.
- Karatzas, I., J. P. Lehoczky, and S. E. Shreve. 1987. Optimal Portfolio and Consumption Decisions for a Small Investor on a Finite Horizon. *SIAM Journal of Control and Optimization* 25: 1557-1586.
- Jackwerth, J., and G. Vilkov. 2013. Asymmetric Volatility Risk: Evidence from Option Markets. *Working paper*.

- Kim, T.S., and E. Omberg. 1996. Dynamic Nonmyopic Portfolio Behavior. *Review of Financial Studies* 9: 141-161.
- Kroll, Y., H. Levy, and H. M. Markowitz. 1984. Mean-Variance Versus Direct Utility Maximization. *Journal of Finance* 39: 47-75.
- Ledoit, O. and M. Wolf. 2004. Honey, I Shrunk the Sample Covariance Matrix. *Journal of Portfolio Management* 30(4): 110-119.
- Leippold, M., F. Trojani, and P. Vanini. 2004. Geometric Approach to Multiperiod Mean-Variance Optimization of Assets and Liabilities. *Journal of Economic Dynamics and Control* 28: 1079-1113.
- Liu, J. 2007. Portfolio Selection in Stochastic Environments. *Review of Financial Studies*: 20, 1-39.
- Liu, J., and J. Pan. 2003. Dynamic Derivative Strategies. *Journal of Financial Economics* 69: 401-430.
- Loistl, O. 1976. The Erroneous Approximation of Expected Utility by Means of a Taylors Series Expansion: Analytic and Computational Results. *American Economic Review* 66: 904-910.
- Longstaff, F. A. and Schwartz, E. S. 2001. Valuing American Options by Simulation: A Simple Least-Squares Approach. *Review of Financial Studies* 14(1): 113-147.
- Malamud, S., and Y. Zhang. 2014. The Option Liquidity Surface. Working paper.
- Markowitz, H. 1952. Portfolio Selection. *Journal of Finance* 7(1): 77-91.
- Markowitz, H. M. 1991. Foundations of Portfolio Theory. *Journal of Finance* 46: 469-477.
- Merton, R.C. 1971. Optimum Consumption and Portfolio Rules in a Continuous Time Model. *Journal of Economic Theory* 3: 373-413.
- Plyakha, Y., and G. Vilkov. 2008. Portfolio Policies with Stock Options. *Working Paper*.
- Samuelson, P. 1970. The Fundamental Approximation Theorem of Portfolio Analysis in Terms of Means, Variances and Higher Moments. *Review of Economic Studies* 37(4): 537-542.
- Song, Z., and D. Xiu. 2013. A Tale of Two Option Markets: Pricing Kernels and Volatility Risk. *Working paper, University of Chicago*.
- Wachter, J.A. 2002. Portfolio and Consumption Decisions under Mean-Reverting Returns: An Explicit Solution for Complete Market. *Journal of Financial and Quantitative Analysis* 37: 63-91.

A Some calculations

Proof of Lemma 6.1. We have

$$\begin{aligned}
E_t[\mathbf{W}_{t+1}(\mathbf{f}_t)] - e^r \beta_t(\mathbf{f}_t) &\approx \sum_{(j,\theta_j), \theta_j > 1} \sum_{\mathbf{k}} \mu_{\mathbf{k}}(X_t) \int_{\mathbb{R}} \eta_{(j,\theta_j-1)}^{(\mathbf{k})}(X_t, y) f_{(j,\theta_j)}(X_t, y) dy \\
&+ \sum_j (E_t[f_{t,(i,1)}(X_t, X_{j,t+1})] - e^r \beta_t(f_{t,(j,1)}(X_t, y))) = \mu(X_t)^T \mathbf{\Gamma}(X_t, \mathbf{f}_t)
\end{aligned} \tag{47}$$

■

B Dynamic Mean Variance Optimal Portfolios

In this section, I will consider the problem of an investor whose goal is to optimize mean-variance preferences over terminal wealth w_T at time t . Investor's time t value function is given by

$$J(w_t, X_t) = E_t[w_T] - \frac{\gamma}{2} \text{Var}_t[w_T].$$

Following Basak and Chabakauri (2010), I assume that the investor is time consistent.⁵⁶ As Basak and Chabakauri (2010) show, the time-consistent value function satisfies the backward recursion

$$J(w_t, X_t) = E_t[J(t+1, w_{t+1}, X_{t+1})] - \frac{\gamma}{2} \text{Var}_t[E_{t+1}[w_T]]. \tag{48}$$

Denote by $\mathbf{f}_t^* = (f_{i,\theta}^*(t, X_t, y))_{(i,\theta)}$ the future optimal policy,⁵⁷ and let⁵⁸

$$\tilde{v}_{t+1}^* \equiv \mathbf{W}_{t+1}(\mathbf{f}_t^*) - e^r \beta_t(\mathbf{f}_t^*)$$

be the excess return from the optimal option portfolio. Then, for any $t \leq T-2$

$$w_T = e^{rT-t-1} w_{t+1} + \sum_{l=1}^k e^{rk-l} \tilde{v}_{t+l+1}^*$$

and I define

$$v_{t+1}^* \equiv E_{t+1} \left[\sum_{l=1}^k e^{rk-l} \tilde{v}_{t+l+1}^* \right].$$

Then,⁵⁹

$$\text{Var}_t[E_{t+1}[w_T]] = e^{r2(T-t-1)} \text{Var}_t[w_{t+1}] + 2e^{rT-t-1} \text{Cov}_t[w_{t+1}, v_{t+1}^*] + \text{Var}_t[v_{t+1}^*]. \tag{49}$$

⁵⁶Our method can also be directly applied to the pre-commitment problem.

⁵⁷Note that, in general, the policy may be non-unique. In the empirical implementation of the solution, one would need to pick a policy that is least sensitive to model errors and is not too costly to trade.

⁵⁸See (2).

⁵⁹These calculations are completely analogous to those in Basak and Chabakauri (2010).

We conjecture that

$$J_{t+1} = e^{rT-t-1}w_{t+1} + L_{t+1}$$

for some \mathcal{F}_{t+1} -measurable random variables L_{t+1} . Then, (48) implies that

$$E_t[\mathbf{W}_{t+1}(f)v_{t+1}^*(X_{t+1})] = \sum_{(i,\theta)} \int V_{(i,\theta),t}^*(y)f_{(i,\theta)}(X_t, y)dy \quad (50)$$

where

$$V_{(i,\theta),t}^*(y) = E_t[v_{t+1}^*(X_{t+1})\eta_{(i,\theta-1)}(X_{t+1}, y)]$$

This quantity has a very clear meaning: this is the expected tomorrow's price of an Arrow security that pays $v_{t+1,T}^*(X_{t+1})$ at time $t+\theta$ if $X_{i,t+\theta} = y$. One could think of it as the cost of hedging the fluctuations in future investment opportunities. The first order conditions take the form⁶⁰

$$\begin{aligned} & (1 + \gamma E_t[v_{t+1,T}^*])E_t[\eta_{(j,\theta_j-1)}(X_{t+1}, y)] - e^r\eta_{(j,\theta_j)}(X_t, y) - \gamma V_{(i,\theta),t+1}^*(y) \\ & = \gamma e^{r(T-t-1)} \int p(X_t, X_{t+1})\eta_{(j,\theta_j-1)}(X_{t+1}, y)\mathbf{W}_{t+1}(\mathbf{f}_t)dX_{t+1}. \end{aligned} \quad (51)$$

Now, we can follow the same argument as in the proof of Theorem 3.7. Using the approximation

$$V_{(i,\theta),t+1}^*(y) \approx \sum_{|\mathbf{k}|\leq K} \nu_{(i,\theta),t}^{(\mathbf{k})} \eta_{(i,\theta-1)}^{(\mathbf{k})}(\bar{X}_t, y) \quad (52)$$

with

$$\nu_{(i,\theta),t}^{(\mathbf{k})} = E_t[v_{t+1}^*(X_{t+1})(X_{t+1} - \bar{X}_t)^{\mathbf{k}}], \quad (53)$$

we arrive at the following result.⁶¹

Theorem B.1 (Efficient Greeks) *The Greeks $\Gamma = (\Gamma_{(\mathbf{k})}(X_t, \mathbf{f}_t))$ of the optimal portfolio satisfy*

$$\Gamma^{\mathbf{b}} = \gamma^{-1} e^{-r(T-t-1)} (\mathbf{A}_{t,K}^T \Sigma_t \mathbf{A}_{t,K})^{-1} \mathbf{A}_{t,K}^T (\mu - \gamma(\text{Cov}_t[(X_{t+1} - \bar{X}_t)^{\mathbf{k}}, v_{t+1}^*])_{\mathbf{k} \in H}) \quad (54)$$

As in Basak and Chabakauri (2010), the multi-period mean-variance optimizing agent behaves as if he were a myopic optimizer, but with an additional exposure to risk given by the present value of future expected returns. The agent hedges this risk by introducing an additional covariance term into his optimal portfolio. The hedging demand term in (54) looks similar to solution looks case when the agent simply trades one period claims on the moments $(X_{t+1} - \bar{X}_t)^{\mathbf{k}}$. Indeed, these moments are capturing changes in option prices driven by the corresponding Greeks. The exact structure of market incompleteness and the relationship between hedgeable Greeks, captured by the matrix $\mathbf{A}_{t,K}$ determines the combinations of moments that approximate the optimal option portfolio.

⁶⁰As above, I use the normalization $E_t[\mathbf{W}_{t+1}(\mathbf{f}_t)] = 0$.

⁶¹Recall that H is the set of hedgeable Greeks.

C Proofs

Proof of Lemma 6.3. For simplicity we omit the indices (i, θ) . To prove the claim, it suffices to show that any function $g(y)$ such that $\tilde{g}(y) = g(y)\Lambda(y)$ is square integrable, can be approximated by a linear combination of optimal claims $\frac{\partial^k}{\partial X_{i,t}^k}\eta(y - X_{i,t})$, $k \geq 0$. Here, we consciously suppress the dependence on $X_{-i,t}$. Equivalently, we need to show that $h(y) \equiv \tilde{g}(y + X_{i,t})$ can be approximated by linear combinations of the derivatives $\frac{\partial^k}{\partial y^k}\eta(y)$. Let \hat{h} and $\hat{\eta}$ be the Fourier transforms of h and η respectively. Then,

$$h(y) \approx \sum_k a_k \frac{\partial^k}{\partial y^k} \eta(y) \Leftrightarrow \frac{\hat{h}(y)}{\hat{\eta}(y)} \approx \sum_k a_k (-iy)^k.$$

and the claim follows because $\frac{\hat{h}(y)}{\hat{\eta}(y)}$ can be approximated with polynomials in y . ■

Proof of Theorem 6.5. Recall the dynamic programming equation

$$\begin{aligned} V_t(\mathbf{f}_{t-1}, X_t) &= E_t[\mathbf{W}_{t+1}(\mathbf{f}_t)] - e^r \beta_t(\mathbf{f}_t) - \frac{1}{2} \gamma \text{Var}_t[\mathbf{W}_{t+1}(\mathbf{f}_t)] \\ &- \frac{1}{2} \sum_{(i,\theta)} \int \Lambda_{(i,\theta)}(X_{i,t}, y) (f_{(i,\theta+1),t-1}(X_{t-1}, y) - f_{(i,\theta),t}(X_t, y))^2 dy \\ &- \frac{1}{2} \sum_i \int \Lambda_{(i,\theta_i,\max)}(X_{i,t}, y) (f_{(i,\theta_i,\max),t}(X_t, y))^2 dy + (1 - \rho) V_{t+1}(X_{t+1}, \mathbf{f}_t) \end{aligned} \quad (55)$$

We have

$$\begin{aligned} &\frac{1}{2} \sum_{(i,\theta)} \int \Lambda_{(i,\theta)}(X_{i,t}, y) (f_{(i,\theta+1),t-1}(X_{t-1}, y) - f_{(i,\theta),t}(X_t, y))^2 dy \\ &\approx \frac{1}{2} \sum_{(i,\theta)} \int (\Lambda_{(i,\theta)}(X_t, y))^{-1} \left(\sum_{\mathbf{k}} ((\Omega^{(j,\theta_j)} \mathbf{a}_{t-1}^s)_{\mathbf{k}} - a_{t,(j,\theta_j)}^{(\mathbf{k})}) \eta_{(j,\theta_j-1)}^{(\mathbf{k})}(X_t, y) \right)^2 dy \\ &= 0.5 (\mathbf{\Omega} \mathbf{a}_{t-1}^s - \mathbf{a}_t)^T \mathbf{B}_t (\mathbf{\Omega} \mathbf{a}_{t-1}^s - \mathbf{a}_t) = 0.5 (\mathbf{B}_t \mathbf{\Psi} \kappa_{t-1}^s - \kappa_t)^T (\mathbf{B}_t)^{-1} (\mathbf{B}_t \mathbf{\Psi} \kappa_{t-1}^s - \kappa_t) \end{aligned} \quad (56)$$

Then using the just derived approximation and substituting the quadratic Ansatz for the value function, I get

$$\begin{aligned} V_t &= \max_{\kappa_t} \left\{ \left(\mathcal{I}^T \mu - (1 - \rho) E_t[\mathcal{C}_{t+1}(X_{t+1})] + \mathbf{B}_t \mathbf{\Psi} \kappa_{t-1}^s \right)^T \kappa_t \right. \\ &\left. - 0.5 \kappa_t^T \left(\gamma \mathcal{I}^T \Sigma_t \mathcal{I} + \mathbf{B}_t^{-1} + (1 - \rho) E_t[\mathcal{A}_{t+1}(X_{t+1})] \right) \kappa_t - 0.5 \kappa_{t-1}^T \mathbf{\Psi}^T \mathbf{B}_t \mathbf{\Psi} \kappa_{t-1} + (1 - \rho) E_t[L_{t+1}] \right\} \end{aligned} \quad (57)$$

The first order condition immediately gives the required expressions. ■

D Figures



Figure 1: Cumulative returns on the equally weighted index portfolio, and the returns on the index portfolio with incorporated views on squared returns. Returns are normalized to have an ex-post annualized volatility of 20%.

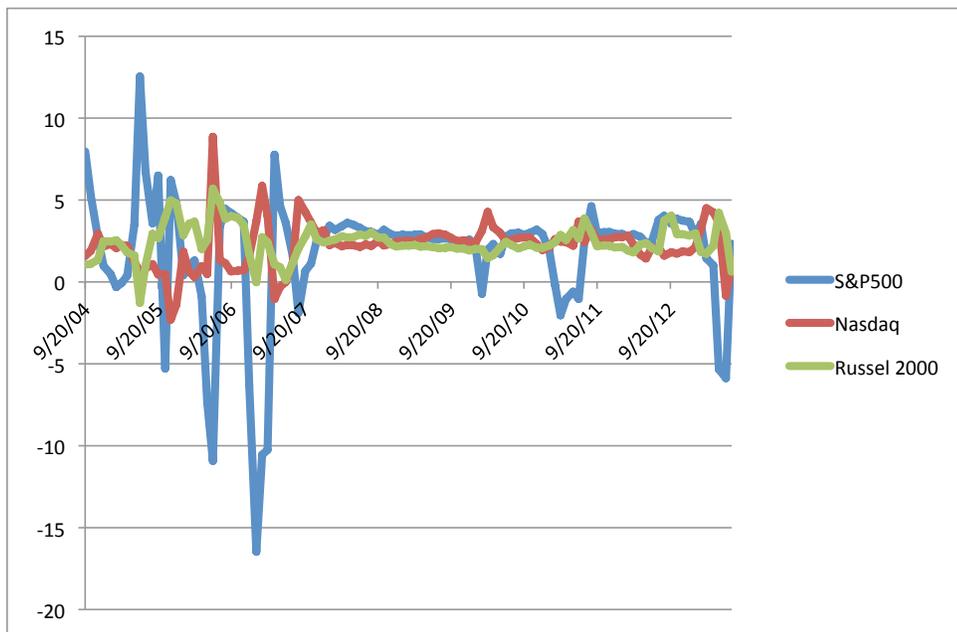


Figure 2: Portfolio weights of the three stock indices in the Black-Litterman approach, using views on squared returns. Only stocks are traded.

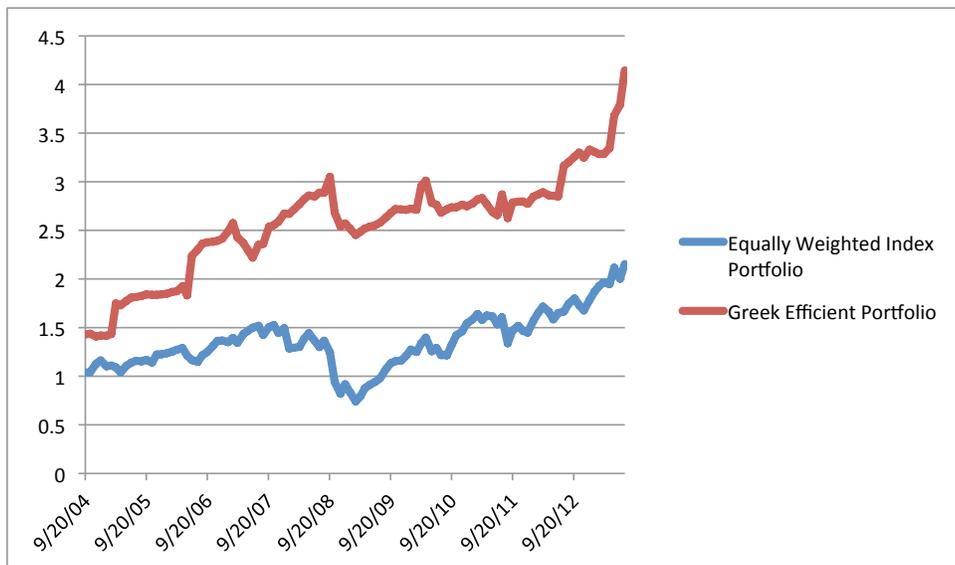


Figure 3: Cumulative returns on the equally weighted portfolio of index returns and the Greek MV-efficient portfolio, after transactions costs. Returns are normalized to have an ex-post annualized volatility of 20%.

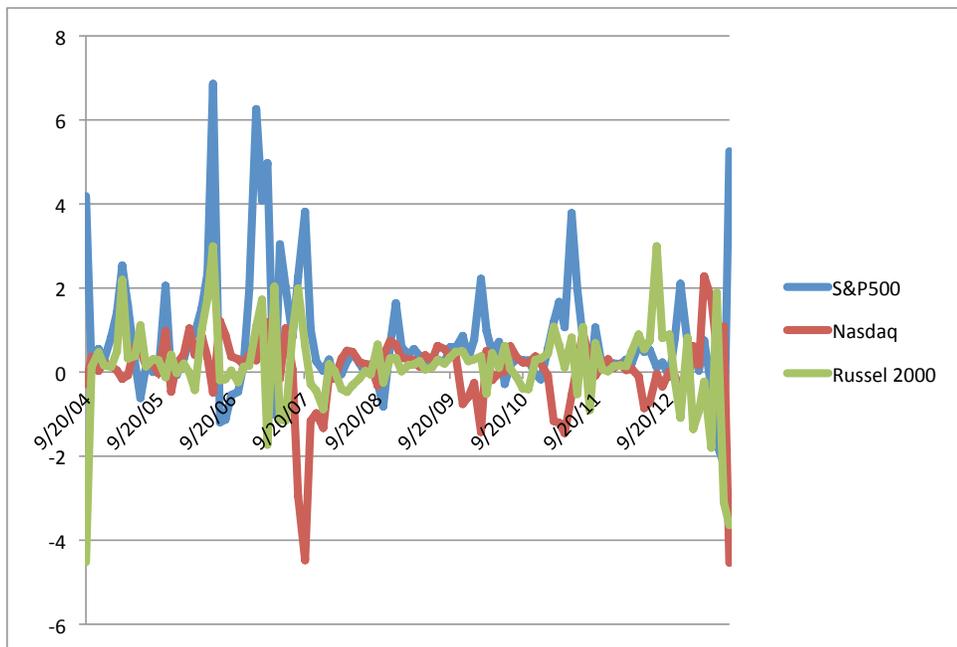


Figure 4: Portfolio weights for the three stock indices in the 6-instrument Greek MV-efficient portfolio. Mean stock returns are assumed to satisfy CAPM. Portfolio weights are normalized so that the Greek efficient portfolio has an ex-post annualized volatility of 20%.

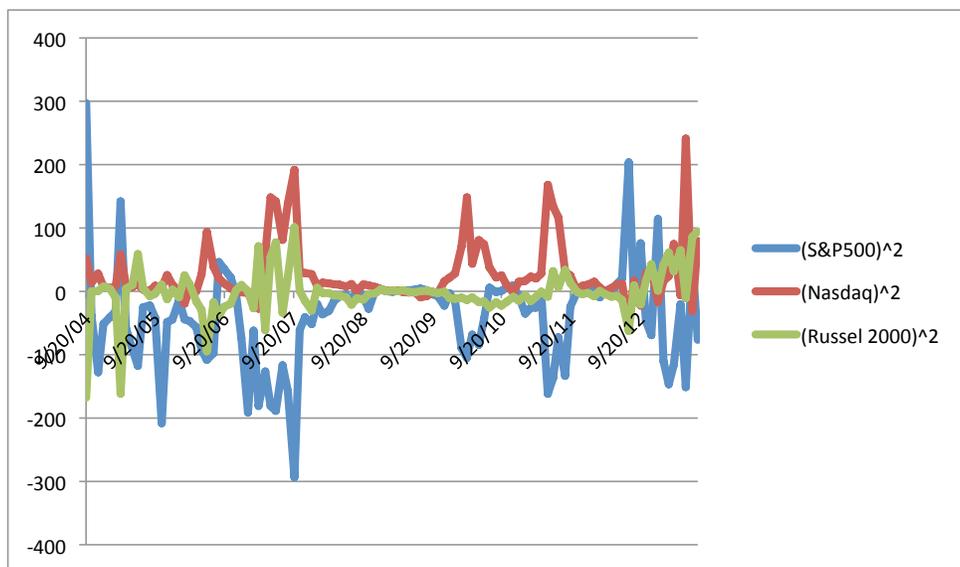


Figure 5: Portfolio weights for squared returns on the three stock indices in the 6-instrument Greek MV-efficient portfolio. Mean stock returns are assumed to satisfy CAPM. Portfolio weights are normalized so that the Greek efficient portfolio has an ex-post annualized volatility of 20%.

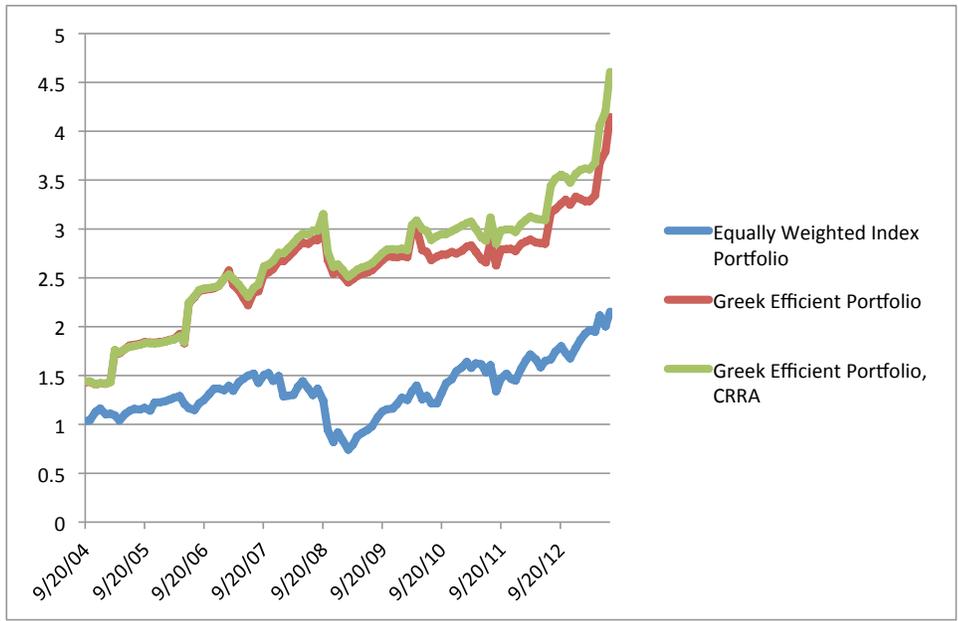


Figure 6: Cumulative returns on the equally weighted portfolio of index returns, the Greek MV-efficient portfolio, and the Greek-CRRA efficient portfolio, after transactions costs. Returns are normalized to have an ex-post annualized volatility of 20%.

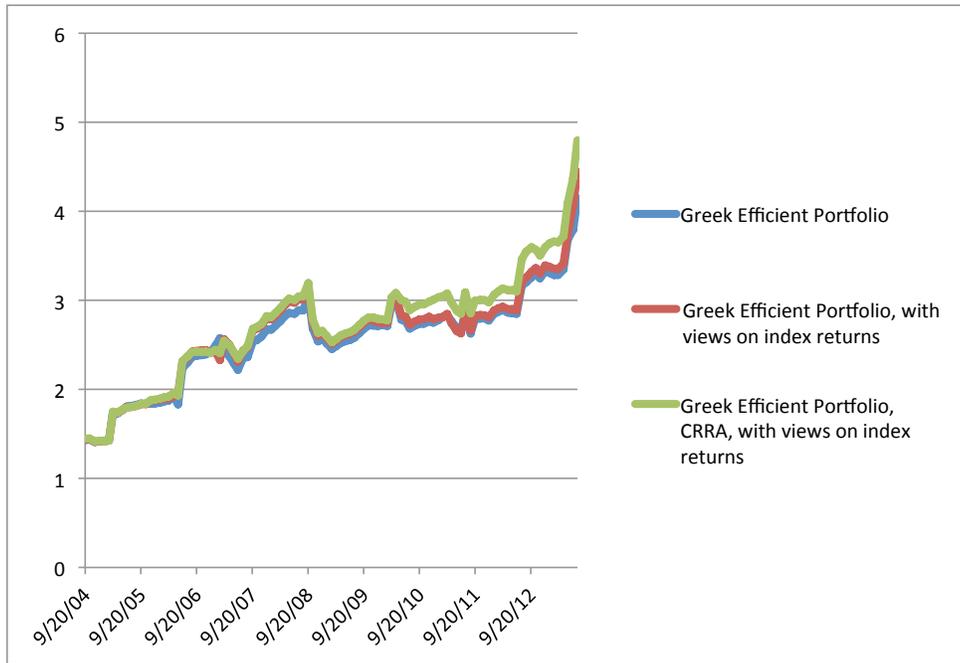


Figure 7: Cumulative returns on index portfolio, the Greek MV-efficient portfolio, and the Greek-CRRA efficient portfolio, after transactions costs. For all three portfolios, forecasts of squared returns are used to update expectations of index returns. Returns are normalized to have an ex-post annualized volatility of 20%.

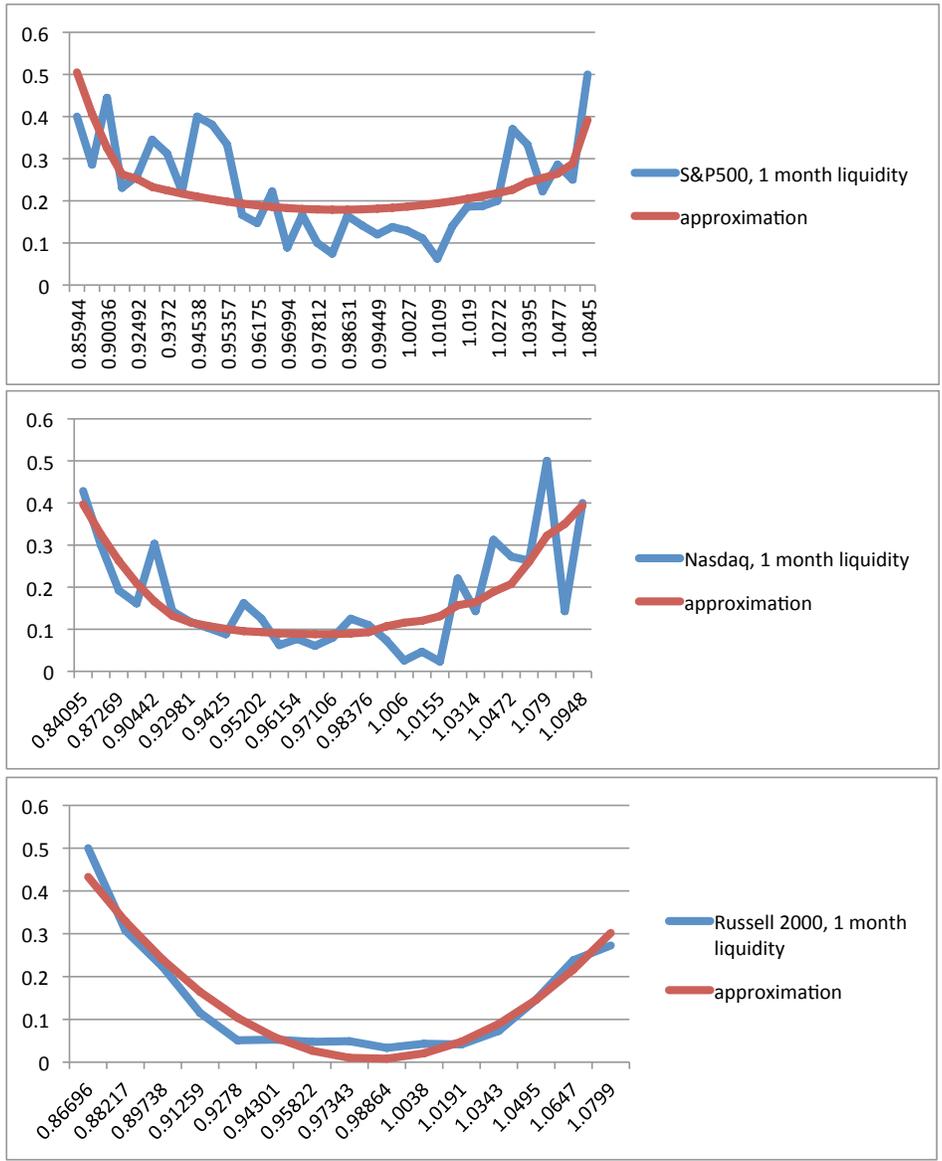


Figure 8: Option liquidity and the quadratic approximation, as a function of strike/price. All three stock indices. Date = 22/08/2005, Maturity = 17/09/2005.

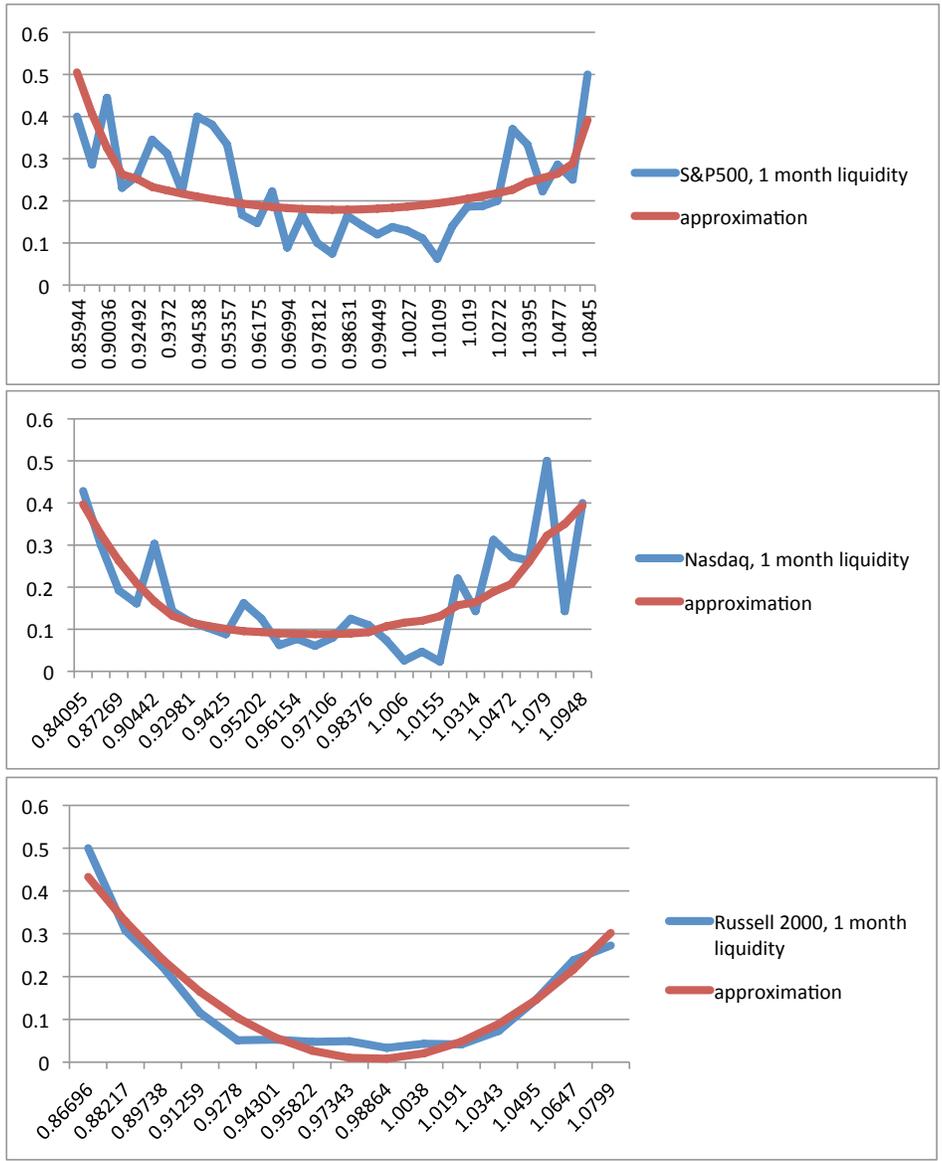


Figure 9: Option liquidity and the quadratic approximation, as a function of strike/price. All three stock indices. Date = 22/08/2005, Maturity = 22/10/2005.

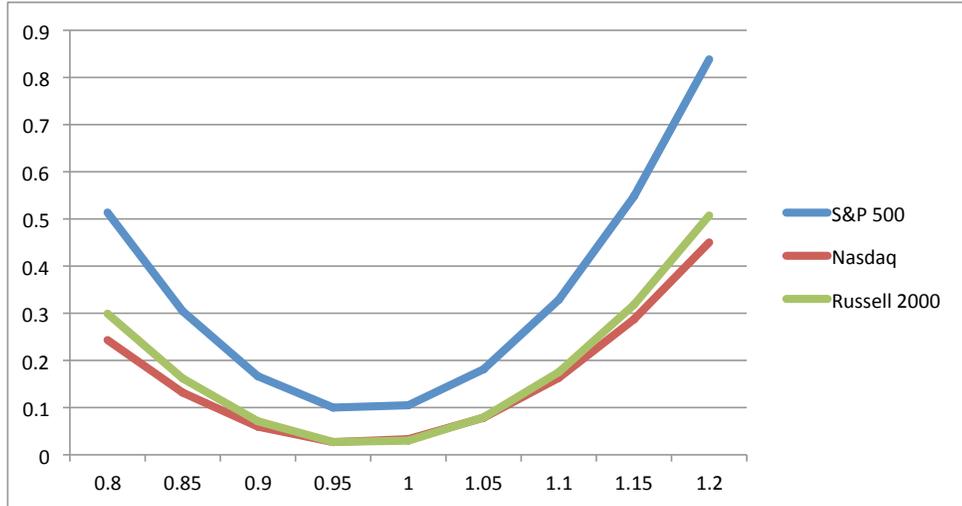


Figure 10: Approximate average option liquidates, as a function of strike/price. All three stock indices.

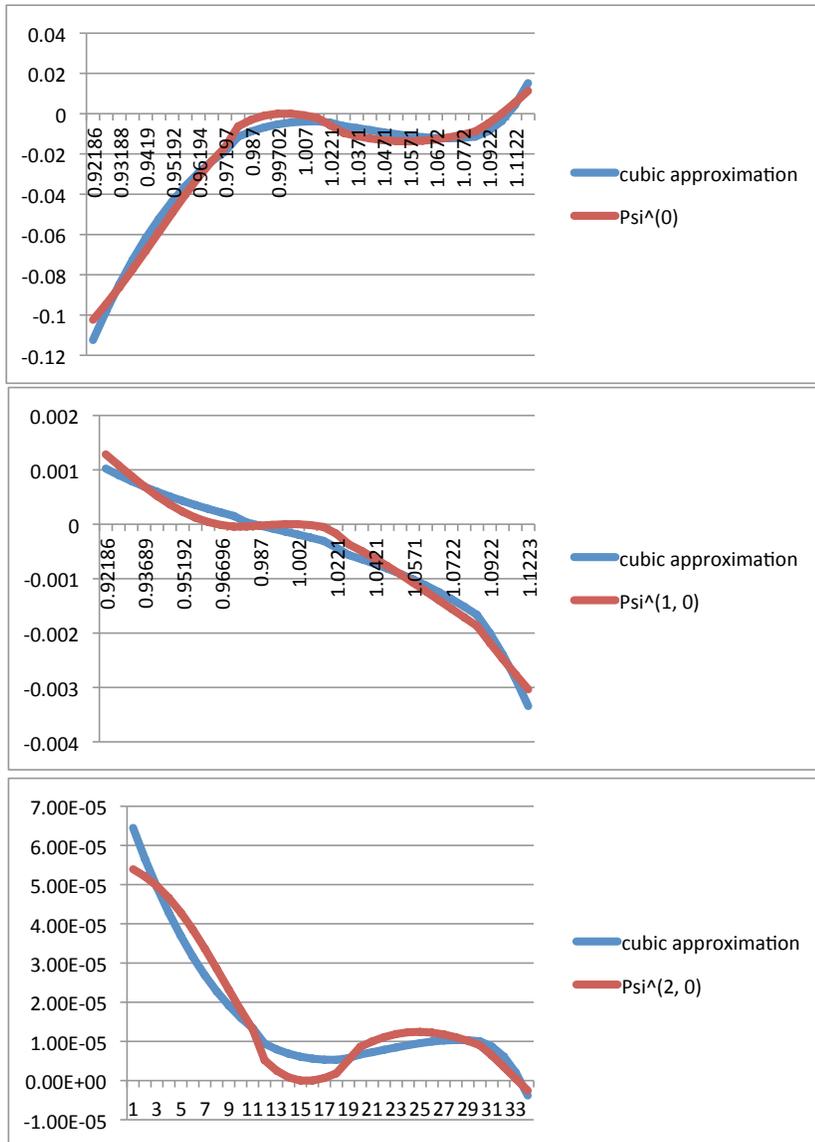


Figure 11: The $\tilde{\Psi}^{(k)}$ claims and their cubic approximations.

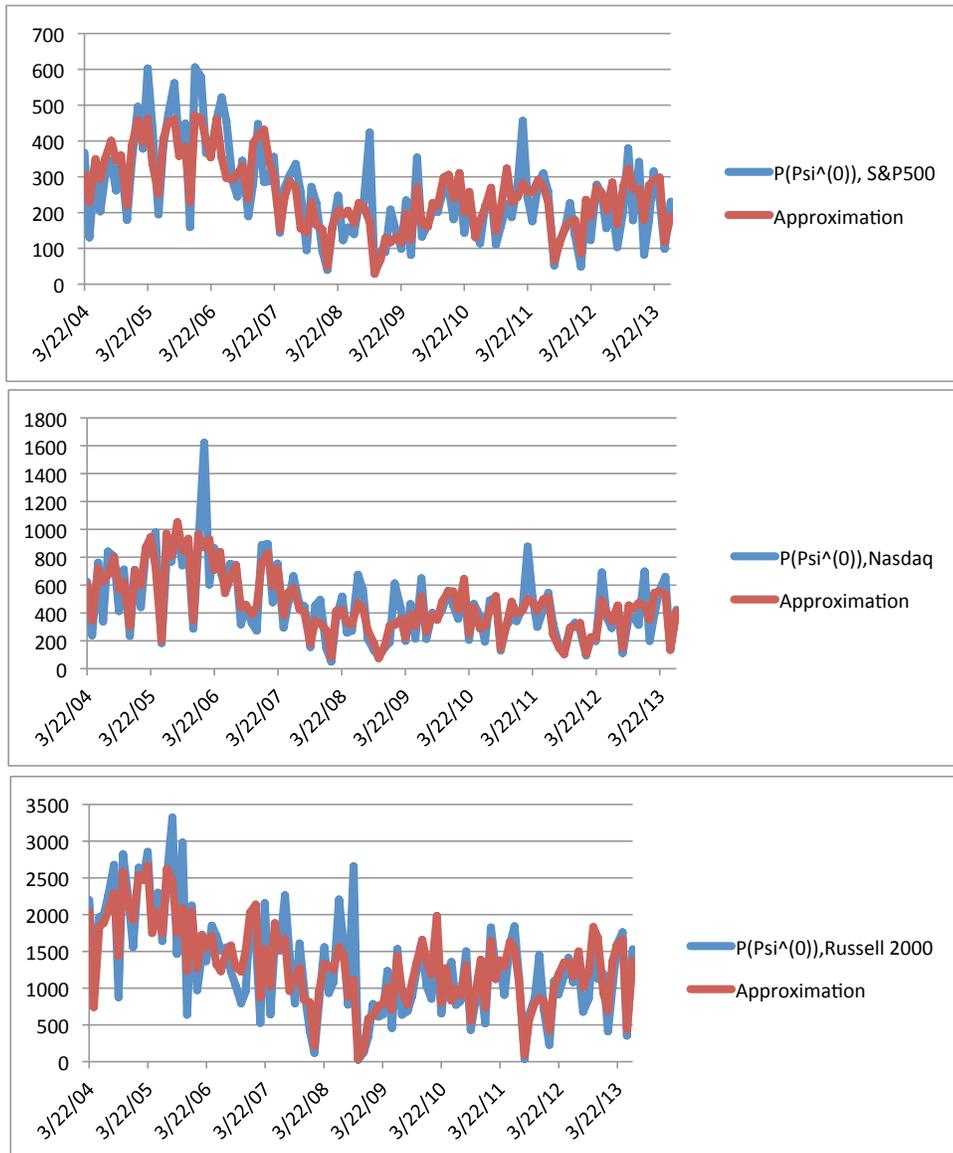


Figure 12: $P_{t+1}(\Psi_{(i,\theta)}^{(0)}(y, X_{i,t}, \sigma_{i,t}))$ and its Black-Scholes-based approximation.

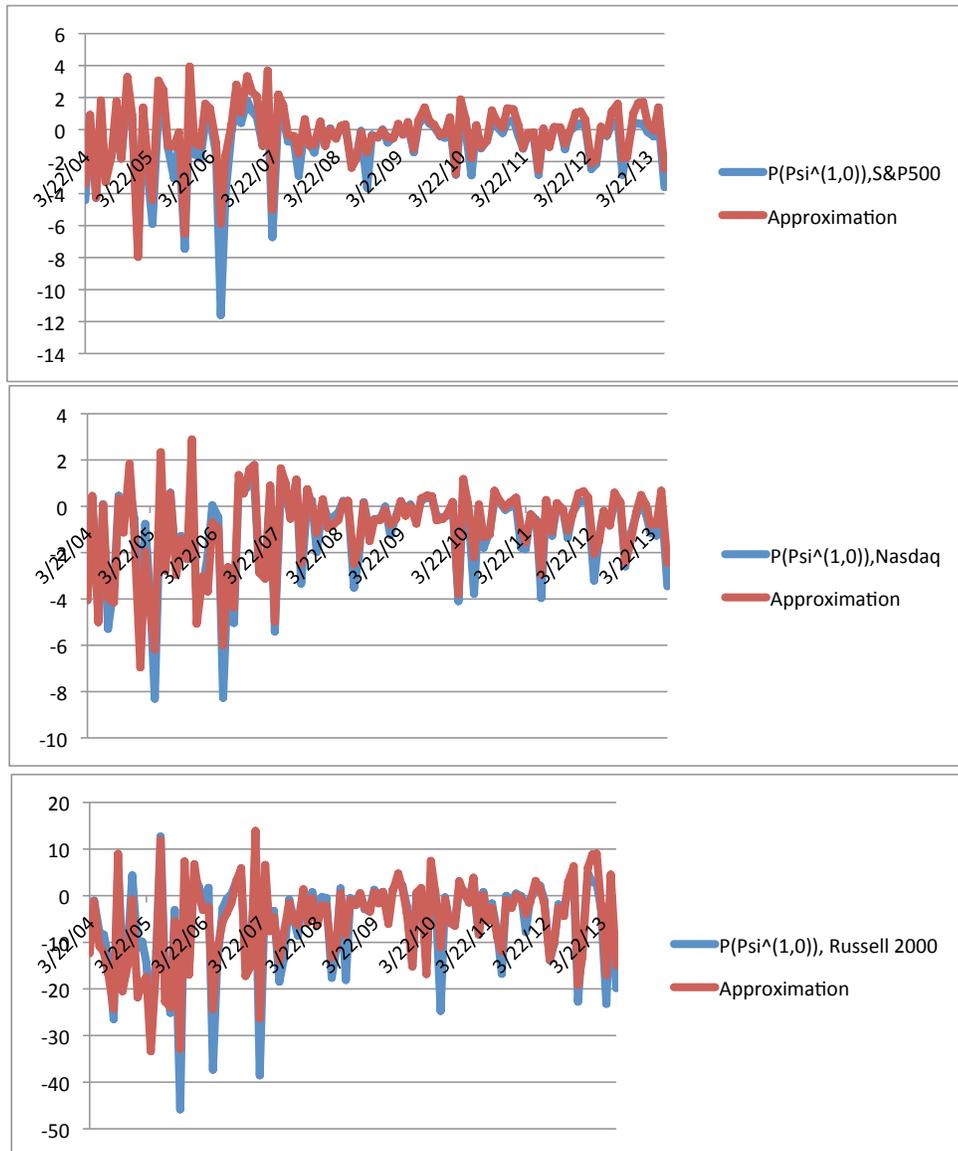


Figure 13: $P_{t+1}(\Psi_{(i,\theta)}^{(1,0)}(y, X_{i,t}, \sigma_{i,t}))$ and its Black-Scholes-based approximation.

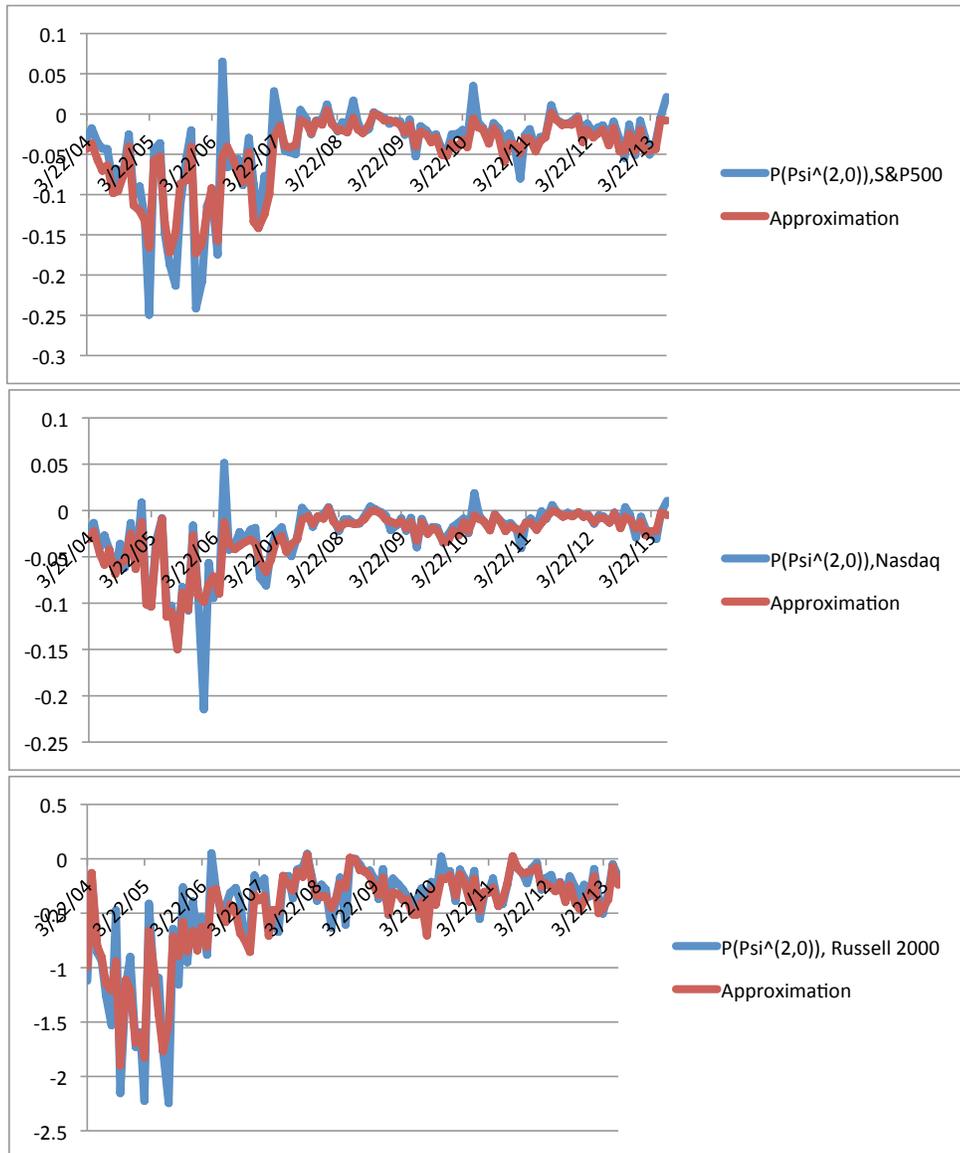


Figure 14: $P_{t+1}(\Psi_{(i,\theta)}^{(2,0)}(y, X_{i,t}, \sigma_{i,t}))$ and its Black-Scholes-based approximation.

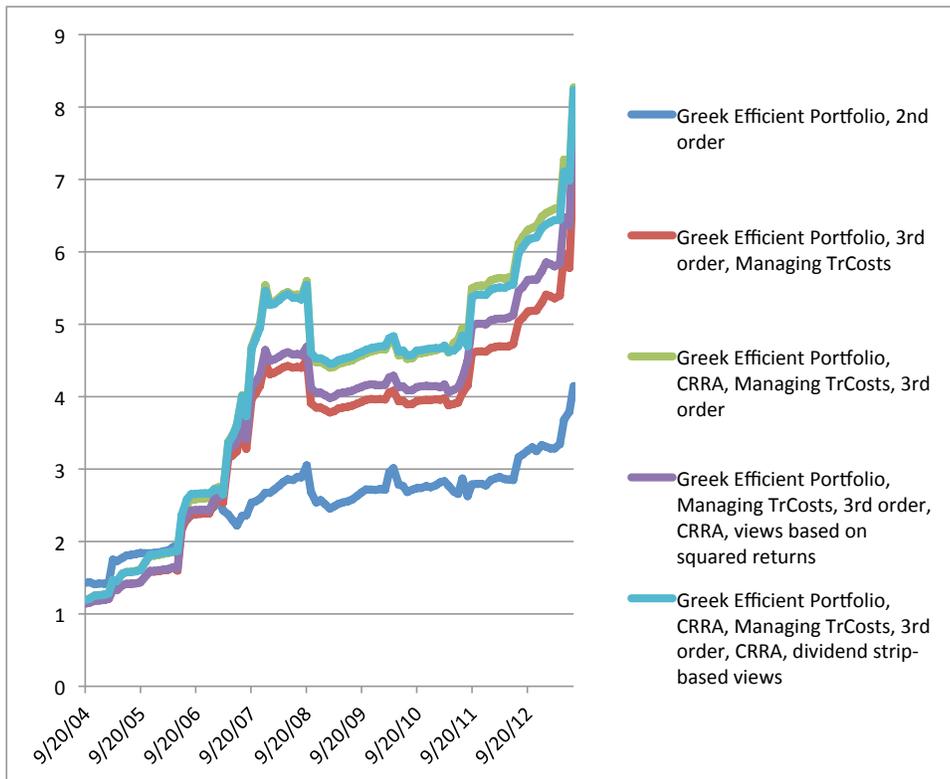


Figure 15: Cumulative returns on the Greek MV-efficient portfolio versus the cumulative returns on the corresponding portfolios that use 3rd order Greeks for risk and transactions costs management. Returns are normalized to have an ex-post annualized volatility of 20%.