SIAM J. FINANCIAL MATH.
Vol. 5, pp. 316–359

Dynamic Portfolio Execution and Information Relaxations*

Martin Haugh† and Chun Wang†

Abstract. We consider a portfolio execution problem where a possibly risk-averse agent needs to trade a fixed number of shares in multiple stocks over a short time horizon. Our price dynamics can capture linear but stochastic temporary and permanent price impacts as well as stochastic volatility. In general it is not possible to solve even numerically for the optimal policy in this model, however, and so we must instead search for good suboptimal policies. Our principal policy is a variant of an open-loop feedback control (OLFC) policy, and we show how the corresponding OLFC value function may be used to construct good primal and dual bounds on the optimal value function. The dual bound is constructed using the recently developed duality methods based on information relaxations. One of the contributions of this paper is the identification of sufficient conditions to guarantee convexity, and hence tractability, of the associated dual problem instances. That said, we do not claim that the only plausible models are those where all dual problem instances are convex. We also show that it is straightforward to include a nonlinear temporary price impact as well as return predictability in our model. We demonstrate numerically that good dual bounds can be computed quickly even when nested Monte Carlo simulations are required to estimate the so-called dual penalties. These results suggest that the dual methodology can be applied in many models where closed-form expressions for the dual penalties cannot be computed.

Key words. portfolio execution, suboptimal control, duality, information relaxations

AMS subject classifications. 91G80, 93E20

DOI. 10.1137/120896761

1. Introduction.

Dynamic portfolio execution is one of the most important practical problems faced by institutional investors and brokerage firms today. This problem requires the purchase or sale of a large number of shares in multiple stocks over a short time horizon. The large number of shares and short time horizon can result in a very significant market impact that can in turn greatly increase the cost of trading. This of course is in contrast to traditional dynamic portfolio optimization problems, where trading is generally assumed to be frictionless. Because of the potentially large trading costs, it is necessary to model market impact explicitly when determining portfolio execution policies.

In this paper we consider the portfolio execution problem of a possibly risk-averse agent who needs to account for temporary and permanent market impact costs, as well as other market features such as stochastic liquidity and return predictability, among others. In general it is not possible to solve for the optimal policy in a model that can capture these features, and so we must instead search for good suboptimal policies. But how good are these suboptimal policies? In this paper we demonstrate the use of duality methods based on information

*Received by the editors October 29, 2012; accepted for publication (in revised form) February 24, 2014; published electronically May 22, 2014.
http://www.siam.org/journals/sifin/5/89676.html

†Department of Industrial Engineering & Operations Research, Columbia University, New York, NY 10027 (mh2078@columbia.edu, cw2519@columbia.edu).

relaxations to evaluate how far from optimality these suboptimal policies can be. We also recognize that even "realistic" models are still only an approximation to the true market dynamics, and we therefore note how these dual methods can also be used to estimate how robust a given policy is to departures from the assumed model.

Our principal model incorporates risk aversion, stochastic return covariances, and temporary and permanent linear price impacts that are also stochastic and time-varying. We propose a variant of an open-loop feedback control (OLFC) policy for solving the resulting portfolio execution problem and use duality methods based on information relaxations to demonstrate how well this policy performs, at least for the parameter settings that we consider. We also show that it is straightforward to include a nonlinear temporary price impact as well as return predictability. Our model therefore allows us to capture a much broader range of market features than can be captured by those models that insist on explicit calculation of the optimal policy. One of the contributions of this paper is the formulation and analysis of the dual problem instances that arise from this model. We discuss the convexity of these dual problem instances and provide some sufficient conditions that guarantee the convexity of all such instances. That said, we also recognize that the economic case for guaranteeing dual convexity of *all* problem instances is not clear, and so we should not necessarily limit ourselves to model conditions that preclude the nonconvexity of *some* dual instances.

In order to apply the aforementioned duality methods it is necessary to use an approximate value function to compute so-called *penalties* which penalize the decision-maker for violating the nonanticipativity constraints. These penalties, however, require us to compute conditional expectations of the approximate value function. Ideally we can compute these expectations analytically, but in many circumstances this may not be possible. This does not limit the applicability of the dual methodology, however, since it is known that being able to compute an unbiased estimate of the conditional expectations is sufficient for computing valid dual bounds. A further contribution of this paper is the use of suitably randomized low-discrepancy sequences (LDS) to estimate these expectations *efficiently*. We show that using randomized LDS to estimate these expectations can still yield tight dual bounds and at only a very modest increase (approx 10%) in computational work.

The literature on controlling execution costs is extensive, and we can barely do justice to it here. Most of the early work, beginning with Bertsimas and Lo [10] and Almgren and Chriss [5], considered single-stock execution problems which they formulated as dynamic control problems. Other related work in this spirit includes, for example, Almgren [3], Huberman and Stanzl [28], and Gatheral and Schied [23]. Generally these papers consider the *macro* or *scheduling* component of the execution problem, that is, the problem of deciding how to slice the order, when the execution algorithm should trade, and in what size. These models might, for example, take a time horizon of 1 day and break this period up into 5-minute time periods. Assuming a trading day of 6.5 hours duration, this yields a finite-horizon control problem with 78 time periods.

More recently, the growth of electronic exchange markets has led to the modeling of limit order book dynamics. These models, beginning with Obizhaeva and Wang [38], and extended by Alfonsi, Fruth, and Schied [2] and Predoiu, Shaikhet, and Shreve [39], model the market impact and the *decay* of market impact at the limit order book level. Often formulated as continuous-time problems, the resulting stochastic control problems generally yield deterministic execution policies due to the assumption of relatively simple price dynamics. These papers can also be viewed as considering the scheduling component of the execution policy, although they are inspired by microfoundations in that they model the transient behavior of the limit order book. They assume that trades are executed as market orders whose price impact depends on the shape of the order book at the time of execution. In contrast, Cont and Kukanov [17] and Huitema [30] consider the use of both limit and market orders. Other papers related to execution in limit order books include, for example, Cont, Stoikov, and Talreja [18] and Gatheral [22], while Tóth et al. [45] focus on market impact rather than execution.

More recently, so-called *dark pools* have grown in importance as an alternative form of trading. There is no order book in dark pool venues, and buyers/sellers are matched electronically without revealing any information. Trades in dark pools generally have a smaller price impact, but order execution is uncertain. Recent works on optimal execution in these venues include, for example, Kratz and Schoneborn [32] and Buti, Rindi, and Werner [14], both of which consider the use of a classical exchange and dark pools for optimal execution.

While most research has focused on the single-stock execution problem, there has also been some work on the *portfolio* execution problem. Early work in this direction, which again focuses on the macrocomponent of the problem, includes Bertsimas, Hummel, and Lo [9] and Almgren and Chriss [4]. More recent work includes Huberman and Stanzl [29], Schied, Schoneborn, and Tehranci [42], and Lim and Wimonkittiwat [35]. Tsoukalas, Wang, and Giesecke [46] consider the portfolio execution problem at the limit order book level and formulate their control problem as an equivalent static convex optimization problem. In another recent paper, Moallemi and Sağlam [37] consider general dynamic portfolio optimization problems, including portfolio execution problems, and propose optimizing over linear rebalancing rules in order to find good suboptimal policies. Their optimization problem is therefore static in nature and amenable to standard convex optimization algorithms. Their paper is quite similar in spirit to our paper in that they also focus on developing good suboptimal policies and also use duality based on information relaxations to demonstrate the effectiveness of these policies.

The use of dual methods based on information relaxations to construct good bounds for dynamic control problems began independently with Haugh and Kogan [26] and Rogers [40] in the context of optimal stopping problems and the pricing of American options. See also Andersen and Broadie [6] and Jamshidian [31]. These techniques were then extended to multiple optimal stopping problems by Meinshausen and Hambly [36] and Schoenmakers [43]. Bender [7] and Bender, Schoenmakers, and Zhang [8] provide further extensions of this work. A significant development came with Brown, Smith, and Sun [13] and Rogers [41], who independently extended these duality techniques to more general stochastic dynamic programs. This has now become a very significant research area. More recent applications and developments can be found in Brown and Smith [12], Desai, Farias, and Moallemi [21], Lai, Margot, and Secomandi [33], Lai et al. [34], Moallemi and Sağlam [37], and Haugh and Lim [27]. The latter paper makes the point that some of the ideas behind these dual techniques are actually not so recent and date back to earlier work, including Davis and Karatzas [19] in the case of American options and Davis and Zervos [20] for linear-quadratic control. Chandramouli and Haugh [15] also showed that the earlier duality results for multiple optimal stopping problems could easily be derived using the more general framework of Brown, Smith, and Sun [13].

The remainder of this paper is organized as follows. We define the basic portfolio execution problem and describe our model in section 2. We also describe there the suboptimal policies that we will consider, focusing in particular on a variant of an OLFC policy. In section 3 we review the duality theory of Brown, Smith, and Sun [13] and show how it can be applied in the context of our portfolio execution problem. We present numerical results in section 4 where we consider a portfolio of 50 stocks and 78 time periods, representing, as stated earlier, a trading frequency of once every 5 minutes. In section 5 we consider how the dual methodology can be extended to other portfolio execution problems, including the problem where there is also a nonlinear temporary price impact. We also include numerical results comparing the performance of the dual bound when the penalties are computed analytically with the performance when the dual penalties are computed via an unbiased Monte Carlo simulation. We conclude in section 6. Appendix A contains the various calculations and derivations required for implementing the OLFC and other policies. Appendix B contains a review of the duality theory based on information relaxations, while Appendix C analyzes the particular dual problem arising from our portfolio execution application. Appendix E contains additional calibration details for the numerical results of section 4.

In Appendix D we consider the model of Bertsimas, Hummel, and Lo [9] (BHL), where the problem of an agent needing to purchase shares in multiple securities assuming a linear temporary price impact and return predictability was considered. The presence of no-sales constraints, which in general are binding due to return predictability, implies that it is hard to solve for optimal policy even numerically. They proposed instead an OLFC policy and conjectured that such a policy should be close to optimal. Using their calibrated model parameters, we use the duality techniques to confirm that the OLFC policy is indeed very close to optimal. While this appendix is stand-alone, we include it for two reasons: (i) it provides another set of numerical experiments (in a portfolio execution context) demonstrating the use of dual methods to confirm the conjecture that a suboptimal policy is close to optimal, and (ii) it provides a simple demonstration of how duality can be used to determine *in advance* whether or not a particular feature, in this case cross-price impacts, is worth accounting for in a portfolio execution policy. We believe the duality methodology is particularly suited for answering such *need-to-model* questions.

2. The portfolio execution problem and model description. We now describe the basic portfolio execution problem of a possibly risk-averse agent who we assume needs to *purchase* a fixed number of shares in a fixed number of assets. We note that we could just as easily handle the problem where the agent needs to *sell* a portfolio of securities as well as the case where the agents needs to purchase one subset of securities and sell another subset.

We assume time is discrete and runs from $t = 0$ to $t = T$ for a total of $T + 1$ time periods. There are n different assets that are traded in the market, and the agent needs to purchase a fixed number of shares in each of the n assets between $t = 0$ and $t = T$. At time t the agent observes the nonimpact asset price vector $\tilde{\mathbf{p}}_t = [\tilde{p}_t^{(1)} \dots \tilde{p}_t^{(n)}]'$, and then determines the decision vector, $\mathbf{s}_t = [s_t^{(1)} \dots s_t^{(n)}]'$, where $s_t^{(i)}$ is the number of shares of the ith asset purchased at time t. We will also use $\mathbf{s}_{u:v}$ to denote the $n(v - u + 1) \times 1$ vector of decision variables, $[\mathbf{s}'_u \dots \mathbf{s}'_v]'$. However, we will simply write $\mathbf{s}$ for $\mathbf{s}_{0:T}$. We use $\mathbf{p}_t = [p_t^{(1)} \dots p_t^{(n)}]'$ to denote the time t vector of transaction prices and note that in general, $\mathbf{p}_t \neq \tilde{\mathbf{p}}_t$ due to the market impact of trading.

We let $\mathbf{w}_t = [w_t^{(1)} \dots w_t^{(n)}]'$, where $w_t^{(i)}$ denotes the remaining number of shares in asset i that must be purchased in periods $t, \dots, T$. We define the execution cost as the difference between the actual cost, $\sum_{t=0}^{T} \mathbf{p}_t' \mathbf{s}_t$, and the benchmark cost, $\tilde{\mathbf{p}}_0' \mathbf{w}_0$, which would prevail if the agent could purchase everything at time $t = 0$ without any market impact. The execution cost is therefore given by

$$\sum_{t=0}^{T} \mathbf{p}_t' \mathbf{s}_t - \tilde{\mathbf{p}}_0' \mathbf{w}_0. \tag{1}$$

In order to model risk aversion we assume that the agent has an exponential utility function, with parameter γ. We assume that $\gamma > 0$ to reflect the fact that we are defining utility over *costs* rather than *wealth*, as is usually the case in portfolio optimization problems. The portfolio execution problem is then stated as

$$\min_{\mathbf{s}_t \in \mathcal{F}_t, t=0,\dots,T} \quad \mathbb{E}_0 \left[\exp\left(\gamma \sum_{t=0}^{T} \mathbf{p}_t' \mathbf{s}_t \right) \right] \tag{2}$$

$$\text{subject to} \quad \text{price and state variable dynamics,}$$

$$\sum_{t=0}^{T} \mathbf{s}_t = \mathbf{w}_0, \tag{3}$$

$$\text{and any other constraints,} \tag{4}$$

where we use $\{\mathcal{F}_t\}_{t=0,\dots,T}$ to denote the filtration generated by the price vectors as well as any other state variables in the model. Note also that the objective function (2) does not include $\tilde{\mathbf{p}}_0' \mathbf{w}_0$ because it is a constant and therefore can be factored out. Several comments are in order:

(i) Risk neutrality can easily be modeled by taking the limit in (2) as γ goes to 0.

(ii) Recall the *CARA* property of exponential utility which, in the context of a standard dynamic portfolio optimization problem, implies that the optimal dollar value invested in risky assets does *not depend* on the current level of wealth. This property is typically viewed as a serious weakness of exponential utility. We see no problem with this assumption in the context of portfolio execution problems, however, because such problems often have a time horizon of just a few hours or at most just a few days.

(iii) Notwithstanding the previous point, we use only exponential utility as a mechanism for trading off execution cost with execution risk. In particular, other utility functions could also be used, although the solution approach of Appendices A.3 and A.4 would no longer be applicable.

2.1. Basic model description. Our main model allows for stochastic and time-varying price impacts as well as stochastic variance-covariance return dynamics. These are important features in practice but are generally ignored in the academic literature. While we suspect that superior or more sophisticated proprietary models may be used by some industry participants, our model is sufficiently rich to demonstrate the broad applicability of the dual methodology as outlined in section 3.

We assume the nonimpact price, $\tilde{\mathbf{p}}_t$, follows a random walk where the dollar return, $\mathbf{r}_{t+1}$, between times t and $t+1$, is normally distributed with mean $\mathbf{0}$ and conditional covariance matrix $\mathbf{\Sigma}_t$. A trading volume of $\mathbf{s}_t$ incurs a permanent price impact of $\mathbf{A}_t\mathbf{s}_t$ and a temporary price impact of $\mathbf{B}_t\mathbf{s}_t$. This results in dynamics of the form

$$\mathbf{p}_t = \tilde{\mathbf{p}}_t + \mathbf{A}_t\mathbf{s}_t + \mathbf{B}_t\mathbf{s}_t, \tag{5}$$

$$\tilde{\mathbf{p}}_{t+1} = \tilde{\mathbf{p}}_t + \mathbf{r}_{t+1} + \mathbf{A}_t\mathbf{s}_t, \tag{6}$$

$$\mathbf{r}_{t+1} = \mathbf{\Sigma}_t^{1/2}\boldsymbol{\epsilon}_{t+1}, \tag{7}$$

$$\mathbf{w}_{t+1} = \mathbf{w}_t - \mathbf{s}_t \tag{8}$$

for $t = 0, \ldots, T$ and with the understanding that $\mathbf{w}_{T+1} = \mathbf{0}$. The $\boldsymbol{\epsilon}_t$'s in (7) are independent and identically distributed (IID) standard normal random vectors.

In general we allow $\mathbf{\Sigma}_t$, $\mathbf{A}_t$, and $\mathbf{B}_t$ to be stochastic. This will allow us to model the tendency for markets to be more liquid around the open and close and less liquid in between. The assumption of a linear permanent price impact is commonly made in the academic literature, where it has been justified by no-arbitrage considerations; see Huberman and Stanzl [28] and Gatheral [22]. Nonetheless we note that permanent price impact is generally assumed to be better approximated by a square-root function in practice. We note that Guéant [25] recently provided a justification for this without introducing arbitrage into the model. Later, in section 5 we will show that a nonlinear temporary price impact can also be included in our model.

When the agent chooses $\mathbf{s}_t$ at time t, we assume that he knows $\mathbf{A}_t$, $\mathbf{B}_t$, $\mathbf{\Sigma}_t$, and $\tilde{\mathbf{p}}_t$. We note that it would also be straightforward to model the possibly more realistic situation where the agent is not assumed to know $\tilde{\mathbf{p}}_t$ when choosing $\mathbf{s}_t$. We will also assume that we can compute $\mathbb{E}_t[\mathbf{A}_j]$ and $\mathbb{E}_t[\mathbf{B}_j]$ for all $j \geq t$. In section 4 we will describe the specific dynamics for $\mathbf{\Sigma}_t$, $\mathbf{A}_t$, and $\mathbf{B}_t$ that we used in our numerical experiments.

In practical applications the agent may also need to include various types of portfolio constraints including no-sales constraints

$$\mathbf{s}_t \geq \mathbf{0} \text{ for } t = 0, \ldots, T \tag{9}$$

as well as sector-balance constraints. These constraints are linear and therefore, as we shall see later, do not impact the convexity of the primal or dual problems that we consider. We do not consider portfolio balance constraints in this paper but note that it is generally straightforward to include them when attempting to construct good suboptimal policies. We also note that the assumption of a risk-averse utility function implicitly incorporates some form of portfolio balance constraints.

2.2. Suboptimal execution policies. In general it is not possible to find the optimal policy for this portfolio execution problem, and so instead we seek good suboptimal policies. For the particular model of section 2.1 we will consider three different policies.

The simple policy. Here the agent buys the same quantity of shares in each of the $T+1$ time periods so that $\mathbf{s}_t = \mathbf{w}_0/(T+1)$.

The risk-neutral policy. In this case the agent ignores the *risk* of the execution costs and uses the trading policy which minimizes the expected execution cost. If the dynamics of $\mathbf{A}_t$

and $\mathbf{B}_t$ are sufficiently tractable, then this trading policy can be determined via dynamic programming when there are no portfolio constraints. The details are in Appendix A.1. In Appendix A.2 we also show that the simple and risk-neutral policies coincide if $\mathbf{A}_t$ and $\mathbf{B}_t$ are martingales and if $\mathbf{A}_t$ and $\mathbf{B}_t$ are symmetric. We note that this condition is of course very unlikely to hold in practice.

The OLFC policy. At each time t the agent assumes that $\boldsymbol{\Sigma}_t$ will remain constant thereafter, in which case he assumes the return series, $\{\mathbf{r}_j\}_{j=t+1}^T$, is IID MVN $(\mathbf{0}, \boldsymbol{\Sigma}_t)$. The agent also assumes the price-impact matrices evolve deterministically and replaces $\mathbf{A}_j$ and $\mathbf{B}_j$ with their time-t conditional expectations, $\mathbb{E}_t[\mathbf{A}_j]$ and $\mathbb{E}_t[\mathbf{B}_j]$, respectively, for $j \geq t$. The OLFC policy therefore solves the following problem at each time $t = 0, \ldots, T$:

$$
\begin{aligned}
(10) \qquad & \min_{\mathbf{s}_{t:T} \in \mathcal{F}_t} \quad \mathbb{E}_t\left[\exp\left(\gamma \sum_{j=t}^{T} \mathbf{p}_j' \mathbf{s}_j\right)\right] \\
& \text{subject to} \quad \mathbf{p}_j = \tilde{\mathbf{p}}_j + \mathbb{E}_t[\mathbf{A}_j]\mathbf{s}_j + \mathbb{E}_t[\mathbf{B}_j]\mathbf{s}_j \quad \text{for } j = t, \ldots, T, \\
& \qquad \tilde{\mathbf{p}}_{j+1} = \tilde{\mathbf{p}}_j + \mathbf{r}_{j+1} + \mathbb{E}_t[\mathbf{A}_j]\mathbf{s}_j \quad \text{for } j = t, \ldots, T, \\
& \qquad \mathbf{r}_j = \boldsymbol{\Sigma}_t^{1/2} \boldsymbol{\epsilon}_j \quad \text{for } j = t+1, \ldots, T, \\
& \qquad \sum_{j=t}^{T} \mathbf{s}_j = \mathbf{w}_t, \\
& \qquad \mathbf{s}_j \geq \mathbf{0} \quad \text{for } j = t, \ldots, T.
\end{aligned}
$$

Let V_t^{ol} denote the optimal value of (10), and let $\mathbf{s}_{t:T}^{ol,t} := [\mathbf{s}_t^{ol,t\,\prime} \ \ldots \ \mathbf{s}_T^{ol,t\,\prime}]'$ denote the corresponding optimal solution. The OLFC policy implements $\mathbf{s}_t^{ol,t}$ at time t and ignores $\mathbf{s}_{t+1}^{ol,t}, \ldots, \mathbf{s}_T^{ol,t}$. Further details on solving this problem are in Appendix A.3. We also note that if we remove the constraint $\mathbf{s}_j \geq \mathbf{0}$, then an analytic solution to the OLFC policy can be found very quickly via dynamic programming. This analytic solution will form the basis for constructing dual penalties as discussed in section 3. Details on this solution can be found in Appendix A.4.

A risk-neutral OLFC policy. As mentioned above, depending on the dynamics of $\mathbf{A}_t$ and $\mathbf{B}_t$, the risk-neutral policy may not be computable. In this case we can compute the risk-neutral OLFC policy by taking the limit of (10) as γ goes to 0. The details are in Appendix A.3.

3. Evaluating suboptimal policies. Given a feasible suboptimal policy to the portfolio execution problem in (2), we can construct an unbiased upper bound, V_{ub}, to the optimal value function, V^*, by simulating multiple paths of the policy and taking the sample average of the realized utility. We can also use dual methods to estimate a lower bound, V_{lb}. Appendix B contains a review of these dual methods that were originally developed independently by Brown, Smith, and Sun [13] and Rogers [41]. In this section we state the dual problem formulation of (2) and briefly discuss the tractability of this dual problem and its implications for modeling various features of the portfolio execution problem.

Let $\tilde{V}_t(\mathbf{s}_{0:t-1})$ be some approximation to the time t optimal value function of the portfolio execution problem. While we typically consider a value function to be a function of the model's state variables, in our duality context it is much more convenient to explicitly recognize the dependence of $\tilde{V}_t$ only on $\mathbf{s}_{0:t-1}$. As reviewed in Appendix B, weak duality implies that

$$V_{lb} \; := \; \mathbb{E}_0\left[\min_{\mathbf{s}\in\mathbb{S}} \left\{ \exp\left(\gamma \sum_{t=0}^{T} \mathbf{p}_t'\mathbf{s}_t\right) + \sum_{t=0}^{T-1} \left(\mathbb{E}_t[\tilde{V}_{t+1}(\mathbf{s}_{0:t})] - \tilde{V}_{t+1}(\mathbf{s}_{0:t})\right) \right\}\right] \tag{11}$$

yields a lower bound on V^*, where $\mathbb{S}$ is the decision space defined by constraints (3) and (4). Moreover, strong duality states that if we take $\tilde{V} = V^*$, then $V_{lb} = V^*$. This suggests that the closer $\tilde{V}_t(\mathbf{s}_{0:t-1})$ is to the optimal value function the tighter the dual bound will be. Note that it is the *shape* of $\tilde{V}$ that is important rather than the absolute level of $\tilde{V}$ because adding any constant to $\tilde{V}$ will have no impact on the dual bound since it will cancel out on the right-hand side of (11). In general for a given $\tilde{V}_t$ we cannot compute V_{lb} in closed form, but in principle, an unbiased estimate of it can be computed via Monte Carlo: we simply simulate M paths of the exogenously specified noise processes, and on each path we solve the minimization problem inside the expectation in (11). If $V_{lb}^{(i)}$ is the optimal solution of the problem on the ith path, then $\sum_{i=1}^{M} V_{lb}^{(i)}/M$ is an unbiased estimate of V_{lb}.

Two key issues arise in applying the duality methodology. First, while $V_{lb}^{(i)}$ can be obtained as the solution of a deterministic dynamic program (DP), solving this DP may, like the primal problem, also be difficult. Instead we prefer to solve it as a static optimization problem, but to do this we would generally prefer this optimization problem to be convex. Even if the term $\exp\left(\gamma \sum_{t=0}^{T} \mathbf{p}_t'\mathbf{s}_t\right)$ is convex in $\mathbf{s}$, there is no guarantee that the objective function will be convex because of the penalty term $\sum_{t=0}^{T-1}(\mathbb{E}_t[\tilde{V}_{t+1}(\mathbf{s}_{0:t})] - \tilde{V}_{t+1}(\mathbf{s}_{0:t}))$. We can overcome this problem, however, by using an approximation $\tilde{V}_{t+1}$ that is linear in $\mathbf{s}_t$. It is important to note that using such an approximation still yields a valid dual bound. This approach was introduced by Brown and Smith [12], and further details are provided in section 3.1.

The second issue that arises is the possibility of not being able to compute an analytic expression for $\mathbb{E}_t[\tilde{V}_{t+1}(\mathbf{s}_{0:t})]$. Note that these conditional expectations are required to evaluate the objective function inside the expectation in (11). As discussed in section 3.1, with our choice of $\tilde{V}_t$ in this paper, we were able to compute these conditional expectations analytically. We will also see in section 5, however, that even if we cannot compute these expectations in closed form, then we can instead use unbiased estimates of them and still obtain valid dual bounds. In a series of numerical experiments we will see that the resulting dual bounds remain very tight and that they are not much more expensive to compute than the bounds we obtain when the conditional expectations are available in closed form. We conclude then that the second issue can also often be overcome.

3.1. The dual problem formulation. Under the model assumptions of section 2.1, we see from (11) that a dual problem instance takes the form

$$\begin{aligned} \min_{\mathbf{s}\in\mathbb{S}} \quad & \exp\left(\gamma \sum_{t=0}^{T} \mathbf{p}_t'\mathbf{s}_t\right) + \sum_{t=0}^{T-1} \left(\mathbb{E}_t[\tilde{V}_{t+1}(\mathbf{s}_{0:t})] - \tilde{V}_{t+1}(\mathbf{s}_{0:t})\right) \\ \text{subject to} \quad & \mathbf{p}_t = \tilde{\mathbf{p}}_t + \mathbf{A}_t\mathbf{s}_t + \mathbf{B}_t\mathbf{s}_t, \\ & \tilde{\mathbf{p}}_{t+1} = \tilde{\mathbf{p}}_t + \mathbf{r}_{t+1} + \mathbf{A}_t\mathbf{s}_t, \end{aligned} \tag{12}$$

where the $\mathbf{r}_t$'s, $\mathbf{A}_t$'s, and $\mathbf{B}_t$'s have been simulated according to the true model dynamics and are all known to the decision-maker, and $\mathbb{S}$ denotes the constraint region defined by (3) and (9). We would like to take $\tilde{V}_t$ to be a linearized version of V_t^{ol} since we expect that the OLFC policy will typically provide a good approximation to V_t^*. In this case, however, it is difficult to compute $\mathbb{E}_t[\tilde{V}_{t+1}(\mathbf{s}_{0:t})]$. While we can overcome this problem using nested Monte Carlos as discussed in section 5, we prefer instead to use a *modified* version of V_t^{ol} which we denote by V_t^{mol}. The precise definition of V_t^{mol} can be found in Appendix C.2, but we note in particular that we can compute an analytic expression for $\mathbb{E}_t[V_{t+1}^{mol}(\mathbf{s}_{0:t})]$.

In order to define the particular choice of $\tilde{V}_t$ that we use in (12) we first define

(13)
$$\hat{V}_{t+1}(\mathbf{s}_{0:t}) := \exp\left(\gamma \sum_{j=0}^{t} \mathbf{p}_j' \mathbf{s}_j\right) V_{t+1}^{mol}$$

(14)
$$= \exp\left(\gamma \sum_{j=0}^{t} \mathbf{p}_j' \mathbf{s}_j + \frac{\gamma}{2}\mathbf{w}_{t+1}' \tilde{\mathbf{G}}_{t+1}\mathbf{w}_{t+1} + \left(\tilde{\mathbf{p}}_t + \mathbf{A}_t\mathbf{s}_t\right)'\mathbf{w}_{t+1}\right) \exp\left(\gamma \mathbf{r}_{t+1}'\mathbf{w}_{t+1}\right),$$

where we have used (C-40) to substitute for V_{t+1}^{mol} in (14) and then used (6) to obtain (14). We note that the dependence of $\hat{V}_{t+1}(\cdot)$ on $\mathbf{s}_{0:t}$ also appears via $\mathbf{w}_{t+1}$ since $\mathbf{w}_{t+1} = \mathbf{w}_0 - \sum_{j=0}^{t} \mathbf{s}_j$. Note also that we want to include the first term on the right-hand side of (14) since this term is present in the value function (see also the remark at the end of Appendix B) of any policy, but we omitted it from (10) since it had no impact on the OLFC *policy* at time $t+1$.

The difficulty with taking $\tilde{V}_{t+1}(\mathbf{s}_{0:t}) = \hat{V}_{t+1}(\mathbf{s}_{0:t})$ as given by (14) is that the resulting objective function in (12) will not in general be convex, even if the first term in (12) is convex. Instead we linearize $\hat{V}_{t+1}$. Recall that $\mathbf{s}_{0:t} = [\mathbf{s}_0' \ldots \mathbf{s}_t']'$ denotes the $n(t+1)\times 1$ vector of decision variables corresponding to the first $t+1$ time periods, and let $\tilde{\mathbf{s}}_{0:t}$ be a fixed $n(t+1)\times 1$ vector. We then take $\tilde{V}_{t+1}(\cdot)$ to be a first order Taylor expansion of $\hat{V}_{t+1}(\cdot)$ about $\tilde{\mathbf{s}}_{0:t}$. In particular, we take

$$\tilde{V}_{t+1}(\mathbf{s}_{0:t}) \; := \; \hat{V}_{t+1}(\tilde{\mathbf{s}}_{0:t}) + \nabla \hat{V}_{t+1}(\tilde{\mathbf{s}}_{0:t})'(\mathbf{s}_{0:t} - \tilde{\mathbf{s}}_{0:t}) \tag{15}$$

and then use this in (12), noting that the linearity of $\tilde{V}_{t+1}$ in $\mathbf{s}_{0:t}$ will preserve convexity of the objective function in (12) if the exponential term there itself is convex.

The question that now arises is how to choose $\tilde{\mathbf{s}}_{0:t}$. Intuitively we would like $\tilde{\mathbf{s}}_{0:t}$ to be as close as possible to the true optimal trade sequence. But of course we do not know the optimal trade sequence, so instead we will take $\tilde{\mathbf{s}}_{0:t} = [\mathbf{s}_0^{ol,0\,\prime} \ldots \mathbf{s}_t^{ol,t\,\prime}]'$, the trade sequence from the unconstrained OLFC policy. This implies that $\tilde{\mathbf{s}}_{0:t}$ is path-dependent so that each dual problem instance will use a different $\tilde{\mathbf{s}}_{0:t}$ to construct the $\tilde{V}_{t+1}$. This requires no additional work, however, since the $\mathbf{s}_t^{ol,t}$'s will have already been computed when simulating the OLFC policy.

In order to compute $\mathbb{E}_t[\tilde{V}_{t+1}(\mathbf{s}_{0:t})] = \mathbb{E}_t[\hat{V}_{t+1}(\tilde{\mathbf{s}}_{0:t})] + \mathbb{E}_t[\nabla \hat{V}_{t+1}(\tilde{\mathbf{s}}_{0:t})'](\mathbf{s}_{0:t} - \tilde{\mathbf{s}}_{0:t})$ we need $\mathbb{E}_t[\hat{V}_{t+1}(\tilde{\mathbf{s}}_{0:t})]$ and $\mathbb{E}_t[\nabla \hat{V}_{t+1}(\tilde{\mathbf{s}}_{0:t})']$ to solve the dual instance in (12). Using (14) we obtain

(16)
$$\mathbb{E}_t[\hat{V}_{t+1}(\tilde{\mathbf{s}}_{0:t})] = \exp\left(\gamma \sum_{j=0}^{t} \mathbf{p}_j'\tilde{\mathbf{s}}_j + \frac{\gamma}{2}\tilde{\mathbf{w}}_{t+1}'\tilde{\mathbf{G}}_{t+1}\tilde{\mathbf{w}}_{t+1} + \left(\tilde{\mathbf{p}}_t + \mathbf{A}_t\tilde{\mathbf{s}}_t\right)'\tilde{\mathbf{w}}_{t+1}\right)\mathbb{E}_t\left[\exp\left(\gamma \mathbf{r}_{t+1}'\tilde{\mathbf{w}}_{t+1}\right)\right],$$

where $\tilde{\mathbf{w}}_{t+1} = \mathbf{w}_0 - \sum_{j=0}^{t} \tilde{\mathbf{s}}_j$. The first exponential term on the right-hand side of (16) is $\mathcal{F}_t$-measurable (since the OLFC policy is $\mathcal{F}_t$-adapted and $\tilde{\mathbf{G}}_{t+1}$ is $\mathcal{F}_t$-measurable) and conditional on $\mathcal{F}_t$, $\mathbf{r}_{t+1}$ is normally distributed with mean $\mathbf{0}$ and covariance matrix $\mathbf{\Sigma}_t$. We can therefore compute $\mathbb{E}_t[\hat{V}_{t+1}(\tilde{\mathbf{s}}_{0:t})]$ analytically and, while the calculations are somewhat tedious, also do the same for $\mathbb{E}_t[\nabla \hat{V}_{t+1}(\tilde{\mathbf{s}}_{0:t})']$. This means in particular that we can compute an analytic expression for the dual penalty term, $\mathbb{E}_t[\tilde{V}_{t+1}(\mathbf{s}_{0:t})] - -\tilde{V}_{t+1}(\mathbf{s}_{0:t})$.

3.2. Dual convexity? The exponential term in (12) is studied in Appendix C, where we establish sufficient conditions on the $\mathbf{A}_t$'s and $\mathbf{B}_t$'s that guarantee its convexity for all dual problem instances. For example, we show that if $\mathbf{A}_t = \mathbf{A}$, a constant matrix, and that $\mathbf{A}+\mathbf{A}'$ is positive definite and $\mathbf{B}_t + \mathbf{B}_t'$ is positive semidefinite for all t, then every dual problem instance will be convex. These conditions guarantee the positive definiteness of $\mathbf{A}_t + \mathbf{A}_t' + \mathbf{B}_t + \mathbf{B}_t'$, which seems reasonable from an economic standpoint.

If the $\mathbf{A}_t$'s are stochastic, however, then economic considerations may still be employed to impose some structure on their dynamics. For example, it seems reasonable to assume that there should always exists a solution to the OLFC problem. Otherwise the OLFC decision-maker would believe (possibly incorrectly since admittedly he makes simplifying assumptions on the model dynamics) that arbitrage opportunities exist. This seems very unlikely in practice, and so we could therefore insist on dynamics for the $\mathbf{A}_t$'s that guarantee the positive definiteness, of the $\mathbf{Q}_{OLFC,t}$'s as defined in (A-19) of Appendix A.3.

More generally, however, it is not clear that we can use economic considerations to justify the convexity of each dual problem instance when the $\mathbf{A}_t$'s are stochastic. After all, consider the example where there is no temporary price impact and the agent needs to purchase shares in just one security over the course of a day. If the $\mathbf{A}_t$'s are stochastic, it is possible on some given realization of the model uncertainty over the course of a day that the price impact is very large in the early periods and approximately zero in the later periods. In the corresponding dual problem instance (with zero dual penalties, say) this could result in arbitrage profits since the agent could drive up the stock price over the course of the day by purchasing many shares when the price impact is large and then selling them all at the end of the day when the price impact is small. The presence of such arbitrage profits would imply the nonconvexity of the corresponding dual problem instance. The primal problem could still be well-posed, however, since the decision-maker does not get to see the realization of the day's uncertainty at time $t = 0$ and therefore would not be in a position to profit from the aforementioned strategy.

In general, then, it seems possible for a plausible model to result in some dual problem instances that are nonconvex. In that case we would need to bound the optimal objective function of these particular dual problems in order to still obtain a valid dual bound. This should be possible in the presence of portfolio constraints such as no-sales constraints since in that case every dual problem would have a compact decision space and therefore have a bounded optimal solution. Nonetheless, this would require more work since we would need to check for convexity of each dual problem and then solve to optimality those problems that are convex and bound the optimal solution to the nonconvex problems. In the numerical results of section 4 we have limited ourselves to a model specification where convexity of all problem instances is guaranteed. But we do acknowledge that this appears to be a strong assumption in models with stochastic price impacts.

3.3. Using the dual formulation to investigate more complex models. We mention at this point just how useful the dual methodology can be. Suppose, for example, that we begin with a simplified model where the nonimpact price, $\tilde{\mathbf{p}}_t$, follows a simple random walk. Suppose also that we have established that all of the corresponding dual problems are convex and therefore very tractable. Consider now adding the following features to this simplified model: (i) we replace the random walk dynamics with stochastic variance-covariance dynamics, and (ii) we introduce new state variables to model return predictability. Then the dual problem instances of this more complex model will remain convex as long as the stochastic volatility and state variable dynamics are not influenced by the decision variables. This follows immediately once we recognize that these new model features affect only the *distribution* of dual problem instances; in particular they do not change the convexity properties of the dual problem instances. Another way to see this is to consider the dual problem instance in (12). We see that $\mathbf{r}_{t+1}$ appears in the constraints for this problem but Σ_t does not. Similarly, if we had a state vector, Z_t, that induced predictability in the returns, then Z_t would also not appear in the dual problem formulation, and so it would not influence the convexity of the dual problems. It would, however, influence the distribution of the dual problems and therefore the actual value of the dual bound.

Note also that in this more complex model we are free to use the original $\tilde{V}_t$ from the simplified model in order to construct dual penalties and a dual bound. Similarly we are also free to use a suboptimal policy for the simplified model to construct a primal bound for the more complex model. If we then find that the resulting duality gap is small, then we know that the new features (as currently calibrated) have little influence and can be safely ignored. We would argue then that the dual methodology can also be employed to determine whether or not certain features that are known to exist in the market are sufficiently important as to require explicit modeling. We do precisely this in section 5.2, where we investigate whether or not it is necessary to include a stochastic component in our temporary price-impact model.

4. Numerical results. In this section we analyze the performance of the primal and dual bounds on a stylized example where the agent needs to purchase 100,000 shares in each of $n = 50$ securities over $T + 1 = 78$ time periods. The 78 time periods are intended to reflect the fact that there are 78 5-minute periods in a trading day of 6.5 hours. We note that each dual problem instance has a total of $50 \times 78 = 3,900$ decision variables and that computing the OLFC policy at each time t involves $(T + 1 - t) \times 50$ decision variables. Of course in practice, portfolio execution problems can involve several hundred assets and computing good suboptimal policies in (almost) real time and evaluating these policies via duality can itself be a significant challenge. If, for example, a portfolio execution problem has 500 assets, then each dual problem instance will have $500 \times 78 = 39,000$ decision variables and solving a large number of these dual problem instances would therefore be computationally demanding. We note, however, that there is no need to compute dual bounds in real time since they are not required to implement a given execution policy. There is therefore no problem with computing dual bounds off-line.

Returning to our example, we take our 50 stocks to be the top 50 stocks by market capitalization in the S&P 500 as of October 12, 2011. We take the initial price vector, $\tilde{\mathbf{p}}_0$, to be the initial prices of these 50 stocks on that date. We consider 10 different values of the risk aversion parameter: $\gamma = k \times 10^{-7}$ for $k = 1, \ldots, 10$. For each value of γ, we simulate $5,000$ sample paths and implement the various suboptimal policies on each sample path. We use only 100 sample paths to estimate the dual bounds since this smaller number was found to estimate the dual bounds with an accuracy comparable to that of the primal bounds. We also used control variates as a variance reduction technique, and the specific details can be found in Appendix A.5. We note that obtaining the primal bound required solving $5,000 \times 78$ optimization problems, whereas obtaining the dual bound required solving just 100 optimization problems. The computational bottleneck was therefore in estimating the primal bound.

4.1. The price-impact dynamics. We assume the price-impact coefficients of a stock are inversely proportional to its average daily trading volume. In particular, we take the permanent price-impact coefficient matrix $\mathbf{A}_t = \mathbf{A}$ to be a constant diagonal matrix so that there is no cross-price impact. The diagonal elements of $\mathbf{A}$ are obtained by assuming that the purchase of 10% of the average daily volume of a stock will incur a permanent price impact of 10 basis points of the stock's initial price. In the case of *Apple Inc.*, for example, this implies a permanent price impact coefficient of 1.7830×10^{-7}. With an initial price of \$407.33, we therefore see that if all $100,000$ shares are purchased immediately, then the permanent price impact will be $1.7830 \times 10^{-7} \times 10^5 = \$0.01783$ per share.

We assume that the temporary price-impact matrix $\mathbf{B}_t$ is driven by a single factor, x_t, according to

$$\mathbf{B}_t = \max(x_t, 0)\mathbf{B}, \tag{17}$$

$$x_t = x_{d,t} + x_{s,t}, \tag{18}$$

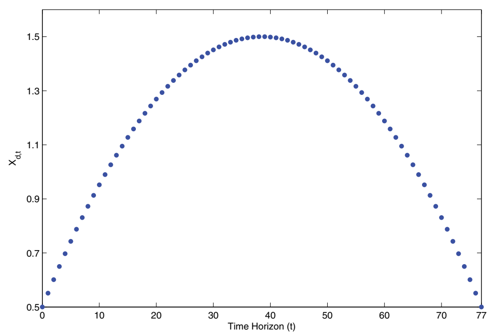
$$x_{s,t} = \rho x_{s,t-1} + \eta_t, \tag{19}$$

where $\mathbf{B}$ is a positive-definite matrix and the max operator in (17) ensures $\mathbf{B}_t$ is positive semidefinite for all t. The state variable x_t is composed of a deterministic component, $x_{d,t}$, and a stochastic component, $x_{s,t}$. The $x_{d,t}$'s are chosen to reflect the fact that the market impact tends to be smaller nearer the open and the close of the trading day. We have assumed $x_{s,t}$ is an AR(1) process so that η_t's in (19) are IID normal random variables with mean 0 and variance σ_η^2. The AR(1) assumption allows us to capture mean-reverting and clustering effects so that periods of low (high) liquidity tend to be followed by periods of low (high) liquidity. We emphasize here, however, that this is a stylized example and in practice considerably more care would be required to specify the dynamics of $\mathbf{B}_t$.

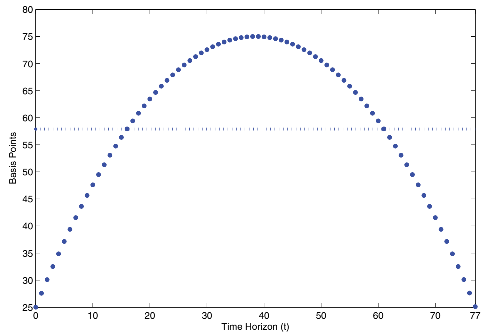
We assume that $\mathbf{B}$ is also a diagonal matrix and represents a temporary price impact of 50 basis points when $x_t = 1$ and 1% of the average daily volume of a stock is purchased immediately. Since $\mathbf{A}$ and $\mathbf{B}$ are both diagonal matrices, we have assumed here that there are no cross-price impacts. We set $x_{s,0} = 0$, $\rho = 0.6$, and $\sigma_\eta^2 = 0.25^2 \times (1 - \rho^2)$. If we ignore the max operator in (17), then these parameter values correspond to a stationary distribution of $\mathbf{B}_t$ where the diagonal elements have a standard deviation equivalent to a temporary price impact of 12.5 basis points.

Figure 1(a) plots $x_{d,t}$ as a function of t, while Figure 1(b) shows the expected temporary price impact of purchasing 1% of the average daily volume shares at time t as a function of t and expressed in basis points. Because this is a stylized example we simply assumed that

$x_{d,t}$ is a simple quadratic function since this is perhaps the simplest way to model the fact that the price impact tends to be lower near the open and the close. Of course we could have just as easily assumed an alternative and possibly more accurate function for $x_{d,t}$. It is also not necessary to restrict ourselves to a scalar process, x_t, and we expect a higher-dimensional process would be more realistic in practice.



(a) Deterministic Component, $x_{d,t}$



(b) Expected Temporary Price Impact

Figure 1. (a) *The deterministic component* $x_{d,t}$ *is assumed to be quadratic in* t. (b) *The expected temporary price impact of purchasing* 1% *of the average daily volume, expressed in basis points, is identical for each of the* n *securities. The dashed line is the time-averaged temporary price impact.*

Returning now to the case of *Apple Inc.*, the purchase of all 100,000 shares at time $t = 0$ incurs a temporary price impact of $0.5 \times 8.9150 \times 10^{-6} \times 10^5 = \$0.4458$ per share. Therefore the total execution cost for this trade is approximately $(0.01783 + 0.4458) \times 100,000 = \$46,358$. The assumption that $\mathbf{A}$ and $\mathbf{B}$ are diagonal matrices is made because of the difficulty of calibrating cross-price impacts, but we do note that in some circumstances it may be worthwhile accounting for them. Table 5 in Appendix E.2 contains the initial prices, average daily volume, and price-impact coefficients that we assumed for the 50 stocks.

Finally we note that dynamics of $\mathbf{A}_t$ and $\mathbf{B}_t$ assumed here satisfy conditions (i) and (ii) of Appendix C.1. We are therefore guaranteed that all dual problem instances will be convex.

4.2. Variance-covariance dynamics. We assume that $\mathbf{\Sigma}_t$ follows an O-GARCH model as in Alexander [1] so that

$$\mathbf{\Sigma}_t = \mathbf{F}\mathbf{\Omega}_t\mathbf{F}' + \mathbf{\Upsilon}, \tag{20}$$

where $\mathbf{\Omega}_t$ is a diagonal matrix, $\mathbf{F}$ is a matrix of factor loadings, and $\mathbf{\Upsilon}$ is a diagonal matrix of idiosyncratic variances. The diagonal elements in $\mathbf{\Omega}_t$ are assumed to follow independent GARCH(1,1) processes. Further details on this model and its calibration can be found in Appendix E.1.

4.3. Results. In our numerical results we report both the average execution cost of each policy as well the certainty equivalent (CE) execution cost. Given an average utility, $\hat{u}$, calculated as the average utility across the simulated sample paths, the CE cost is defined as the execution cost (in basis points), c_e, which yields the same utility with certainty if all shares are purchased immediately. That is, c_e solves

$$\hat{u} = \exp(\gamma \tilde{\mathbf{p}}_0' \mathbf{w}_0 (1 + c_e/10,000)). \tag{21}$$

The CE cost is much easier to interpret than the average utility, and so we prefer to report the former together with the average execution cost. We note of course that the performance of any given policy is measured by the CE cost and *not* the average execution cost, which we would naturally expect to increase with the level of risk aversion, γ.

In Figure 2 we have plotted the CE costs for the simple, OLFC, and risk-neutral OLFC policies. We have also plotted the CE cost for the dual bound constructed using the penalties based on the modified OLFC policy as explained in section 3.1. In this case we see that there is no discernible gap between the OLFC primal bound and the dual bound, which implies that the OLFC policy is actually very close to optimal. For low levels of risk aversion the simple and risk-neutral OLFC policies lose only around five basis points with respect to the OLFC policy, but this number increases (as expected) to approximately 60 basis points as γ increases.

We have also plotted the average execution cost as a function of γ in Figure 3(a). We see that the risk-neutral OLFC policy has the lowest average cost, which is as expected since the risk-neutral agent cares only about minimizing execution cost and is indifferent to execution risk. Since the simple policy is deterministic and the risk-neutral OLFC policy does not depend on γ, we see that their expected execution costs do not vary with γ. As γ increases the agent becomes more risk averse and therefore prefers to buy more shares in earlier time periods. This results in a higher price impact which is reflected by the higher average execution cost for the OLFC policy.

In Figure 3(b) we have plotted the mean-standard deviation frontier corresponding to the OLFC policy. As expected, we see that a higher average execution cost is accompanied by a lower standard deviation. We also see that the simple policy and risk-neutral OLFC policy have low average execution costs but high standard deviations.

5. Extensions and other portfolio execution problems. In this section we briefly describe how our model can be extended to include a nonlinear temporary price impact as well as predictable state variables with linear dynamics. We also discuss limit order book models and explain how the dual technology can also handle these problems. We begin, however, by considering the case where dual penalties cannot be calculated explicitly. We will see that we can still quickly compute good dual bounds in that case using Monte Carlo methods. While this is not a "model extension" it clearly broadens the range of penalties that can be used when constructing dual bounds.

5.1. Computing dual bounds when penalties cannot be computed explicitly. We noted in section 3.1 that we needed to compute terms of the form $\mathbb{E}_t[\hat{V}_{t+1}(\tilde{\mathbf{s}}_{0:t})]$ and $\mathbb{E}_t[\nabla \hat{V}_{t+1}(\tilde{\mathbf{s}}_{0:t})']$ in order to solve a given dual problem instance as in (12). Brown, Smith, and Sun [13] showed that if we could instead compute only unbiased (conditional on $\mathcal{F}_T$) estimates of these terms, then we would still obtain a valid, albeit more conservative, dual bound. The downside of this approach is that we require Monte Carlo simulation to estimate these conditional expectations, and so constructing the dual bound therefore requires nested Monte Carlos, which can be very demanding from a computational standpoint.

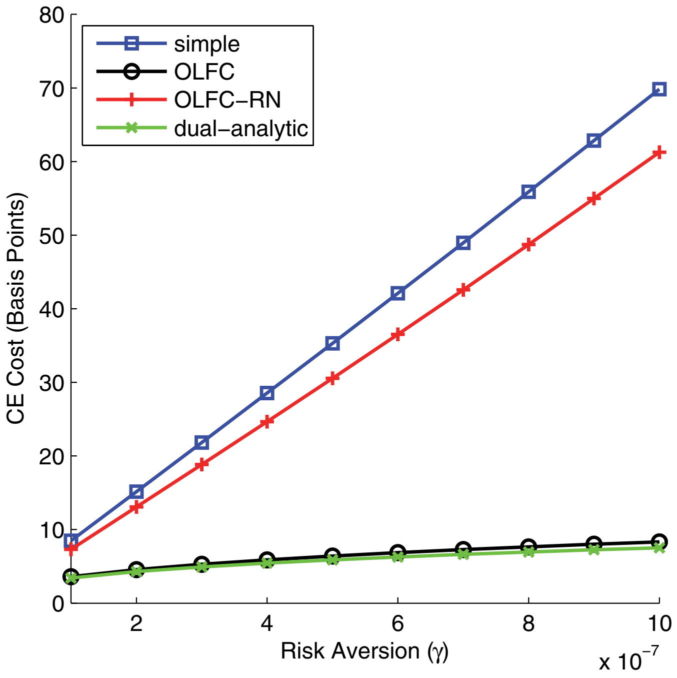
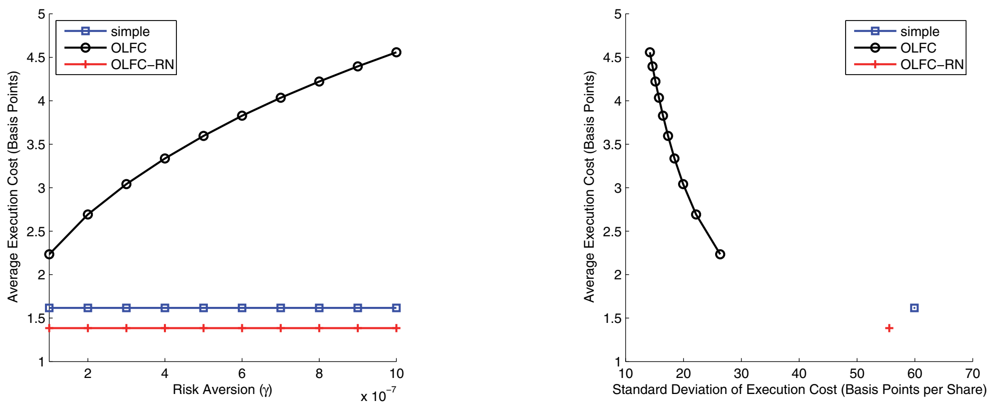


Figure 2. *Certainty equivalent costs.*



(a) Average Execution Cost (b) Mean-Standard Deviation Frontier

Figure 3. *Average cost and mean-standard deviation frontier.*

In this section we will use suitably randomized *low-discrepancy sequences* (LDS)[1] to perform the nested simulations, and we will compare the resulting dual bound to the dual bound

[1]See, for example, Chapter 5 of Glasserman [24] for an introduction to LDS and the randomization technique we describe here.

that we obtained when the conditional expectations are computed in closed form. We will see that we can compute LDS-based dual bounds that are virtually indistinguishable from the original dual bounds and that these new bounds require only 10% additional work.

We note that each of the two conditional expectations may be written in the form $\theta = \mathbb{E}[g(\mathbf{r})]$ for some function $g(\cdot)$ and where $\mathbf{r}$ is an n-dimensional multivariate normal random vector with mean vector $\mathbf{0}$ and covariance matrix $\boldsymbol{\Sigma}$. The precise form of $g(\cdot)$ is clear from (16) in the case of $\mathbb{E}_t[\hat{V}_{t+1}(\tilde{\mathbf{s}}_{0:t})]$ and is also easy to determine in the case of $\mathbb{E}_t[\nabla\hat{V}_{t+1}(\tilde{\mathbf{s}}_{0:t})']$. Using the inverse-transform approach to Monte Carlo we can then easily rewrite θ as

$$\theta = \mathbb{E}[g(\mathbf{U})], \tag{22}$$

where $\mathbf{U}$ is an n-dimensional vector of independent $U(0,1)$ random variables. We could perform the integration in (22) using LDS, which have very nice theoretical properties and which often work particularly well for high-dimensional integrals. The problem with doing this is that the resulting estimate of θ is deterministic and in particular will have some (although presumably small) bias. As a result we could no longer conclude that our estimate of V_{lb} would indeed be an unbiased lower bound. We can overcome this problem by randomizing the LDS in the following manner. In particular, we set

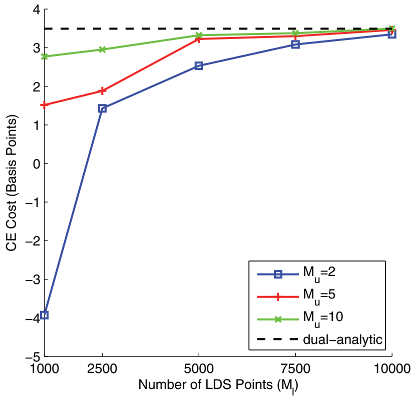
$$\hat{\theta} := \frac{1}{M_u}\sum_{j=1}^{M_u}\left(\frac{1}{M_l}\sum_{i=1}^{M_l} g\big((\mathbf{l}_i + \mathbf{U}_j)\mathrm{mod}\ \mathbf{1}\big)\right), \tag{23}$$

where $(\mathbf{l}_1, \ldots, \mathbf{l}_{M_l})$ is a series of n-dimensional low discrepancy points, and $(\mathbf{U}_1, \ldots, \mathbf{U}_{M_u})$ is a series of independent n-dimensional uniform random vectors. Note also that the mod-operator in (23) applies componentwise and that the inner summation in (23) uses the same uniform vector, $\mathbf{U}_j$, for all M_l samples. In practice, M_u can be very small, e.g., 5 or 10, whereas M_l might be on the order of 10^4 or higher. That $\hat{\theta}$ is an unbiased estimator for θ follows from the fact that if $\mathbf{U}_j$ is an independent n-dimensional uniform random vector, then so too is $(\mathbf{l}_i + \mathbf{U}_j)\mathrm{mod}\ \mathbf{1}$ for all i. If we set $\hat{\theta}_j = \frac{1}{M_l}\sum_{i=1}^{M_l} g\big((\mathbf{l}_i + \mathbf{U}_j)\mathrm{mod}\ \mathbf{1}\big)$, then we see that $\hat{\theta} = \frac{1}{M_u}\sum_{j=1}^{M_u}\hat{\theta}_j$ is the mean of M_u IID random variables, and so confidence intervals can be constructed in the usual way if so desired.

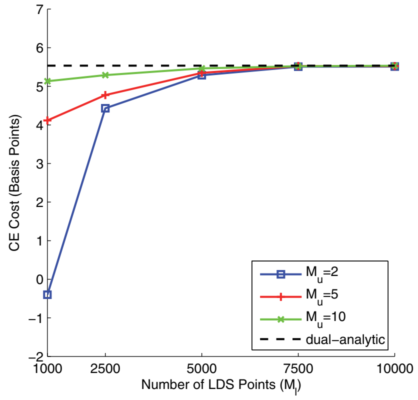
Numerical results. We considered the same portfolio execution problem that we studied in section 4. In Figure 4 we plot the performance of the dual bound as estimated using our LDS scheme for various values of M_u, M_l, and γ. Our LDS was an n-dimensional *Sobol* sequence that we generated using the LDS functionality of MATLAB. We skipped the first 1000 points, retained every 101st point thereafter, and also applied the so-called Matousek–Affine–Owen scrambling scheme.

The dashed lines in Figure 4 represent the dual bound that was computed using the analytic expressions for $\mathbb{E}_t[\hat{V}_{t+1}(\tilde{\mathbf{s}}_{0:t})]$ and $\mathbb{E}_t[\nabla\hat{V}_{t+1}(\tilde{\mathbf{s}}_{0:t})']$. These dashed lines are therefore the same bounds that we calculated in section 4. As expected, we see that the new LDS-based dual bounds are conservative so that they are all at or below the corresponding dashed line. It is also not surprising that as we increase M_l and M_u the bounds improve to the point that with $M_l = 10{,}000$ and $M_u = 10$ the bound is virtually indistinguishable from the dashed lines. (In fact in a couple of cases, e.g., $\gamma = 7 \times 10^{-7}$ and $M_u = 2$, the LDS-based bound was marginally higher than the dashed line, but the difference was well within the Monte Carlo standard error. Recall that even the dashed line has some statistical error since it is computed as the average over 100 dual problem instances. The LDS-based bounds use the same 100 dual paths, and so these bounds are also exposed to this error as well as the nested simulation error of (23).)

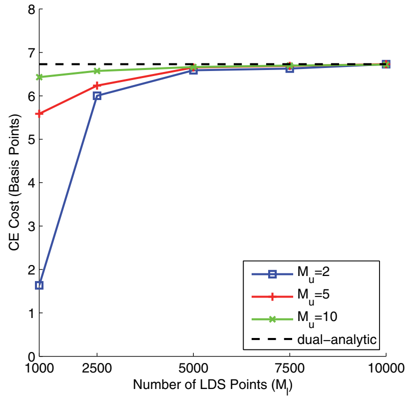
The surprising feature of these results is how little additional work is required. For example, it typically took approximately 150 seconds to solve a dual problem instance when we used the analytic expressions for the dual penalties. When we used the LDS to estimate these penalties with $M_l = 10,000$ and $M_u = 10$ the running time increased by approximately 15 seconds for a relative increase of just 10%. This was possible because it is straightforward to generate and store in advance all of the $(\mathbf{l}_i + \mathbf{U}_j) \bmod \mathbf{1}$ vectors. Moreover, the $g(\cdot)$ function was easy to evaluate since it could be computed explicitly. Of course, the use of exponential utility would force us to consider only those distributions that have a moment generating function.



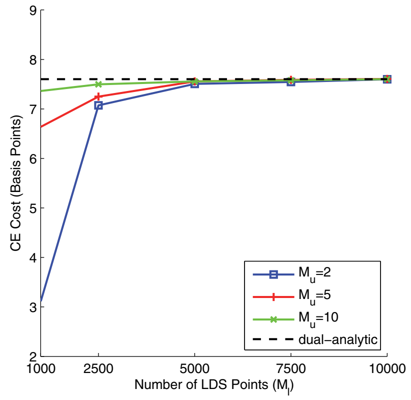
(a) $\gamma = 1 \times 10^{-7}$



(b) $\gamma = 4 \times 10^{-7}$



(c) $\gamma = 7 \times 10^{-7}$



(d) $\gamma = 10 \times 10^{-7}$

Figure 4. *Dual bounds obtained using randomized LDS to estimate dual penalties.*

Other applications. The results of Figure 4 suggest that the randomized LDS approach can be used efficiently to generate very good dual bounds. One possible application would be for solving portfolio execution problems where we suspect that the random return vectors are only approximately normally distributed. In particular, we may suspect that the true return vectors have fatter tails. In that case we could, for example, use the same OLFC policy (which assumed normal returns) as before, but in order to evaluate it using the dual formulation we would need to estimate $\mathbb{E}_t[\hat{V}_{t+1}(\tilde{\mathbf{s}}_{0:t})]$ and $\mathbb{E}_t[\nabla \hat{V}_{t+1}(\tilde{\mathbf{s}}_{0:t})']$ using the *true* return distributions. If these terms cannot be computed analytically, then we could use the approach outlined in this section. If the resulting duality gap, i.e., the difference between the primal and dual bounds, were sufficiently small, then we would know that we could safely ignore the nonnormality of returns. Otherwise, we would need to adapt our policy to take this nonnormality into account.

5.2. Investigating the temporary price-impact model. In section 3.3 we discussed how the dual methodology could be used to determine whether or not certain market features require explicit modeling. We demonstrate this idea here by investigating the features of our temporary price-impact model from section 4.1. Suppose the agent suspects that only the time-varying feature of the temporary price impact is important and that the stochastic component can be safely ignored by the execution policy. To investigate this conjecture we consider two additional policies. The first policy, which we call the OLFC-TV policy, assumes the temporary price impact is time-varying but deterministic. In particular the OLFC-TV policy assumes the temporary price impact at time t is given by $\mathrm{E}_0[\mathbf{B}_t]$. The second policy, which we call the OLFC-C policy, assumes the temporary price-impact is constant and equal to its time average across all time periods. Note that the "true" temporary price-impact model remains as described by (17)–(19).

The performance of these policies is shown in Figure 5 together with the original OLFC policy and our dual bound. Not surprisingly we see that the OLFC policy outperforms the OLFC-TV policy, which in turn outperforms the OLFC-C policy. It is interesting to note the degree of outperformance, however, particularly for higher levels of risk aversion where the OLFC-TV policy performs almost as well as the OLFC policy and much better than the OLFC-C. In this case we could argue that the agent's conjecture was justified, although to draw this conclusion more generally we would need a more careful specification and calibration of the temporary price-impact model.

5.3. Including a nonlinear temporary price impact. In section 2, we introduced a model with linear price impacts, but of course in practice the true price dynamics are generally more complex. In this section we therefore investigate the case where the price dynamics also include a temporary nonlinear price impact. (We cannot include a permanent nonlinear price impact in our modeling framework since this will introduce the possibility of arbitrage; see Huberman and Stanzl [28], for example.) We let $h_t(\cdot)$ denote this nonlinear temporary price-impact function at time t so that (5) is now replaced by

$$\mathbf{p}_t = \tilde{\mathbf{p}}_t + \mathbf{A}_t \mathbf{s}_t + \mathbf{B}_t \mathbf{s}_t + h_t(\mathbf{s}_t). \tag{24}$$

The primal problem. If the agent ignores the nonlinear impact, he will implement a portfolio execution policy that he perceives as being a good policy for the linear price-impact model of section 2. The original OLFC policy of section 2 could play the role of this "good"

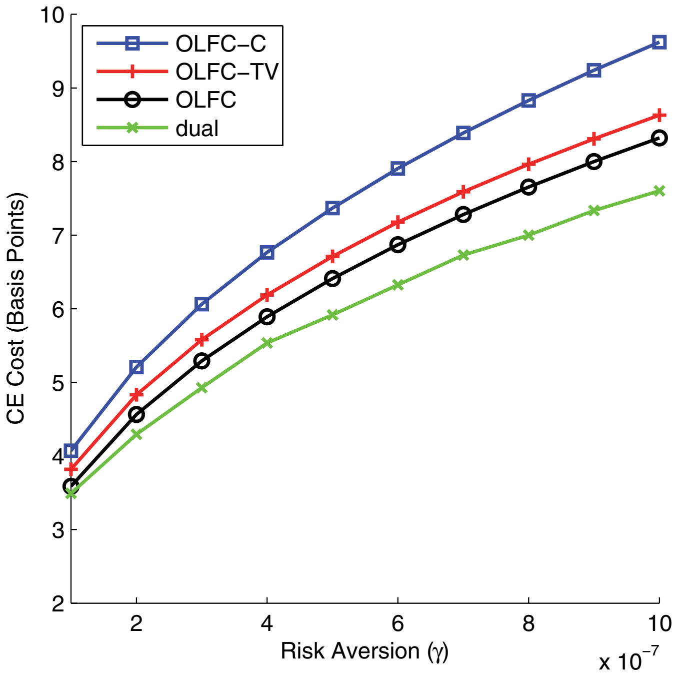


Figure 5. *Certainty equivalent costs.*

policy. Note that the dual methodology can be used to determine whether or not the nonlinear price impact is significant and therefore needs to be accounted for in the policy. In particular, if we find the OLFC policy is close to optimal even in the presence of the nonlinear price impact, then the latter need not be modeled explicitly.

In practice, of course, we would expect the original OLFC policy to be far from optimal, implying that the nonlinear price impact is significant. In that case the agent should explicitly account for the nonlinear impact in constructing the OLFC policy. This would be quite straightforward, and we could formulate the problem analogously to (A-18). We could also find similar expressions to (26) and (27) below for the gradient and Hessian of the OLFC objective function, and it would be reasonable to insist that this Hessian be positive definite for all $\mathbf{s}$ in order to guarantee the convexity of the OLFC problem.

The dual problem. Calculating dual bounds for the nonlinear impact model proceeds exactly as in section 3.1. Many dual problem instances are simulated, and each results in an optimization problem of the form

$$(25)\qquad \begin{aligned} \min_{\mathbf{s}\in\mathbb{S}} \quad & \exp\left(\gamma \sum_{t=0}^{T} \mathbf{p}_t' \mathbf{s}_t\right) + \sum_{t=0}^{T-1} \left(\mathbb{E}_t[\tilde{V}_{t+1}(\mathbf{s}_{0:t})] - \tilde{V}_{t+1}(\mathbf{s}_{0:t})\right) \\ \text{subject to} \quad & \mathbf{p}_t = \tilde{\mathbf{p}}_t + \mathbf{A}_t \mathbf{s}_t + \mathbf{B}_t \mathbf{s}_t + h_t(\mathbf{s}_t), \\ & \tilde{\mathbf{p}}_{t+1} = \tilde{\mathbf{p}}_t + \mathbf{r}_{t+1} + \mathbf{A}_t \mathbf{s}_t, \end{aligned}$$

where $\tilde{V}_{t+1}$ is some suitably linearized approximate value function. The analogue to (C-35) under the nonlinear temporary price-impact model is

$$f(\mathbf{s}) := \exp\left(\gamma \sum_{t=0}^{T} \left(\tilde{\mathbf{p}}_0 + \sum_{i=0}^{t-1} \mathbf{A}_i \mathbf{s}_i + \sum_{i=1}^{t} \mathbf{r}_i + \mathbf{A}_t \mathbf{s}_t + \mathbf{B}_t \mathbf{s}_t + h_t(\mathbf{s}_t)\right)' \mathbf{s}_t\right)$$
$$= \exp(\gamma \tilde{\mathbf{p}}_0' \mathbf{w}_0)\ \exp\left(\gamma \left(\frac{1}{2}\mathbf{s}'\mathbf{Q}\mathbf{s} + \mathbf{c}_{p,0}'\mathbf{s} + \tilde{h}(\mathbf{s})\right)\right).$$

The gradient vector and Hessian matrix of f are

$$\nabla f(\mathbf{s}) = \gamma f(\mathbf{s})(\mathbf{Q}\mathbf{s} + \mathbf{c}_{p,0} + \nabla \tilde{h}(\mathbf{s})), \tag{26}$$

$$Hf(\mathbf{s}) = \gamma f(\mathbf{s})\Big(\mathbf{Q} + H\tilde{h}(\mathbf{s}) + \gamma(\mathbf{Q}\mathbf{s} + \mathbf{c}_{p,0} + \nabla\tilde{h}(\mathbf{s}))(\mathbf{Q}\mathbf{s} + \mathbf{c}_{p,0} + \nabla\tilde{h}(\mathbf{s}))'\Big), \tag{27}$$

where $\tilde{h}(\mathbf{s}) := \sum_{t=0}^{T} h_t'(\mathbf{s}_t)\mathbf{s}_t$, $\nabla\tilde{h}(\mathbf{s})$ is the gradient of $\tilde{h}$, and $H\tilde{h}(\mathbf{s})$ is the Hessian matrix of $\tilde{h}$. A sufficient condition to guarantee that $Hf(\mathbf{s})$ is positive definite (so that f is therefore convex) is that $\tilde{h}$ is convex and that $\mathbf{Q}$ is positive definite.

5.4. Predictable state dynamics. Some of the earlier papers such as Bertsimas and Lo [10] and Bertsimas, Hummel, and Lo [9] (some details are available in Appendix D) allow for a predictable component in the return dynamics. In particular, they assumed the price dynamics depended on a state variable, X_t, which itself had linear dynamics. Assuming the agent is risk-neutral, they were able to solve recursively for the optimal solution in the same manner as the analysis of Appendix A.4.

We could include a similar state variable or vector in our model. Except for some specific circumstances, we would not be able to solve for the optimal feedback control policy, but we could still solve for the OLFC policy in that case. It should also be clear that we can again use a modified version of the OLFC value function to compute dual penalties and therefore obtain valid dual bounds. We do not pursue this any further since allowing for predictable price dynamics is generally of less interest than the accurate modeling of price impact and liquidity effects.

5.5. Limit order book models. Building on the earlier single-stock work of Obizhaeva and Wang [38], Alfonsi, Fruth, and Scheid [2] and Predoiu, Shaikhet, and Shreve [39], Tsoukalas, Wang, and Giesecke [46] formulate the portfolio execution problem at the limit order book (LOB) level. Here we briefly discuss this model and note that the dual methodology may also be applied in this context.

Their model assumes a two-sided block-shaped order book with infinite depth and time-invariant density. The midprice is a random walk with zero drift. A buy (sell) order $\mathbf{s}_t^+$ ($\mathbf{s}_t^-$) is executed as a market order against available inventory in the ask (bid) side of the LOB, creating a linear temporary price impact that moves the best ask (bid) price away from the midprice. The order also creates a linear permanent price by moving the midprice up (down). Newly arriving limit orders replenish the inventory and cause the distance between the best ask (bid) price and the midprice to decay exponentially. The execution problem is then formulated as

$$\max_{\mathbf{s}_t^+,\mathbf{s}_t^- \in \mathcal{F}_t, t=0,\ldots,T} \quad \mathbb{E}_0\left[-\exp\left(-\gamma\sum_{t=0}^{T}\left((\mathbf{p}_t^b - \mathbf{B}^b\mathbf{s}_t^-)'\mathbf{s}_t^- - (\mathbf{p}_t^a + \mathbf{B}^a\mathbf{s}_t^+)'\mathbf{s}_t^+\right)\right)\right] \tag{28}$$

$$\text{subject to} \quad \mathbf{p}_t^a = \tilde{\mathbf{p}}_0 + \sum_{j=1}^{t}\mathbf{r}_j + \frac{1}{2}\boldsymbol{\delta}_t + \mathbf{A}\sum_{j=0}^{t-1}(\mathbf{s}_j^+ - \mathbf{s}_j^-) + \mathbf{d}_t^a, \tag{29}$$

$$\mathbf{p}_t^b = \tilde{\mathbf{p}}_0 + \sum_{j=1}^{t}\mathbf{r}_j - \frac{1}{2}\boldsymbol{\delta}_t + \mathbf{A}\sum_{j=0}^{t-1}(\mathbf{s}_j^+ - \mathbf{s}_j^-) + \mathbf{d}_t^b, \tag{30}$$

$$\mathbf{d}_t^a = (\mathbf{d}_{t-1}^a + \mathbf{B}^a\mathbf{s}_{t-1}^+ - \mathbf{A}(\mathbf{s}_{t-1}^+ - \mathbf{s}_{t-1}^-))\exp(-\boldsymbol{\rho}^a\Delta t), \tag{31}$$

$$\mathbf{d}_t^b = (\mathbf{d}_{t-1}^b - \mathbf{B}^b\mathbf{s}_{t-1}^- - \mathbf{A}(\mathbf{s}_{t-1}^+ - \mathbf{s}_{t-1}^-))\exp(-\boldsymbol{\rho}^b\Delta t), \tag{32}$$

$$\sum_{t=0}^{T}(\mathbf{s}_t^+ - \mathbf{s}_t^-) = \mathbf{w}_0 \quad \text{and} \quad \mathbf{s}_t^+, \mathbf{s}_t^- \geq \mathbf{0}, \tag{33}$$

where $\mathbf{p}_t^a$ and $\mathbf{p}_t^b$ are the best ask/bid prices in the LOB at time t, $\tilde{\mathbf{p}}_0$ is the initial midprice, and $\boldsymbol{\delta}_t$ is the bid-ask spread. $\mathbf{A}$ is the permanent price-impact coefficient matrix, while $\mathbf{B}^a$ and $\mathbf{B}^b$ are the temporary price-impact coefficients for buy and sell orders, respectively. Each of these matrices is assumed to be deterministic. In addition, $\mathbf{B}^a$ and $\mathbf{B}^b$ are diagonal matrices with diagonal elements equal to the inverse of the order book density on the ask and bid sides, respectively. $\mathbf{d}_t^a$ and $\mathbf{d}_t^b$ track the deviation of the current ask and bid prices from their steady state levels, and they decay exponentially at constant speeds $\boldsymbol{\rho}^a$ and $\boldsymbol{\rho}^b$ due to order book replenishment. Δt is the length of a time period. The only uncertainty comes from the returns, $\mathbf{r}_j$, which are IID normally distributed with zero mean and a constant covariance matrix. The objective is to maximize an exponential utility function over the terminal wealth which is equivalent to minimizing the exponential utility function over cost in our model.

Tsoukalas, Wang, and Giesecke [46] show that the optimal execution policy is deterministic and find an equivalent quadratic formulation for the problem. Such a result is quite standard in the portfolio optimization literature, and indeed the equivalence of our formulations in Appendices A.3 and A.4 are in the same spirit. But once we begin to relax some of the assumptions in this model, then solving for the optimal policy even numerically becomes a very challenging task and we again find ourselves in the situation of needing to construct and evaluate good suboptimal policies. For example if we introduce a state vector, Z_t, that drives the dynamics of the $\mathbf{r}_t$'s and the temporary price-impact matrices, then a dual problem instance will take the form

$$\max_{\mathbf{s}} \quad -\exp\left(-\gamma\sum_{t=0}^{T}\left((\mathbf{p}_t^b - \mathbf{B}_t^b\mathbf{s}_t^-)'\mathbf{s}_t^- - (\mathbf{p}_t^a + \mathbf{B}_t^a\mathbf{s}_t^+)'\mathbf{s}_t^+\right)\right) + \sum_{t=0}^{T-1}\left(\mathbb{E}_t[\tilde{V}_{t+1}(\mathbf{s}_{0:t})] - \tilde{V}_{t+1}(\mathbf{s}_{0:t})\right) \tag{34}$$

subject to (29) to (33) but with $\mathbf{B}^a$ and $\mathbf{B}^b$ replaced by $\mathbf{B}_j^a$ and $\mathbf{B}_j^b$, where $j = t$ or $t-1$ as appropriate. Note again that Z_t does not appear explicitly in any dual problem instance, but instead it changes the distribution of these instances. We see that the dual problem in (34) is similar to the dual problem of section 3 and that we could approach this problem in a similar manner.

6. Conclusions and future research.

Any realistic model of portfolio execution should be able to handle features such as stochastic variance-covariance dynamics, return predictability, time-of-day effects, and stochastic liquidity/price impacts, and possible risk aversion. Any model which includes these features will not in general be analytically tractable, and so it will be necessary instead to construct good suboptimal policies. It is important that these suboptimal policies can be properly evaluated, and in this paper we have demonstrated the use of duality methods to do this. In particular, our model is capable of capturing all of the above effects and our OLFC policy appears to be capable of generating very good primal bounds, while a variation of the associated OLFC value function also leads to very tight dual bounds. (Of course we have shown this to actually be the case only for the parameter settings we considered, but we suspect it to be true more generally for realistic parameter settings.) While these dual methods have become quite standard in a relatively short period of time, implementing them is not straightforward in general since the dual optimization problems can be quite complex and establishing convexity or nonconvexity may also require some effort.

An additional contribution of this paper is that valid and tight dual bounds can still be computed efficiently even when the dual penalties are not explicitly available and need to be estimated via Monte Carlo. The numerical experiments of section 5.1 where we used randomized LDS have demonstrated the validity of this broadly applicable approach.

We have also demonstrated some useful properties of the dual problem in the context of portfolio execution problems. For example, we noted that the convexity of the dual problem does not depend on variance-covariance dynamics or state variable dynamics as long as these dynamics are not influenced by the execution policy. We have also noted how the duality technology can be used to determine in advance whether a market feature, e.g., the cross price impact of Appendix D or the nonlinear price impacts of section 5.3, need to be accounted for by the portfolio execution policy.

There are several possible future research directions, and these include conducting a more detailed calibration and empirical evaluation of the model we have proposed in this paper. Another direction is to use duality to study the impact of using misspecified return distributions. For example, the assumption of (conditional) multivariate normality has been made throughout this paper and throughout the literature on single-stock and portfolio execution. This of course is due to the tractability of the normal distribution. In practice, however, we might expect conditional returns to be only approximately normal. We could still construct feasible execution policies in this case by *pretending* that everything was normally distributed, after which we could use the true model to simulate sample paths to estimate valid primal and dual bounds. Of course when computing the penalty terms we require expectations of the linearized approximate value functions and these expectations must be evaluated under the *true* model. If they are not available explicitly, then they can be estimated via nested Monte Carlos as described in section 5.

Another potentially interesting direction is the problem of multiportfolio execution, where firms need to solve the portfolio execution problem simultaneously for multiple portfolios rather than just one portfolio. This raises the question as to how we should prioritize the portfolios. We could, for example, solve each portfolio execution problem separately and ignore all knowledge of the other portfolio trades that need to be executed. We would expect this to be far from optimal in general, although it would be necessary to first define what is meant by "optimal." Towards this end, we would also need to determine how to allocate the total execution costs fairly across each of the portfolios, which is difficult when impact costs are nonlinear. The multiportfolio execution problem is nontrivial and, depending on the solution technique, can draw on concepts from cooperative or noncooperative game theory. It was first studied by O'Cinneide, Scherer, and Xu [16], who proposed optimizing the total "social welfare" without accounting for fairness. Stubbs and Vandenbussche [44] study the pros and cons of both the cooperative and noncooperative solution techniques in a static rather than dynamic context. Once the multiportfolio execution problem has been properly formulated, however, then it will be necessary to determine good execution policies. The duality technology might be very useful in assessing the quality of these policies.

A final direction for future research is understanding just how plausible it is to impose conditions that guarantee the convexity of all dual problem instances. There should be little difficulty in justifying this assumption when the price impacts are assumed to be deterministic. Huberman and Stanzl [28] provide such a justification in the single-stock case, for example. When the price impacts are stochastic, however, insisting on the positive definiteness of the matrix, $\mathbf{Q}$, in (C-37) on all dual problem instances is not so easy to justify from an economic standpoint. From an intuitive viewpoint it seems plausible to insist that $\mathbf{Q}$ should be positive definite in some *average* sense (as implied by the positive definiteness of (A-19), for example), but there is no reason to assume that all realizations of $\mathbf{Q}$ need to be positive definite to avoid the absence of arbitrage.

If we were to consider models where dual convexity was not guaranteed, then it should still be possible to obtain a valid dual bound. In particular, on each dual sample path we first check whether $\mathbf{Q}$ is positive definite. If it is, then we solve this dual problem instance as before. If $\mathbf{Q}$ is not positive definite, then the dual optimization problem is difficult to solve in general. But we can still bound the optimal value of such a problem instance. For example, the exponential term in (12) is bounded below by zero, and if the constraint set is a bounded polyhedron (as will typically be the case in practice), then LP techniques could be used to bound the second term in (12). If relatively few of the $\mathbf{Q}$'s fail to be positive definite, then we may still be able to obtain a good dual bound in this manner. Another possibility would be to form the concave dual (see Boyd and Vanderberghe [11], for example) of the nonconvex dual instance. The optimal solution (or indeed a good feasible solution) to this concave problem will also provide a valid lower bound, although it may be difficult to compute.

Ultimately we believe the goal is to construct realistic models that can include the various market effects that are found in practice, that can handle hundreds of securities or more, that are straightforward to calibrate, and for which good suboptimal policies can easily be found. Because it is impossible to ever know the true market dynamics, we believe these suboptimal policies need to be robust to deviations from the assumed model and that duality based on information relaxations can play a key role in assessing this robustness.

Appendix A. Solving for the suboptimal policies. Here we consider the approaches we follow for obtaining the various policies outlined in section 2.2. In each of Appendices A.1, A.2, and A.4 we assume that no-short-sales constraints are *not* imposed. When we discuss convexity and existence of solutions we restrict ourselves to positive definite matrices rather than positive semidefinite matrices for ease of exposition.

A.1. The risk-neutral policy. Let $V_t(\cdot)$ denote the time t value function for the problem faced by a risk-neutral agent. At time T the agent needs to buy all remaining shares so that $\mathbf{s}_T = \mathbf{w}_T$. We also have

$$\begin{aligned} V_T(\tilde{\mathbf{p}}_T, \mathbf{w}_T, \mathbf{A}_T, \mathbf{B}_T) &= \mathbf{p}_T' \mathbf{s}_T = (\tilde{\mathbf{p}}_T + \mathbf{A}_T \mathbf{w}_T + \mathbf{B}_T \mathbf{w}_T)' \mathbf{w}_T \\ &= \frac{1}{2} \mathbf{w}_T' (\mathbf{A}_T + \mathbf{A}_T' + \mathbf{B}_T + \mathbf{B}_T') \mathbf{w}_T + \tilde{\mathbf{p}}_T' \mathbf{w}_T \\ &= \frac{1}{2} \mathbf{w}_T' \mathbf{G}_T \mathbf{w}_T + \tilde{\mathbf{p}}_T' \mathbf{w}_T, \end{aligned} \tag{A-1}$$

where $\mathbf{G}_T := \mathbf{A}_T + \mathbf{A}_T' + \mathbf{B}_T + \mathbf{B}_T'$. At time $T-1$ we need to solve

$$V_{T-1}(\tilde{\mathbf{p}}_{T-1}, \mathbf{w}_{T-1}, \mathbf{A}_{T-1}, \mathbf{B}_{T-1}) = \min_{\mathbf{s}_{T-1}} \mathbb{E}_{T-1} \Big[\mathbf{p}_{T-1}' \mathbf{s}_{T-1} + V_T(\tilde{\mathbf{p}}_T, \mathbf{w}_T, \mathbf{A}_T, \mathbf{B}_T) \Big] \tag{A-2}$$

$$\begin{aligned} &= \min_{\mathbf{s}_{T-1}} \frac{1}{2} \mathbf{s}_{T-1}' \mathbf{N}_{ss,T-1} \mathbf{s}_{T-1} - (\mathbf{N}_{ws,T-1} \mathbf{w}_{T-1})' \mathbf{s}_{T-1} \\ &\quad + \frac{1}{2} \mathbf{w}_{T-1}' \mathbb{E}_{T-1}[\mathbf{G}_T] \mathbf{w}_{T-1} + \tilde{\mathbf{p}}_{T-1}' \mathbf{w}_{T-1}, \end{aligned} \tag{A-3}$$

where we have used (A-1) to substitute for V_T in (A-2) and then used the price dynamics of section 2.1 and $\mathbf{w}_T = \mathbf{w}_{T-1} - \mathbf{s}_{T-1}$ in obtaining (A-3) and where we have defined

$$\mathbf{N}_{ss,T-1} := \mathbb{E}_{T-1}[\mathbf{G}_T] + \mathbf{B}_{T-1} + \mathbf{B}_{T-1}', \tag{A-4}$$

$$\mathbf{N}_{ws,T-1} := \mathbb{E}_{T-1}[\mathbf{G}_T] - \mathbf{A}_{T-1}'. \tag{A-5}$$

Assuming $\mathbf{N}_{ss,T-1}$ is positive definite, then the objective function in (A-3) is convex and the optimal solution to this problem is given by

$$\mathbf{s}_{T-1}^* = \mathbf{N}_{ss,T-1}^{-1} \mathbf{N}_{ws,T-1} \mathbf{w}_{T-1} \tag{A-6}$$

with the optimal value function satisfying

$$V_{T-1}(\tilde{\mathbf{p}}_{T-1}, \mathbf{w}_{T-1}, \mathbf{A}_{T-1}, \mathbf{B}_{T-1}) = \frac{1}{2} \mathbf{w}_{T-1}' \mathbf{G}_{T-1} \mathbf{w}_{T-1} + \tilde{\mathbf{p}}_{T-1}' \mathbf{w}_{T-1},$$

where

$$\mathbf{G}_{T-1} := \mathbb{E}_{T-1}[\mathbf{G}_T] - \mathbf{N}_{ws,T-1}' \mathbf{N}_{ss,T-1}^{-1} \mathbf{N}_{ws,T-1}. \tag{A-7}$$

Continuing in this manner we obtain, for $t = T-2, \ldots, 0$,

$$\begin{aligned} V_t(\tilde{\mathbf{p}}_t, \mathbf{w}_t, \mathbf{A}_t, \mathbf{B}_t) &= \min_{\mathbf{s}_t} \mathbb{E}_t \Big[\mathbf{p}_t' \mathbf{s}_t + V_{t+1}(\tilde{\mathbf{p}}_{t+1}, \mathbf{w}_{t+1}, \mathbf{A}_{t+1}, \mathbf{B}_{t+1}) \Big] \\ &= \min_{\mathbf{s}_t} \frac{1}{2} \mathbf{s}_t' \mathbf{N}_{ss,t} \mathbf{s}_t - (\mathbf{N}_{ws,t} \mathbf{w}_t)' \mathbf{s}_t + \frac{1}{2} \mathbf{w}_t' \mathbb{E}_t[\mathbf{G}_{t+1}] \mathbf{w}_t + \tilde{\mathbf{p}}_t' \mathbf{w}_t, \end{aligned}$$

where

$$\text{(A-8)} \qquad \mathbf{N}_{ss,t} := \mathbb{E}_t[\mathbf{G}_{t+1}] + \mathbf{B}_t + \mathbf{B}_t',$$
$$\text{(A-9)} \qquad \mathbf{N}_{ws,t} := \mathbb{E}_t[\mathbf{G}_{t+1}] - \mathbf{A}_t'.$$

Again assuming $\mathbf{N}_{ss,t}$ is positive definite, the optimal solution satisfies

$$\text{(A-10)} \qquad \mathbf{s}_t^* = \mathbf{N}_{ss,t}^{-1}\mathbf{N}_{ws,t}\mathbf{w}_t$$

and the optimal value function satisfies

$$\text{(A-11)} \qquad V_t(\tilde{\mathbf{p}}_t, \mathbf{w}_t, \mathbf{A}_t, \mathbf{B}_t) = \frac{1}{2}\mathbf{w}_t'\mathbf{G}_t\mathbf{w}_t + \tilde{\mathbf{p}}_t'\mathbf{w}_t,$$

where

$$\text{(A-12)} \qquad \mathbf{G}_t \;:=\; \mathbb{E}_t[\mathbf{G}_{t+1}] - \mathbf{N}_{ws,t}'\mathbf{N}_{ss,t}^{-1}\mathbf{N}_{ws,t}.$$

We note that the optimal trading quantities, $\mathbf{s}_t^*$, depend on $\mathbf{w}_t$ only, and not on the price vector $\tilde{\mathbf{p}}_t$ or the stochastic variance-covariance process. It is not path-independent, however, due to its dependence on $\mathbf{A}_t$ and $\mathbf{B}_t$, which are in general stochastic. Note that if any $\mathbf{N}_{ss,t}$ fails to be positive definite, then the agent's time t objective function will be unbounded from below (in the absence of constraints) and economic considerations alone would imply that this possibility should be ruled out. One way to do this would be to first impose sufficient structure on the dynamics of $\mathbf{A}_t$ and $\mathbf{B}_t$ so that $\mathbb{E}_t[\mathbf{G}_{t+1}]$ can be computed analytically. We could then look to impose additional conditions that guarantee the positive definiteness of the $\mathbf{N}_{ss,t}$'s.

Regardless, we can implement the policy given by (A-10) only if we can solve for the $\mathbf{G}_t$'s analytically. One situation where it is straightforward to determine the optimality of (A-10) and actually implement the policy is when the the $\mathbf{A}_t$'s and $\mathbf{B}_t$'s are deterministic. Indeed if we assume the $\mathbf{A}_t$'s and $\mathbf{B}_t$'s are constant across time, then it is possible to determine explicit expressions for $\mathbf{s}_t^*$. In Appendix A.2 we also identify a particular case where the risk-neutral and simple policies coincide.

In the numerical results of section 4 we assumed that the $\mathbf{B}_t$'s were stochastic, and so we were not able to implement the risk-neutral policy. Instead we implemented the risk-neutral OLFC policy as discussed in section 2.2 and at the end of Appendix A.3.

A.2. When the simple and risk-neutral policies coincide. Suppose the following two conditions both hold:

(i) $\mathbf{A}_t$ and $\mathbf{B}_t$ are martingales: $\mathbb{E}_t[\mathbf{A}_{t+1}] = \mathbf{A}_t$ and $\mathbb{E}_t[\mathbf{B}_{t+1}] = \mathbf{B}_t$.

(ii) $\mathbf{A}_t$ and $\mathbf{B}_t$ are symmetric: $\mathbf{A}_t = \mathbf{A}_t'$ and $\mathbf{B}_t = \mathbf{B}_t'$.

Then the simple and risk-neutral policies coincide, and we can show this by induction as follows. At time T, we have $\mathbf{G}_T = 2\mathbf{A}_T + 2\mathbf{B}_T$ and $\mathbf{s}_T^* = \mathbf{w}_T$. Under conditions (i) and (ii), (A-4) to (A-7) reduce to

$$\begin{aligned}
\mathbf{N}_{ss,T-1} &= \mathbb{E}_{T-1}[\mathbf{G}_T] + \mathbf{B}_{T-1} + \mathbf{B}'_{T-1} = 2\mathbf{A}_{T-1} + 4\mathbf{B}_{T-1},\\
\mathbf{N}_{ws,T-1} &= \mathbb{E}_{T-1}[\mathbf{G}_T] - \mathbf{A}'_{T-1} = \mathbf{A}_{T-1} + 2\mathbf{B}_{T-1},\\
\mathbf{s}^*_{T-1} &= \mathbf{N}^{-1}_{ss,T-1}\mathbf{N}_{ws,T-1}\mathbf{w}_{T-1} = \frac{1}{2}\mathbf{w}_{T-1},\\
\mathbf{G}_{T-1} &= \mathbb{E}_{T-1}[\mathbf{G}_T] - \mathbf{N}'_{ws,T-1}\mathbf{N}^{-1}_{ss,T-1}\mathbf{N}_{ws,T-1} = \left(1+\frac{1}{2}\right)\mathbf{A}_{T-1} + \frac{1}{2}2\mathbf{B}_{T-1}
\end{aligned}$$

and in this case the simple and risk-neutral policies do indeed coincide. Suppose now that at time $t+1$ we have

$$\mathbf{G}_{t+1} = \left(1 + \frac{1}{T-(t+1)+1}\right)\mathbf{A}_{t+1} + \frac{1}{T-(t+1)+1}2\mathbf{B}_{t+1}$$

and

$$\mathbf{s}^*_{t+1} = \frac{1}{T-(t+1)+1}\mathbf{w}_{t+1}.$$

Then (A-8), (A-9), (A-12), and (A-10) yield

$$\begin{aligned}
\mathbf{N}_{ss,t} &= \mathbb{E}_t[\mathbf{G}_{t+1}] + \mathbf{B}_t + \mathbf{B}'_t = \left(1+\frac{1}{T-t}\right)\mathbf{A}_t + \left(1+\frac{1}{T-t}\right)2\mathbf{B}_t,\\
\mathbf{N}_{ws,t} &= \mathbb{E}_t[\mathbf{G}_{t+1}] - \mathbf{A}'_t = \frac{1}{T-t}\mathbf{A}_t + \frac{1}{T-t}2\mathbf{B}_t,\\
\mathbf{G}_t &= \mathbb{E}_t[\mathbf{G}_{t+1}] - \mathbf{N}'_{ws,t}\mathbf{N}^{-1}_{ss,t}\mathbf{N}_{ws,t} = \left(1+\frac{1}{T-t+1}\right)\mathbf{A}_t + \frac{1}{T-t+1}2\mathbf{B}_t,
\end{aligned}$$

$$\mathbf{s}^*_t = \mathbf{N}^{-1}_{ss,t}\mathbf{N}_{ws,t}\mathbf{w}_t = \frac{1}{T-t+1}\mathbf{w}_t, \tag{A-13}$$

and by (A-13) we see the inductive step is complete.

A.3. Computing the OLFC policy when $\mathbf{s}_j \geq 0$ is imposed. We now consider calculation of the OLFC when nonnegativity constraints on the $\mathbf{s}_j$'s are imposed. It should be clear that imposing additional convex constraints on the $\mathbf{s}_j$'s would also be straightforward. We first note that (5) and (6) imply

$$\mathbf{p}_j = \tilde{\mathbf{p}}_t + \sum_{i=t}^{j-1}\mathbf{A}_i\mathbf{s}_i + \sum_{i=t+1}^{j}\mathbf{r}_i + \mathbf{A}_j\mathbf{s}_j + \mathbf{B}_j\mathbf{s}_j \quad \text{for } j = t,\ldots,T. \tag{A-14}$$

The OLFC policy assumes the price-impact matrices evolve deterministically by taking their time-t conditional expectations. We therefore substitute (A-14) into the objective function in (10) but with $\mathbf{A}_j$ and $\mathbf{B}_j$ replaced by $\mathbb{E}_t[\mathbf{A}_j]$ and $\mathbb{E}_t[\mathbf{B}_j]$, respectively. We then see the OLFC problem can be formulated as

(A-15)

$$\min_{\mathbf{s}_{t:T}\geq 0} \mathbb{E}_t\left[\exp\left(\gamma\tilde{\mathbf{p}}_t'\mathbf{w}_t + \gamma\sum_{j=t}^{T}\left(\sum_{i=t}^{j-1}\mathbb{E}_t[\mathbf{A}_i]\mathbf{s}_i + \mathbb{E}_t[\mathbf{A}_j]\mathbf{s}_j + \mathbb{E}_t[\mathbf{B}_j]\mathbf{s}_j\right)'\mathbf{s}_j + \gamma\sum_{j=t+1}^{T}\mathbf{r}_j'\mathbf{w}_j\right)\right]$$

(A-16)

$$\equiv \min_{\mathbf{s}_{t:T}\geq 0} \exp\left(\gamma\sum_{j=t}^{T}\left(\sum_{i=t}^{j-1}\mathbb{E}_t[\mathbf{A}_i]\mathbf{s}_i + \mathbb{E}_t[\mathbf{A}_j]\mathbf{s}_j + \mathbb{E}_t[\mathbf{B}_j]\mathbf{s}_j\right)'\mathbf{s}_j\right)\mathbb{E}_t\left[\exp\left(\gamma\sum_{j=t+1}^{T}\mathbf{r}_j'\mathbf{w}_j\right)\right],$$

where we have used the fact that $\mathbf{w}_j = \sum_{i=j}^{T}\mathbf{s}_i$ for $j = t, \ldots, T$ and switched the order of summation to obtain the last term in (A-15). The OLFC policy also assumes the $\mathbf{r}_j$'s are IID normal with mean vector $\mathbf{0}$ and conditional covariance matrix $\mathbf{\Sigma}_j = \mathbf{\Sigma}_t$ for $j = t+1, \ldots, T$. Under this assumption we have

$$\mathbb{E}_t\left[\exp\left(\gamma\sum_{j=t+1}^{T}\mathbf{r}_j\mathbf{w}_j\right)\right] = \exp\left(\sum_{j=t+1}^{T}\frac{1}{2}\gamma^2\mathbf{w}_j'\mathbf{\Sigma}_t\mathbf{w}_j\right)$$

so that (A-16) becomes

$$\text{(A-17)}\quad \min_{\mathbf{s}_{t:T}\geq 0} \exp\left(\gamma\sum_{j=t}^{T}\left(\sum_{i=t}^{j-1}\mathbb{E}_t[\mathbf{A}_i]\mathbf{s}_i + \mathbb{E}_t[\mathbf{A}_j]\mathbf{s}_j + \mathbb{E}_t[\mathbf{B}_j]\mathbf{s}_j\right)'\mathbf{s}_j + \frac{1}{2}\gamma^2\sum_{j=t+1}^{T}\mathbf{w}_j'\mathbf{\Sigma}_t\mathbf{w}_j\right).$$

Because the exponential function is monotonic increasing, we can ignore it so that solving (A-17) reduces to the following constrained quadratic programming problem:

$$\text{(A-18)}\quad \begin{aligned} &\min_{\mathbf{s}_{t:T}} \quad \frac{1}{2}[\mathbf{s}_t' \ldots \mathbf{s}_T'](\mathbf{Q}_{OLFC,t} + \mathbf{Q}_{\mathbf{\Sigma},t})[\mathbf{s}_t' \ldots \mathbf{s}_T']' \\ &\text{subject to} \quad \sum_{j=t}^{T}\mathbf{s}_j = \mathbf{w}_t, \\ &\qquad\qquad \mathbf{s}_j \geq \mathbf{0} \quad \text{for } j = t, \ldots, T, \end{aligned}$$

where

(A-19)

$$\mathbf{Q}_{OLFC,t} := \begin{bmatrix} \mathbb{E}_t[\mathbf{A}_t + \mathbf{A}_t' + \mathbf{B}_t + \mathbf{B}_t'] & \mathbb{E}_t[\mathbf{A}_t'] & \ldots & \mathbb{E}_t[\mathbf{A}_t'] \\ \mathbb{E}_t[\mathbf{A}_t] & \mathbb{E}_t[\mathbf{A}_{t+1} + \mathbf{A}_{t+1}' + \mathbf{B}_{t+1} + \mathbf{B}_{t+1}'] & \ldots & \mathbb{E}_t[\mathbf{A}_{t+1}'] \\ \ldots & \ldots & \ldots & \ldots \\ \mathbb{E}_t[\mathbf{A}_t] & \mathbb{E}_t[\mathbf{A}_{t+1}] & \ldots & \mathbb{E}_t[\mathbf{A}_T + \mathbf{A}_T' + \mathbf{B}_T + \mathbf{B}_T'] \end{bmatrix},$$

and

$$\mathbf{Q}_{\mathbf{\Sigma},t} := \gamma\begin{bmatrix} \mathbf{O} & \mathbf{O} & \mathbf{O} & \ldots & \mathbf{O} \\ \mathbf{O} & \mathbf{\Sigma}_t & \mathbf{\Sigma}_t & \ldots & \mathbf{\Sigma}_t \\ \mathbf{O} & \mathbf{\Sigma}_t & 2\mathbf{\Sigma}_t & \ldots & 2\mathbf{\Sigma}_t \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \mathbf{O} & \mathbf{\Sigma}_t & 2\mathbf{\Sigma}_t & \ldots & (T-t)\mathbf{\Sigma}_t \end{bmatrix},$$

where $\mathbf{O}$ denotes the $n \times n$ zero matrix. We first note that $\mathbf{Q}_{\mathbf{\Sigma},t}$ may be expressed as a positively weighted sum of positive semidefinite matrices according to

$$\mathbf{Q}_{\mathbf{\Sigma},t} = \gamma \begin{bmatrix} \mathbf{O} & \mathbf{O} & \mathbf{O} & \dots & \mathbf{O} \\ \mathbf{O} & \mathbf{\Sigma}_t & \mathbf{\Sigma}_t & \dots & \mathbf{\Sigma}_t \\ \mathbf{O} & \mathbf{\Sigma}_t & \mathbf{\Sigma}_t & \dots & \mathbf{\Sigma}_t \\ \dots & \dots & \dots & \dots & \dots \\ \mathbf{O} & \mathbf{\Sigma}_t & \mathbf{\Sigma}_t & \dots & \mathbf{\Sigma}_t \end{bmatrix} + \gamma \begin{bmatrix} \mathbf{O} & \mathbf{O} & \mathbf{O} & \dots & \mathbf{O} \\ \mathbf{O} & \mathbf{O} & \mathbf{O} & \dots & \mathbf{O} \\ \mathbf{O} & \mathbf{O} & \mathbf{\Sigma}_t & \dots & \mathbf{\Sigma}_t \\ \dots & \dots & \dots & \dots & \dots \\ \mathbf{O} & \mathbf{O} & \mathbf{\Sigma}_t & \dots & \mathbf{\Sigma}_t \end{bmatrix} + \dots + \gamma \begin{bmatrix} \mathbf{O} & \mathbf{O} & \mathbf{O} & \dots & \mathbf{O} \\ \mathbf{O} & \mathbf{O} & \mathbf{O} & \dots & \mathbf{O} \\ \mathbf{O} & \mathbf{O} & \mathbf{O} & \dots & \mathbf{O} \\ \dots & \dots & \dots & \dots & \dots \\ \mathbf{O} & \mathbf{O} & \mathbf{O} & \dots & \mathbf{\Sigma}_t \end{bmatrix},$$

and so it follows that $\mathbf{Q}_{\mathbf{\Sigma},t}$ itself is positive semidefinite. As stated in section 3.1 it seems appropriate to insist on the positive definiteness of $\mathbf{Q}_{OLFC,t}$ so that the OLFC decision-maker does not perceive that there are arbitrage opportunities in the market. In Appendix C.1 we will provide sufficient conditions to guarantee that $\mathbf{Q}_{OLFC,t}$ is indeed positive definite. If these conditions are satisfied (as they are in the numerical experiments of section 4), then $\mathbf{Q}_{OLFC,t} + \mathbf{Q}_{\mathbf{\Sigma},t}$ will be positive definite so that the OLFC policy can be found by solving a constrained convex quadratic optimization problem.

We also note that the *risk-neutral OLFC* policy can be obtained by taking $\gamma = 0$ and then solving the optimization problem in (A-18). This results in an objective function of the form

$$\min_{\mathbf{s}_{t:T}} \quad \frac{1}{2}[\mathbf{s}_t' \, \dots \, \mathbf{s}_T'] \mathbf{Q}_{OLFC,t} [\mathbf{s}_t' \, \dots \, \mathbf{s}_T']'.$$

A.4. Computing the OLFC policy when $\mathbf{s}_j \geq \mathbf{0}$ is not imposed. If we drop the non-negativity constraints on the $\mathbf{s}_j$'s, then at each time t the OLFC policy can be computed analytically and very quickly using dynamic programming. Recall the agent assumes that $\mathbf{\Sigma}_j = \mathbf{\Sigma}_t$ for all $j \geq t$ so that he assumes

$$\mathbb{E}_j[\exp(\gamma \mathbf{r}_{j+1}' \mathbf{w})] = \exp\left(\frac{1}{2}\gamma^2 \mathbf{w}' \mathbf{\Sigma}_t \mathbf{w}\right) \tag{A-20}$$

for all $j \geq t$ and all constant vectors, $\mathbf{w}$. We use $V_j^{ol}(\cdot)$ for $j \geq t$ to denote the expected utility of trading from time j onwards as perceived by the agent at time t. The dynamics of section 2.1, but now taking $\mathbf{\Sigma}_j = \mathbf{\Sigma}_t$ and replacing the $\mathbf{A}_j$'s and $\mathbf{B}_j$'s by their time-t conditional expectations for all $j \geq t$, imply that

$$\begin{aligned} V_T^{ol}(\tilde{\mathbf{p}}_T, \mathbf{w}_T) &= \exp\left(\gamma(\tilde{\mathbf{p}}_T + \mathbb{E}_t[\mathbf{A}_T]\mathbf{w}_T + \mathbb{E}_t[\mathbf{B}_T]\mathbf{w}_T)'\mathbf{w}_T\right) \\ &= \exp\left(\gamma\left(\frac{1}{2}\mathbf{w}_T'\mathbf{G}_T\mathbf{w}_T + \tilde{\mathbf{p}}_T'\mathbf{w}_T\right)\right), \end{aligned}$$

where $\mathbf{G}_T := \mathbb{E}_t[\mathbf{A}_T + \mathbf{A}_T' + \mathbf{B}_T + \mathbf{B}_T']$. At time $T-1$ the agent needs to solve

$$V_{T-1}^{ol}(\tilde{\mathbf{p}}_{T-1}, \mathbf{w}_{T-1}) = \min_{\mathbf{s}_{T-1}} \; \mathbb{E}_{T-1}[\exp(\gamma \mathbf{p}_{T-1}' \mathbf{s}_{T-1}) V_T^{ol}(\tilde{\mathbf{p}}_T, \; \mathbf{w}_T)],$$

and using (A-20) this problem reduces to

$$\min_{\mathbf{s}_{T-1}} \exp\Bigg(\gamma\Bigg(\frac{1}{2}\mathbf{s}'_{T-1}\mathbf{N}_{ss,T-1}\mathbf{s}_{T-1} - (\mathbf{N}_{ws,T-1}\mathbf{w}_{T-1})'\mathbf{s}_{T-1} + \frac{1}{2}\mathbf{w}'_{T-1}(\mathbf{G}_T + \gamma\Sigma_t)\mathbf{w}_{T-1} + \tilde{\mathbf{p}}'_{T-1}\mathbf{w}_{T-1}\Bigg)\Bigg)$$

$$\equiv \min_{\mathbf{s}_{T-1}} \frac{1}{2}\mathbf{s}'_{T-1}\mathbf{N}_{ss,T-1}\mathbf{s}_{T-1} - (\mathbf{N}_{ws,T-1}\mathbf{w}_{T-1})'\mathbf{s}_{T-1} + \frac{1}{2}\mathbf{w}'_{T-1}(\mathbf{G}_T + \gamma\Sigma_t)\mathbf{w}_{T-1} + \tilde{\mathbf{p}}'_{T-1}\mathbf{w}_{T-1}, \tag{A-21}$$

where

$$\mathbf{N}_{ss,T-1} := \mathbf{G}_T + \gamma\boldsymbol{\Sigma}_t + \mathbb{E}_t[\mathbf{B}_{T-1} + \mathbf{B}'_{T-1}],$$
$$\mathbf{N}_{ws,T-1} := \mathbf{G}_T + \gamma\boldsymbol{\Sigma}_t - \mathbb{E}_t[\mathbf{A}'_{T-1}].$$

Assuming $\mathbf{N}_{ss,T-1}$ is positive definite, then the optimal solution to (A-21) is

$$\mathbf{s}^*_{T-1} = \mathbf{N}^{-1}_{ss,T-1}\mathbf{N}_{ws,T-1}\mathbf{w}_{T-1}$$

with corresponding value function

$$V^{ol}_{T-1}(\tilde{\mathbf{p}}_{T-1}, \mathbf{w}_{T-1}) \;=\; \exp\left(\gamma\left(\frac{1}{2}\mathbf{w}'_{T-1}\mathbf{G}_{T-1}\mathbf{w}_{T-1} + \tilde{\mathbf{p}}'_{T-1}\mathbf{w}_{T-1}\right)\right),$$

where

$$\mathbf{G}_{T-1} \;:=\; \mathbf{G}_T + \gamma\boldsymbol{\Sigma}_t - \mathbf{N}'_{ws,T-1}\mathbf{N}^{-1}_{ss,T-1}\mathbf{N}_{ws,T-1}.$$

Continuing backwards in this manner we find for $j = T-2, \ldots, t$ that

$$\begin{aligned} V^{ol}_j(\tilde{\mathbf{p}}_j, \mathbf{w}_j) &= \min_{\mathbf{s}_j} \mathbb{E}_j[\exp(\gamma\mathbf{p}'_j\mathbf{s}_j)V_{j+1}(\tilde{\mathbf{p}}_{j+1}, \mathbf{w}_{j+1})] \\ &= \min_{\mathbf{s}_j} \exp\left(\gamma\left(\frac{1}{2}\mathbf{s}'_j\mathbf{N}_{ss,j}\mathbf{s}_j - (\mathbf{N}_{ws,j}\mathbf{w}_j)'\mathbf{s}_j + \frac{1}{2}\mathbf{w}'_j(\mathbf{G}_{j+1} + \gamma\Sigma_t)\mathbf{w}_j + \tilde{\mathbf{p}}'_j\mathbf{w}_j\right)\right) \\ &\equiv \min_{\mathbf{s}_j} \frac{1}{2}\mathbf{s}'_j\mathbf{N}_{ss,j}\mathbf{s}_j - (\mathbf{N}_{ws,j}\mathbf{w}_j)'\mathbf{s}_j + \frac{1}{2}\mathbf{w}'_j(\mathbf{G}_{j+1} + \gamma\Sigma_t)\mathbf{w}_j + \tilde{\mathbf{p}}'_j\mathbf{w}_j, \end{aligned} \tag{A-22}$$

where

$$\mathbf{N}_{ss,j} := \mathbf{G}_{j+1} + \gamma\boldsymbol{\Sigma}_t + \mathbb{E}_t[\mathbf{B}_j + \mathbf{B}'_j],$$
$$\mathbf{N}_{ws,j} := \mathbf{G}_{j+1} + \gamma\boldsymbol{\Sigma}_t - \mathbb{E}_t[\mathbf{A}'_j].$$

Again assuming $\mathbf{N}_{ss,j}$ is positive definite, the optimal value of $\mathbf{s}_j$ in (A-22) is given by

$$\mathbf{s}^*_j = \mathbf{N}^{-1}_{ss,j}\mathbf{N}_{ws,j}\mathbf{w}_j$$

with corresponding value function

$$V_j^{ol}(\tilde{\mathbf{p}}_j, \mathbf{w}_j) = \exp\left(\gamma\left(\frac{1}{2}\mathbf{w}_j'\mathbf{G}_j\mathbf{w}_j + \tilde{\mathbf{p}}_j'\mathbf{w}_j\right)\right), \tag{A-23}$$

where

$$\mathbf{G}_j := \mathbf{G}_{j+1} + \gamma\mathbf{\Sigma}_t - \mathbf{N}_{ws,j}'\mathbf{N}_{ss,j}^{-1}\mathbf{N}_{ws,j}.$$

The OLFC policy implements $\mathbf{s}_t^*$ at time t, and we note that $\mathbf{s}_t^*$ depends on $\mathbf{w}_t$ only, and not on the price, $\tilde{\mathbf{p}}_t$. The OLFC policy is path-dependent, however, because of the dependence of $\mathbf{N}_{ss,t}$ and $\mathbf{N}_{ws,t}$ on $\mathbf{\Sigma}_t$, $\mathbf{A}_t$, and $\mathbf{B}_t$, which in general are stochastic. As mentioned in section 5, it is also possible to model return predictability via linear state variable dynamics and still compute the OLFC value function and policy analytically. This would induce an explicit path dependence of $\mathbf{s}_t^*$ on these state variables.

Note that the DP formulation here is equivalent to the static problem formulation in Appendix A.3 if we drop the nonnegativity constraints on the $\mathbf{s}_t^*$'s there. It follows then that we can check the positive definiteness of the $\mathbf{N}_{ss,j}$'s for $j \geq t$ by confirming that $\mathbf{Q}_{OLFC,t}+\mathbf{Q}_{\mathbf{\Sigma},t}$ is positive definite. We can do the latter using, for example, the results of Appendix C.1.

A.5. Using control variates to estimate the primal bounds. In order to reduce the number of Monte Carlo paths that we used for estimating the primal bound, V_{ub}, we considered two possible control variates. First note that if a trading sequence $\mathbf{s} := [\mathbf{s}_0' \dots \mathbf{s}_T']'$ is deterministic, then the expected execution cost can be computed analytically under the price-impact model of section 2. In particular, it is easy to check that

$$\begin{aligned}\mathbb{E}_0\left[\sum_{t=0}^{T}\mathbf{p}_t'\mathbf{s}_t\right] &= \mathbb{E}_0\left[\sum_{t=0}^{T}\left(\tilde{\mathbf{p}}_0 + \sum_{i=0}^{t-1}\mathbf{A}_i\mathbf{s}_i + \sum_{i=1}^{t}\mathbf{r}_i + \mathbf{A}_t\mathbf{s}_t + \mathbf{B}_t\mathbf{s}_t\right)'\mathbf{s}_t\right]\\ &= \frac{1}{2}\mathbf{s}'\mathbb{E}_0[\mathbf{Q}]\mathbf{s} + \tilde{\mathbf{p}}_0'\mathbf{w}_0,\end{aligned} \tag{A-24}$$

where $\mathbf{Q}$ is as defined in Appendix C.1. If we can evaluate $\mathbb{E}_0[\mathbf{Q}]$, then by (A-24) we can use $\sum_{t=0}^{T}\mathbf{p}_t'\mathbf{s}_t$ as a control variate. However, as the level of risk aversion (as measured by γ) increases, the variance reduction that it achieves will not be as effective.

The expected utility of a deterministic trading sequence cannot be computed analytically when there are stochastic variance-covariance dynamics and stochastic linear price-impact dynamics. However, if we assume the volatility, $\mathbf{\Sigma}$, is constant, and $\mathbf{A}_t$ and $\mathbf{B}_t$ evolve deterministically according to their time $t = 0$ conditional expectations, $\mathbb{E}_0[\mathbf{A}_t]$ and $\mathbb{E}_0[\mathbf{B}_t]$, then it is easy to check that

$$\begin{aligned}\mathbb{E}_0\left[\exp\left(\gamma\sum_{t=0}^{T}\mathbf{p}_t'\mathbf{s}_t\right)\right] &= \mathbb{E}_0\left[\exp\left(\gamma\sum_{t=0}^{T}\left(\tilde{\mathbf{p}}_0 + \sum_{i=0}^{t-1}\mathbb{E}_0[\mathbf{A}_i]\mathbf{s}_i + \sum_{i=1}^{t}\mathbf{r}_i\right.\right.\right.\\ &\qquad\left.\left.\left. + \mathbb{E}_0[\mathbf{A}_t]\mathbf{s}_t + \mathbb{E}_0[\mathbf{B}_t]\mathbf{s}_t\right)'\mathbf{s}_t\right)\right]\\ &= \exp(\gamma\tilde{\mathbf{p}}_0'\mathbf{w}_0)\exp\left(\frac{\gamma}{2}\mathbf{s}'(\mathbf{Q}_{OLFC,0} + \mathbf{Q}_{\mathbf{\Sigma},0})\mathbf{s}\right).\end{aligned} \tag{A-25}$$

We can therefore use $\exp(\gamma \sum_{t=0}^{T} \mathbf{p}_t' \mathbf{s}_t)$ and (A-25) as a control variate. Note that we can do this even when each of $\mathbf{\Sigma}_t$, $\mathbf{A}_t$, and $\mathbf{B}_t$ evolves stochastically in the "true" model by simulating in parallel the model with deterministic price impacts and variance-covariance dynamics that yields (A-25). In our numerical experiments we took $\mathbf{\Sigma} = \mathbf{\Sigma}_0$ and took $\mathbf{s}$ to be the optimal (deterministic) policy as calculated in Appendix A.4 with $t = 0$. We generally find (A-25) to be much more effective than (A-24), and it resulted in a variance reduction for the primal bounds on the order of 75% to 95% depending on the value of γ.

It is worth mentioning that these control variates provided little benefit when estimating the dual bounds. Since the calculation of the primal bound was the computational bottleneck, however, there was no need to construct good control variates for the dual problem.

Appendix B. Review of duality based on information relaxations. We begin with a general finite-horizon discrete-time DP with a probability space, $(\Omega, \mathcal{F}, \mathbb{P})$. Time is indexed by the set $\mathcal{T} := \{0, \ldots, T\}$, and the evolution of information is described by the filtration $\mathbb{F} = \{\mathcal{F}_0, \ldots, \mathcal{F}_T\}$ with $\mathcal{F} = \mathcal{F}_T$. We make the usual assumption that $\mathcal{F}_0 = \{\emptyset, \Omega\}$ so that the decision-maker starts out with no information regarding the outcome of uncertainty. There is a state vector, $x_t \in \mathbb{R}^n$, whose dynamics satisfy

$$x_{t+1} = f_t(x_t, u_t, w_{t+1}), \quad t = 0, \ldots, T-1, \tag{B-26}$$

where $u_t \in U_t(x_t) \subseteq \mathbb{R}^m$ is the control taken at time t, w_{t+1} is an $\mathcal{F}_{t+1}$-measurable random disturbance, and $U_t(x_t)$ is the feasible control set at time t. A *feasible* policy, $u := (u_0, \ldots, u_T)$ is one where each individual action satisfies $u_t \in U_t(x_t)$ for all t. We let $\mathcal{U}$ denote the set of such policies. A feasible *adapted* policy is a feasible policy that is $\mathcal{F}_t$-adapted. We let $\mathcal{U}_{\mathbb{F}}$ denote the set of all such $\mathcal{F}_t$-adapted policies. For example, in the context of the portfolio execution problem of this paper, a feasible but not $\mathcal{F}_t$-adapted strategy would be an execution schedule that satisfies the nonnegativity constraints in all time periods but where the number of shares purchased in a given period is allowed to depend on prices in later periods. The objective is to select a feasible adapted policy, u, to minimize the total loss,

$$g(u) := \sum_{t=0}^{T} g_t(x_t, u_t),$$

where we assume without loss of generality that each $g_t(x_t, u_t)$ is $\mathcal{F}_t$-measurable. In particular, the decision-maker's problem is then given by

$$V_0^*(x_0) := \inf_{u \in \mathcal{U}_{\mathbb{F}}} \mathbb{E}_0 \left[\sum_{t=0}^{T} g_t(x_t, u_t) \right], \tag{B-27}$$

where the expectation in (B-27) is taken over the set of possible outcomes, $w = (w_1, \ldots, w_T) \in \Omega$. Letting V_t^* denote the time t value function for the problem (B-27), the associated dynamic programming recursion is given by

$$V_t^*(x_t) := \inf_{u_t \in U_t(x_t)} \left\{ g_t(x_t, u_t) + \mathbb{E}_t \left[V_{t+1}^* \left(f_t(x_t, u_t, w_{t+1}) \right) \right] \right\}, \quad t = 0, \ldots, T, \tag{B-28}$$

with the understanding that $V^*_{T+1} \equiv 0$ (and therefore does not depend on w_{T+1}, which is undefined). In practice, of course, it is often too difficult or time-consuming to perform the iteration in (B-28). This can occur, for example, if the state vector, x_t, is high-dimensional or if the constraints imposed on the controls are too complex or difficult to handle. In such circumstances, we must be satisfied with suboptimal policies. These policies are generally easy to simulate and can therefore be used to construct unbiased upper bounds on the optimal value function, $V_0^*(x_0)$. Duality methods based on information relaxations can be used to construct lower bounds on $V_0^*(x_0)$.

We will now briefly describe the theory behind these duality methods. Brown, Smith, and Sun [13] should be consulted for further details. We omit most of the technical details, and we will consider only *perfect information* relaxations since these relaxations are usually most useful in practice and are all that we will require in this paper. Let $\mathcal{S}$ denote the space of real-valued measurable functions that are defined on the state space $\mathbb{R}^n$. We now define an operator Δ that maps $\mathcal{S}$ to the space of real-valued measurable functions on $\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^m$ according to

$$\text{(B-29)} \qquad (\Delta V_t)(x_t, x_{t-1}, u_{t-1}) := V_t(x_t) \; - \; \mathbb{E}\left[V_t(x_t) \,|\, x_{t-1}, u_{t-1}\right].$$

Loosely speaking, Δ is an operator on (approximate) value functions. Note in particular that $\mathbb{E}_0\left[(\Delta V_t)(x_t, x_{t-1}, u_{t-1})\right] = 0$ for all integrable V_t. Let $\mathcal{D}$ be the space of real-valued functions on $\mathbb{R}^n \times \mathcal{T}$ such that if $V \in \mathcal{D}$, then $V_t := V(\cdot, t)$ is measurable and $\mathbb{E}_0\left[|V_t(x_t)|\right] < \infty$ for all $t \in \mathcal{T}$ and all feasible policies $u \in \mathcal{U}$. We now define an operator $F : \mathcal{D} \to \mathcal{S}$ according to

(B-30)

$$FV(x) \; := \; \mathbb{E}_0\left[\inf_{u \in \mathcal{U}} \left\{ g_0(x_0, u_0) \; + \; \sum_{t=1}^{T} \left(g_t(x_t, u_t) \; - \; (\Delta V_t)(x_t, x_{t-1}, u_{t-1})\right)\right\} \,\middle|\, x_0 = x\right].$$

Note that the infimum in (B-30) is over the space $\mathcal{U}$ of feasible policies and not the space $\mathcal{U}_{\mathbb{F}}$ of feasible adapted policies. Our first result is weak duality.

Theorem B.1 (weak duality). *$V_0^*(x_0) \; \geq \; FV(x_0)$ for all $V \in \mathcal{D}$.*

Proof. Using the definition of $V_0^*(x_0)$ in (B-27) and the fact that the (ΔV_t)'s have zero mean we have

(B-31)

$$\begin{aligned}
V_0^*(x) &= \inf_{u \in \mathcal{U}_{\mathbb{F}}} \mathbb{E}_0\left[\sum_{t=0}^{T} g_t(x_t, u_t) \; - \; \sum_{t=1}^{T} (\Delta V_t)(x_t, x_{t-1}, u_{t-1}) \,\middle|\, x_0 = x\right] \\
&= \inf_{u \in \mathcal{U}_{\mathbb{F}}} \mathbb{E}_0\left[g_0(x_0, u_0) \; + \; \sum_{t=1}^{T} \left(g_t(x_t, u_t) \; - \; (\Delta V_t)(x_t, x_{t-1}, u_{t-1})\right) \,\middle|\, x_0 = x\right] \\
&\geq \mathbb{E}_0\left[\inf_{u \in \mathcal{U}} \left\{g_0(x_0, u_0) \; + \; \sum_{t=1}^{T} \left(g_t(x_t, u_t) \; - \; (\Delta V_t)(x_t, x_{t-1}, u_{t-1})\right)\right\} \,\middle|\, x_0 = x\right] \\
\text{(B-32)} \qquad &= FV(x_0). \qquad \blacksquare
\end{aligned}$$

Theorem B.1 suggests that we can compute a lower bound on $V_0^*(x_0)$ by evaluating $FV(x_0)$ for any $V \in \mathcal{D}$. It gives no guidance, however, on how to choose V so that the lower bound is as large as possible. In fact we can formulate the following dual problem:

$$\sup_{V \in \mathcal{D}} FV(x_0). \tag{B-33}$$

Theorem B.2 is a strong duality result which states that the dual problem is solved by taking $V = V^*$ and that there is no duality gap between the primal DP and the dual problem.

Theorem B.2 (strong duality). *For all x, $V_0^*(x) = FV^*(x)$.*

Proof. By weak duality we need only show that $FV^*(x) \geq V_0^*(x)$. We have

$$\begin{aligned}
FV^*(x) &= \mathbb{E}_0\left[\inf_{u \in \mathcal{U}} \left\{ g_0(x_0, u_0) \ + \ \sum_{t=1}^{T} (g_t(x_t, u_t) \ - \ (\Delta V_t^*)(x_t, x_{t-1}, u_{t-1})) \right\} \middle| \ x_0 = x \right] \\
&= \mathbb{E}_0\left[\inf_{u \in \mathcal{U}} \left\{ V_0^*(x_0) \ + \ \sum_{t=1}^{T-1} \left[g_t(x_t, u_t) \ + \ \mathbb{E}\left[V_{t+1}^*(x_{t+1}) \,|\, x_t, u_t\right] - V_t^*(x_t)\right] \right.\right. \\
&\qquad \left.\left. + \ g_T(x_T, u_T) - V_T^*(x_T) \right\} \middle| \ x_0 = x \right] \\
&\geq V_0^*(x), \qquad \text{(B-34)}
\end{aligned}$$

where the inequality in (B-34) follows because of (B-28) and since $V_T^*(x_T) = g_T(x_T, u_T)$. ■

Strong duality suggests that we might be able to obtain good dual bounds on the optimal value function, $V_0^*(x)$, if we can find $V \approx V^*$ and then compute $FV(x)$. The numerical experiments in this paper and in the literature cited in section 1 support this claim.

Remark 1. *The portfolio execution problem that we consider in this paper has an objective function of the form $\mathbb{E}_0[\exp(\gamma \sum_{t=0}^{T} \mathbf{p}_t' \mathbf{s}_t)]$. This is not in the form $\mathbb{E}_0[\sum_{t=0}^{T} g_t(x_t, u_t)]$ as we assumed in (B-27), but this does not present any problems because we can introduce an additional state variable, say z_t, with dynamics $z_{t+1} = z_t + \mathbf{p}_{t+1}' \mathbf{s}_{t+1}$ for $t = 0, \ldots, T-1$ with $z_0 := \mathbf{p}_0' \mathbf{s}_0$. We can then take $g_0 \equiv g_1 \equiv \cdots \equiv g_{T-1} \equiv 0$ and $g_T := \exp(\gamma z_T)$. The portfolio execution problem now has an objective function of the appropriate form. We also note that given any policy, say θ, the corresponding time $t+1$ value function, V_{t+1}^θ, can be written as $V_{t+1}^\theta = \exp(\gamma \sum_{j=0}^{t} \mathbf{p}_j' \mathbf{s}_j)\, \mathbb{E}_{t+1}^\theta[\exp(\gamma \sum_{j=t+1}^{T} \mathbf{p}_j' \mathbf{s}_j)]$, where the expectation $\mathbb{E}_{t+1}^\theta[\cdot]$ is with respect to the probability measure induced by the policy θ. This also explains why we want to include the first term on the right-hand side of (14).*

Appendix C. The dual problem for the portfolio execution problem. We define $f(\mathbf{s}) := \exp(\gamma \sum_{t=0}^{T} \mathbf{p}_t' \mathbf{s}_t)$ from section 3.1. We can substitute for $\mathbf{p}_t$ using (A-14) to obtain

$$\begin{aligned}
f(\mathbf{s}) &= \exp\left(\gamma \sum_{t=0}^{T} \left(\tilde{\mathbf{p}}_0 + \sum_{i=0}^{t-1} \mathbf{A}_i \mathbf{s}_i + \sum_{i=1}^{t} \mathbf{r}_i + \mathbf{A}_t \mathbf{s}_t + \mathbf{B}_t \mathbf{s}_t\right)' \mathbf{s}_t\right) && \text{(C-35)} \\
&= \exp(\gamma \tilde{\mathbf{p}}_0' \mathbf{w}_0)\ \exp\left(\gamma \left(\frac{1}{2} \mathbf{s}' \mathbf{Q} \mathbf{s} + \mathbf{c}_{p,0}' \mathbf{s}\right)\right), && \text{(C-36)}
\end{aligned}$$

where $\mathbf{Q}$ is an $n(T+1)\times n(T+1)$ symmetric matrix given by

$$\text{(C-37)}\qquad \mathbf{Q}=\begin{bmatrix}\mathbf{A}_0+\mathbf{A}_0'+\mathbf{B}_0+\mathbf{B}_0' & \mathbf{A}_0' & \dots & \mathbf{A}_0'\\ \mathbf{A}_0 & \mathbf{A}_1+\mathbf{A}_1'+\mathbf{B}_1+\mathbf{B}_1' & \dots & \mathbf{A}_1'\\ \dots & \dots & \dots & \dots\\ \mathbf{A}_0 & \mathbf{A}_1 & \dots & \mathbf{A}_T+\mathbf{A}_T'+\mathbf{B}_T+\mathbf{B}_T'\end{bmatrix}$$

and $\mathbf{c}_{p,0}$ is a $n(T+1)\times 1$ vector given by

$$\mathbf{c}_{p,0}=\begin{bmatrix}\mathbf{0}\\ \mathbf{r}_1\\ \mathbf{r}_1+\mathbf{r}_2\\ \dots\\ \mathbf{r}_1+\mathbf{r}_2+\dots+\mathbf{r}_t\\ \dots\\ \mathbf{r}_1+\mathbf{r}_2+\dots+\mathbf{r}_t+\dots+\mathbf{r}_T\end{bmatrix}.$$

Using (C-36) we see that the gradient vector and Hessian matrix of f satisfy

$$\begin{aligned}\nabla f(\mathbf{s})&=\gamma f(\mathbf{s})(\mathbf{Q}\mathbf{s}+\mathbf{c}_{p,0}),\\ Hf(\mathbf{s})&=\gamma f(\mathbf{s})\Big(\mathbf{Q}+\gamma(\mathbf{Q}\mathbf{s}+\mathbf{c}_{p,0})(\mathbf{Q}\mathbf{s}+\mathbf{c}_{p,0})'\Big).\end{aligned}$$

If $\mathbf{Q}$ is positive definite, then the Hessian matrix $H(f(\mathbf{s}))$ is also positive definite and $f(\mathbf{s})$ is therefore convex. This follows because $\gamma>0$, $f(\mathbf{s})>0$ for all $\mathbf{s}$ and because $(\mathbf{Q}\mathbf{s}+\mathbf{c}_{p,0})(\mathbf{Q}\mathbf{s}+\mathbf{c}_{p,0})'$, as the outerproduct of a column vector, is positive semidefinite for all $\mathbf{s}$. It therefore follows that if $\mathbf{Q}$ is positive definite, then *all* dual problem instances will be convex. In Appendix C.1, we provide some sufficient conditions that ensure $\mathbf{Q}$ will be positive definite.

C.1. Conditions that guarantee the positive definiteness of $\mathbf{Q}$ and $\mathbf{Q}_{OLFC,t}$. The positive definiteness of $\mathbf{Q}$ depends on the dynamics of $\mathbf{A}_t$ and $\mathbf{B}_t$. Here we provide two sufficient conditions that guarantee this.

(i) The permanent price impact is constant over time $\mathbf{A}_t=\mathbf{A}$, and $\mathbf{A}+\mathbf{A}'$ is positive definite.
(ii) $\mathbf{B}_t+\mathbf{B}_t'$ are positive semidefinite.

Assuming that conditions (i) and (ii) hold we can then write (C-37) as

(C-38)
$$\mathbf{Q}=\begin{bmatrix}\frac{\mathbf{A}+\mathbf{A}'}{2} & \mathbf{A}' & \dots & \mathbf{A}'\\ \mathbf{A} & \frac{\mathbf{A}+\mathbf{A}'}{2} & \dots & \mathbf{A}'\\ \dots & \dots & \dots & \dots\\ \mathbf{A} & \mathbf{A} & \dots & \frac{\mathbf{A}+\mathbf{A}'}{2}\end{bmatrix}+\begin{bmatrix}\frac{\mathbf{A}+\mathbf{A}'}{2}+\mathbf{B}_0+\mathbf{B}_0' & & & \\ & \frac{\mathbf{A}+\mathbf{A}'}{2}+\mathbf{B}_1+\mathbf{B}_1' & & \\ & & \dots & \\ & & & \frac{\mathbf{A}+\mathbf{A}'}{2}+\mathbf{B}_T+\mathbf{B}_T'\end{bmatrix}.$$

The second matrix on the right-hand side of (C-38) is positive definite since it is block diagonal and each $\frac{1}{2}(\mathbf{A}+\mathbf{A}')+\mathbf{B}_t+\mathbf{B}_t'$ is positive definite. The first matrix is positive semidefinite since for any given vector $\mathbf{z} = [\mathbf{z}_1 \ \dots \ \mathbf{z}_T]'$, where each $\mathbf{z}_i$ is an $n \times 1$ vector, we have

(C-39)
$$[\mathbf{z}_1 \ \dots \ \mathbf{z}_T]' \begin{bmatrix} \frac{\mathbf{A}+\mathbf{A}'}{2} & \mathbf{A}' & \dots & \mathbf{A}' \\ \mathbf{A} & \frac{\mathbf{A}+\mathbf{A}'}{2} & \dots & \mathbf{A}' \\ \dots & \dots & \dots & \dots \\ \mathbf{A} & \mathbf{A} & \dots & \frac{\mathbf{A}+\mathbf{A}'}{2} \end{bmatrix} \begin{bmatrix} \mathbf{z}_1 \\ \mathbf{z}_2 \\ \vdots \\ \mathbf{z}_T \end{bmatrix} = (\mathbf{z}_1+\mathbf{z}_2+\dots+\mathbf{z}_T)' \frac{\mathbf{A}+\mathbf{A}'}{2} (\mathbf{z}_1+\mathbf{z}_2+\dots+\mathbf{z}_T),$$

which is nonnegative since $\frac{\mathbf{A}+\mathbf{A}'}{2}$ is positive definite. This implies $\mathbf{Q}$ is positive definite.

$\mathbf{Q}_{OLFC,t}$ in (A-19) has a similar structure to $\mathbf{Q}$. Given that the $\mathbf{A}_t$'s are constant we have $\mathbb{E}_t[\mathbf{A}_j] = \mathbf{A}$, and then

$$\mathbf{Q}_{OLFC,t} = \begin{bmatrix} \mathbf{A}+\mathbf{A}'+\mathbb{E}_t[\mathbf{B}_t+\mathbf{B}_t'] & \mathbf{A}' & \dots & \mathbf{A}' \\ \mathbf{A} & \mathbf{A}+\mathbf{A}'+\mathbb{E}_t[\mathbf{B}_{t+1}+\mathbf{B}_{t+1}'] & \dots & \mathbf{A}' \\ \dots & \dots & \dots & \dots \\ \mathbf{A} & \mathbf{A} & \dots & \mathbf{A}+\mathbf{A}'+\mathbb{E}_t[\mathbf{B}_T+\mathbf{B}_T'] \end{bmatrix}.$$

If the dynamics of $\mathbf{B}_t$ ensure that $\mathbb{E}_t[\mathbf{B}_j+\mathbf{B}_j']$ are positive semidefinite, $\mathbf{Q}_{OLFC,t}$ is guaranteed to be positive definite by the same argument we used for (C-38).

C.2. Computing dual penalties. We saw in section 3.1 that we would like to take $\tilde{V}_t$ (required for (12)) to be a linearized version of an approximate value function, $\hat{V}_t$, as given by (15). However, if we construct $\hat{V}_t$ using the OLFC value function, V_t^{ol}, and take $\hat{V}_{t+1}(\mathbf{s}_{0:t}) = \exp(\gamma \sum_{j=0}^{t} \mathbf{p}_j'\mathbf{s}_j)V_{t+1}^{ol}$, then we cannot compute $\mathbb{E}_t[\tilde{V}_{t+1}(\mathbf{s}_{0:t})]$ analytically. This is because V_{t+1}^{ol} is calculated under the assumption that the conditional covariance matrix remains constant at $\boldsymbol{\Sigma}_{t+1}$ and that the price-impact matrices for all $j \geq t+1$ are $\mathbb{E}_{t+1}[\mathbf{A}_j]$'s and $\mathbb{E}_{t+1}[\mathbf{B}_j]$'s. None of these terms is adapted to $\mathcal{F}_t$, and so computing $\mathbb{E}_t[\tilde{V}_{t+1}(\mathbf{s}_{0:t})]$ analytically is not possible in general. We could in theory use the Monte Carlo approach of section 5 to overcome this problem, but in this particular case it would be prohibitively expensive to do so. This is because we would need to solve an OLFC optimization problem at each simulated point at time $t+1$.

Instead we simply modify the assumptions of the OLFC policy and assume that at time $t+1$ the conditional covariance matrix remains constant at $\boldsymbol{\Sigma}_t$ (rather than $\boldsymbol{\Sigma}_{t+1}$) and that the price-impact matrices are given by $\mathbb{E}_t[\mathbf{A}_j]$ and $\mathbb{E}_t[\mathbf{B}_j]$ (rather than $\mathbb{E}_{t+1}[\mathbf{A}_j]$ and $\mathbb{E}_{t+1}[\mathbf{B}_j]$) for all $j \geq t+1$. Using precisely the same DP approach of Appendix A.4 we obtain the *modified* OLFC value function

(C-40)
$$V_{t+1}^{mol} = \exp\left(\gamma\left(\frac{1}{2}\mathbf{w}_{t+1}'\tilde{\mathbf{G}}_{t+1}\mathbf{w}_{t+1} + \tilde{\mathbf{p}}_{t+1}'\mathbf{w}_{t+1}\right)\right),$$

where the $\tilde{\mathbf{G}}_t$'s are the analogue of the $\mathbf{G}_t$'s in Appendix A.4. The important feature of (C-40) is that $\tilde{\mathbf{G}}_{t+1} \in \mathcal{F}_t$ so that $\mathbb{E}_t[V_{t+1}^{mol}]$ can be computed in closed form.

C.3. Solving dual problem instances. Recall that we define $f(\mathbf{s}) := \exp(\gamma \sum_{t=0}^{T} \mathbf{p}_t'\mathbf{s}_t)$. From (12) and (15) it follows that each dual problem instance that we need to solve has an objective function equal to $f(\mathbf{s})$ plus a linear function of $\mathbf{s}$. While all of the problem instances in our numerical experiments will be convex, the exponential operator combined with the linear term can be a source of difficulty. In addition, the permanent price impact implies that $\mathbf{p}_t$ is a function of $\mathbf{s}_{0:t-1}$ so that we cannot reduce the problem to a separable convex problem which is easily solved. Instead we solve each dual problem instance by using the following algorithm:

1. Choose a starting point, $\hat{\mathbf{s}}$.
2. Approximate $f(\mathbf{s})$ with a second order Taylor expansion about $\hat{\mathbf{s}}$ to obtain
$$\hat{f}(\mathbf{s}) \;:=\; f(\hat{\mathbf{s}}) + \nabla f(\hat{\mathbf{s}})'(\mathbf{s}-\hat{\mathbf{s}}) + \frac{1}{2}(\mathbf{s}-\hat{\mathbf{s}})'\mathrm{H}f(\hat{\mathbf{s}})(\mathbf{s}-\hat{\mathbf{s}}), \tag{C-41}$$
where $\nabla f(\hat{\mathbf{s}})$ and $\mathrm{H}f(\hat{\mathbf{s}})$ are, respectively, the gradient vector and Hessian matrix of f evaluated at $\hat{\mathbf{s}}$.
3. Solve
$$\min_{\mathbf{s}\in\mathbb{S}} \quad \hat{f}(\mathbf{s}) + \sum_{t=0}^{T-1}\left(\mathbb{E}_t[\tilde{V}_{t+1}(\mathbf{s}_{0:t})] - \tilde{V}_{t+1}(\mathbf{s}_{0:t})\right), \tag{C-42}$$
which is a constrained convex quadratic programming problem and therefore easy to solve. Let $\mathbf{s}^{opt}$ be the optimal solution.
4. Evaluate the objective value in (12) at $\mathbf{s}^{opt}$, and stop if we have converged to within a given error tolerance. Otherwise set $\hat{\mathbf{s}} = \mathbf{s}^{opt}$ and return to step 2.

In our numerical experiments we use an absolute error tolerance of 10^{-5}. Depending on the level of risk aversion, γ, this corresponds to a relative error tolerance between 10^{-4} and 10^{-5}. Typically we find that convergence occurs after just two or three iterations.

Appendix D. The model of Bertsimas, Hummel, and Lo. Bertsimas, Hummel, and Lo [9] (BHL) was one of the earliest papers to consider the portfolio execution problem. We use the calibrated model parameters of BHL to (i) investigate their conjecture that an OLFC policy should be close to optimal and (ii) provide a simple demonstration of how duality can be used to determine *in advance* whether or not a particular market feature, in this case cross-price impacts, is worth accounting for in a portfolio execution policy. In particular, by considering how small the calibrated cross-price-impact parameters are in BHL, it is reasonable to conjecture that simply following the optimal single stock execution policies should be close to optimal. We confirm this by using the single-stock value functions to construct a dual penalty which we then use to bound how far the single-stock policy is from optimality. This example therefore serves as an illustration of how a policy and the dual penalty from a given model can be used to determine *in advance* whether or not a more complicated model even needs to be considered. Note that we claim only to show that cross-price-impact parameters are insignificant *within* the model of BHL and do not make this claim more generally.

D.1. Model description. The linear percentage price-impact model of BHL has dynamics

$$\mathbf{p}_t = \tilde{\mathbf{p}}_t + \boldsymbol{\delta}_t, \tag{D-43}$$
$$\boldsymbol{\delta}_t = \tilde{\mathbf{P}}_t(\mathbf{A}\tilde{\mathbf{P}}_t\mathbf{s}_t + \mathbf{B}_t\mathbf{x}_t), \tag{D-44}$$
$$\tilde{\mathbf{p}}_{t+1} = \exp(\mathrm{Diag}(\boldsymbol{\epsilon}_{t+1}))\tilde{\mathbf{p}}_t, \tag{D-45}$$
$$\mathbf{x}_{t+1} = \mathbf{C}\mathbf{x}_t + \boldsymbol{\eta}_{t+1}, \tag{D-46}$$
$$\mathbf{w}_{t+1} = \mathbf{w}_t - \mathbf{s}_t, \tag{D-47}$$

where $\tilde{\mathbf{P}}_t := \mathrm{Diag}[\tilde{\mathbf{p}}_t]$, a diagonal matrix with $\tilde{\mathbf{p}}_t$ on the diagonal. The execution price $\mathbf{p}_t$ is the sum of the no-impact price, $\tilde{\mathbf{p}}_t$, and the price impact, $\boldsymbol{\delta}_t$. $\tilde{\mathbf{p}}_t$ follows a vector-geometric Brownian motion, where the $\boldsymbol{\epsilon}_t$'s are IID normal with mean vector $\boldsymbol{\mu}_{\boldsymbol{\epsilon}}$ and covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{\epsilon}}$. The price impact, $\boldsymbol{\delta}_t$, is temporary and assumes the percentage impact is a linear function of the dollar values, $\tilde{\mathbf{P}}_t\mathbf{s}_t$, and an information vector, $\mathbf{x}_t$, that represents market or private information and is assumed to follow a vector autoregressive process with a one period time-lag. The $\boldsymbol{\eta}_t$'s are IID multivariate normal with mean $\mathbf{0}$ and covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{\eta}}$.

BHL's objective was to minimize the expected execution cost, which leads to the problem formulation

$$\begin{aligned} \min_{\mathbf{s}_t \in \mathcal{F}_t, t=1,\ldots,T} \quad & \mathbb{E}_0\left[\sum_{t=0}^{T} \mathbf{p}_t' \mathbf{s}_t\right] \\ \text{subject to} \quad & \sum_{t=0}^{T} \mathbf{s}_t = \mathbf{w}_0 \end{aligned}$$

and dynamics (D-43)–(D-47). This problem can be solved using dynamic programming and explicit solutions for the optimal value function as well as the optimal trading quantities that can be found in BHL. When nonnegativity constraints $\mathbf{s}_t \geq 0$ are imposed, however, then it is not possible to solve this problem explicitly. Moreover, because of the return predictability induced by $\mathbf{x}_t$, we expect the nonnegativity constraints to be binding in general.

D.2. Suboptimal policies. We consider several different suboptimal policies that can be employed when nonnegativity constraints are imposed.

The simple policy. The agent buys the same quantity of shares in each of the $T+1$ time periods so that $\mathbf{s}_t = \mathbf{w}_0/(T+1)$.

A one-step look-ahead policy. At each time t the agent solves

$$\begin{aligned} \min_{\mathbf{s}_t \in \mathcal{F}_t} \quad & \mathbb{E}_t[\mathbf{p}_t' \mathbf{s}_t + V_{t+1}(\tilde{\mathbf{p}}_{t+1}, \mathbf{x}_{t+1}, \mathbf{w}_{t+1})] \\ \text{subject to} \quad & \mathbf{0} \leq \mathbf{s}_t \leq \mathbf{w}_t \end{aligned} \tag{D-48}$$

and implements the optimal solution, say $\mathbf{s}_t^{os}$. We take V_{t+1} to be the value function for the unconstrained problem which can be computed analytically via a matrix recursion. While the calculations are somewhat tedious the expectation of V_{t+1} conditional on $\mathcal{F}_t$ can be computed in closed form.

A rolling OLFC policy. At each time t the agent first solves

$$\begin{aligned} \min_{\mathbf{s}_{t:T} \in \mathcal{F}_t} \quad & \mathbb{E}_t\left[\sum_{j=t}^{T} \mathbf{p}_j' \mathbf{s}_j\right] \\ \text{subject to} \quad & \sum_{j=t}^{T} \mathbf{s}_j = \mathbf{w}_t, \\ & \mathbf{s}_j \geq \mathbf{0} \quad \text{for } j = t, \ldots, T. \end{aligned} \tag{D-49}$$

Let V_t^{ol} denote the optimal value of (D-49), and let $\mathbf{s}_{t:T}^{ol,t} := [\mathbf{s}_t^{ol,t\,\prime} \;\ldots\; \mathbf{s}_T^{ol,t\,\prime}]'$ denote the corresponding optimal solution. The OLFC policy implements $\mathbf{s}_t^{ol,t}$ at time t and ignores $\mathbf{s}_{t+1}^{ol,t}, \ldots, \mathbf{s}_T^{ol,t}$.

The single-stock execution policy. Here the agent simply assumes the cross-price impacts are zero, in which case the problem decouples into n separate single-stock problems. The agent solves each of these problems using a simple approximate dynamic programming algorithm and implements the resulting policy. This policy was suggested by the fact that the diagonal elements of the matrix $\mathbf{A}$, as calibrated by BHL, were typically an order of magnitude larger than the off-diagonal elements. This can be seen from Table 2 in Appendix D.4, where we display a 10×10 submatrix of the price-impact matrix, $\mathbf{A}$. We therefore expected this policy to perform well. Moreover, this case provides a clear example of where the duality technology could be used to determine whether or not a given policy, i.e., the n single-stock policies, should be adapted to account for a new feature, i.e., the cross-price-impact costs, that are known to exist in the marketplace.

D.3. The dual problem. A dual problem instance is obtained by simulating paths of $\tilde{\mathbf{p}}_t$ and $\mathbf{x}_t$ and then solving the resulting deterministic optimization problem with a dual-feasible penalty that is obtained from some approximation to the value function. A dual problem instance therefore takes the form

$$\text{(D-50)} \qquad \begin{aligned} \min_{\mathbf{s}_t \in \mathbb{S}} \quad & \sum_{t=0}^{T} \mathbf{p}_t' \mathbf{s}_t + \sum_{t=0}^{T-1} \Big(\mathbb{E}_t[\tilde{V}_{t+1}(\mathbf{s}_{0:t})] - \tilde{V}_{t+1}(\mathbf{s}_{0:t}) \Big) \\ \text{subject to} \quad & \mathbf{p}_t = \tilde{\mathbf{p}}_t + \tilde{\mathbf{P}}_t(\mathbf{A}_t \tilde{\mathbf{P}}_t \mathbf{s}_t + \mathbf{B}_t \mathbf{x}_t), \end{aligned}$$

where the $\tilde{\mathbf{p}}_t$'s and $\mathbf{x}_t$'s have been simulated according to the true model dynamics and are known to the decision-maker at time $t = 0$. We consider only $\tilde{V}_{t+1}(\mathbf{s}_{0:t})$'s that are linear in the actions, $\mathbf{s}_{0:t}$, and this can be achieved in the manner described in section 3.1. In the numerical results below, we compute dual bounds using penalties constructed from (i) the unconstrained value function, V_t, and (ii) the value function, V_t^{ss}, obtained by summing together the n single-stock value functions.

D.4. Numerical results. We use the same calibration for $\{\mathbf{A}, \mathbf{B}, \mathbf{C}, \boldsymbol{\mu}_{\boldsymbol{\epsilon}}, \boldsymbol{\Sigma}_{\boldsymbol{\epsilon}}, \boldsymbol{\Sigma}_{\boldsymbol{\eta}}\}$ as given by BHL. The agent therefore needs to purchase $100,000$ shares in each of 25 securities. There are $T + 1 = 20$ time periods. Primal bounds are computed on the basis of $10,000$ Monte Carlo paths, while only 100 paths were required to estimate the dual bounds. Monte Carlo results are displayed in Table 1. We use V^s, V^{os}, V^{ol}, and V^{ss} to denote the estimated average execution cost of the simple one-step look-ahead OLFC and single-stock policies, respectively. The mean and standard errors of the reported execution costs are cents per share. We also report runtimes (in seconds) for each of the policies with the exception of the simple policy, whose value can be determined analytically. Note that these runtimes are for the entire 10,000 paths. We see that the OLFC policy yields the lowest execution cost but that all policies provide very similar performance. We note, however, that computing the OLFC policy is particularly time-consuming relative to the other policies.

Table 1 also reports two dual bounds, V_d^{unc} and V_d^{ss}, that were obtained by using the unconstrained value function and the single-stock approximate dynamic programming value

Table 1

Simulation results for the portfolio execution problem with nonnegativity constraints.

	Suboptimal policy				Dual bound		
	V^s	V^{os}	V^{ol}	V^{ss}	V_0	V_d^{unc}	V_d^{ss}
Mean	5.364	4.123	3.979	4.194	2.023	3.976	3.975
Std error	0	0.0523	0.0501	0.0542	0	0.00797	0.00404
Runtimes	N/A	324	7465	13.81	N/A	14.51	14.54

Table 2

Price-impact coefficients for a 10×10 *subset of the* 25 *stocks in BHL. All values have been multiplied by* 10^{10}.

	AHP	AN	BLS	CHV	DD	DIS	DOW	F	FNM	GE
AHP	12.40	−1.69	−1.99	1.04	−1.07	1.09	1.26	1.37	−1.97	−0.17
AN	−1.32	10.10	−1.96	1.09	−1.63	−1.24	2.68	−2.03	−0.80	−0.30
BLS	3.49	0.83	14.40	2.26	−2.45	4.25	−1.02	−5.80	1.68	1.18
CHV	2.09	−0.17	2.34	21.20	2.84	1.62	−0.03	3.60	−3.35	0.91
DD	−0.93	0.66	6.18	1.55	11.70	−1.17	1.00	0.67	1.92	2.24
DIS	1.98	2.73	−0.30	−2.59	0.92	19.30	2.22	0.13	8.27	2.37
DOW	−0.79	0.18	−3.88	−0.23	−1.59	−0.93	7.21	−3.18	1.66	0.05
F	−0.23	0.66	3.29	−0.65	4.24	1.74	3.15	21.90	2.12	0.65
FNM	0.01	−0.92	−0.43	3.27	0.39	−3.57	3.40	2.78	13.70	−0.88
GE	0.77	−0.45	0.53	0.03	−0.41	−0.26	−1.10	0.12	1.60	5.06

function, respectively, to construct dual penalties, $\tilde{V}_t$, in (D-50). We also report the unconstrained optimal value function, V_0, which itself is a lower bound on the optimal execution cost for the constrained problem. We see the two dual bounds generated are very similarly and are very close indeed to the primal bounds.

Confirming the conjecture of BHL. These results allow us to confirm the conjecture of BHL, namely that the OLFC policy should be close to optimal. Indeed the average cost-per-share of the OLFC policy, 3.979 cents, is only .003 cents greater than the best dual bound.

Using the dual methodology to determine whether cross-price impacts need to be included. We also note that if this problem were ever encountered in practice, then it would be possible to "solve" it without constructing any of the more elaborate policies that explicitly account for the cross-price impacts. To see this note that the average execution cost of the single-stock policy is 4.194 cents per share and that the dual bound constructed by using the sum of the single-stock value functions to construct dual penalties is 3.975 cents per share. This implies that building a more elaborate execution policy that incorporates the cross-price impacts will be worth at most $4.194 - 3.975 = .219$ cents per share. As such, we may decide that it is not worth explicitly accounting for the cross-price impacts, at least in this model. This is one of the principal benefits of the dual technology.

We also mention here that we performed additional numerical experiments that included sector balance constraints in the problem formulation. We again found that the various policies performed very well, at least using the calibrated parameters of BHL.

Appendix E. Additional calibration details for section 4.

E.1. Variance-covariance dynamics. As stated in section 4.2 we assume that $\mathbf{\Sigma}_t$ follows an O-GARCH model as in Alexander [1] so that

$$\mathbf{\Sigma}_t = \mathbf{F}\mathbf{\Omega}_t\mathbf{F}' + \mathbf{\Upsilon}, \tag{E-51}$$

where $\mathbf{\Omega}_t$ is a diagonal matrix, $\mathbf{F}$ is a matrix of factor loadings, and $\mathbf{\Upsilon}$ is a diagonal matrix of idiosyncratic variances. The diagonal elements in $\mathbf{\Omega}_t$ are assumed to follow independent GARCH(1,1) processes.

Our calibration of the variance-covariance dynamics is very simple and is also based on Alexander [1]. We first compute the variance-covariance matrix, $\mathbf{\Sigma}$, of standardized 5-minute returns, treating the observations for each security as IID. We then perform a principal components analysis on $\mathbf{\Sigma}$ and then assume the first k principal components follow independent GARCH$(1,1)$ models. We emphasize here that we do not claim that this is a particularly good choice of model. Indeed we expect the proprietary models that are used in practice to be much better than our model, which is intended only to help demonstrate how the duality techniques can be used.

We use 5-minute return data on the 50 stocks between October 12 and October 25, 2011, which corresponds to a total of $d = 10$ trading days. Ignoring time of day and other effects we obtain a series of $10 \times 78 = 780$ observations that we initially treat as IID. After normalizing each time series by subtracting the mean return, we compute the covariance matrix, $\mathbf{\Sigma}$, of the $n = 50$ normalized observation series. We then perform a principal component analysis on $\mathbf{\Sigma}$ and obtain

$$\mathbf{\Sigma} = \mathbf{\Gamma}\mathbf{\Lambda}\mathbf{\Gamma}', \tag{E-52}$$

where $\mathbf{\Gamma}$ is the matrix of eigenvectors, say $\mathbf{c}_1, \ldots, \mathbf{c}_n$, and $\mathbf{\Lambda}$ is the diagonal matrix of corresponding eigenvalues, $\lambda_1, \ldots, \lambda_n$, arranged in decreasing order. We let $\mathbf{F}$ be the $n \times k$ matrix containing the k eigenvectors, $\mathbf{c}_1, \ldots, \mathbf{c}_k$, corresponding to the k largest eigenvalues. We now approximate the covariance matrix with

$$\mathbf{\Sigma} \approx \mathbf{F}\mathbf{\Omega}\mathbf{F}' + \mathbf{\Upsilon}, \tag{E-53}$$

where $\mathbf{\Omega} = \mathrm{diag}(\lambda_1, \ldots, \lambda_k)$ and $\mathbf{\Upsilon}$ is a diagonal matrix chosen to ensure that the diagonal terms on both sides of (E-53) agree. Note that λ_i is the variance of the ith principal component, $\mathbf{c}_i$. The eigenvalue analysis is shown in Table 3, where we see that the first six principal components explain more than 90% of the total variance in the 50 return series. We therefore chose $k = 6$. The problem with (E-53) is that it provides a static description for the covariance matrix, $\mathbf{\Sigma}$. We can use it to create a dynamic model, however, by setting

$$\mathbf{\Sigma}_t \approx \mathbf{F}\mathbf{\Omega}_t\mathbf{F}' + \mathbf{\Upsilon}, \tag{E-54}$$

where the diagonal elements of $\mathbf{\Omega}_t$ are assumed to follow independent GARCH$(1,1)$ processes. In particular, let $\sigma_{i,t}$ denote the value of the ith diagonal element of $\mathbf{\Omega}_t$. Then the GARCH$(1,1)$ model for $\sigma_{i,t}$ assumes

$$\sigma_{i,t+1}^2 = \omega_i + \alpha_i c_{i,t}^2 + \beta_i \sigma_{i,t}^2, \tag{E-55}$$

Table 3

Eigenvalue analysis.

Component	Eigenvalue	Cumulative explained variance
$\mathbf{c}_1$	1.2879	0.5776
$\mathbf{c}_2$	0.3261	0.7239
$\mathbf{c}_3$	0.1801	0.8047
$\mathbf{c}_4$	0.1399	0.8674
$\mathbf{c}_5$	0.0491	0.8894
$\mathbf{c}_6$	0.0348	0.9051

Table 4

GARCH(1, 1) parameter estimates for the top six principal components.

	ω		α		β	
	Coefficient	t-stat	Coefficient	t-stat	Coefficient	t-stat
$\mathbf{c}_1$	0.1889	2.381	0.1563	3.988	0.7001	8.165
$\mathbf{c}_2$	0.1360	2.423	0.2519	4.198	0.3284	1.548
$\mathbf{c}_3$	0.0231	3.386	0.0885	3.822	0.7801	14.882
$\mathbf{c}_4$	0.0114	3.357	0.1671	5.02	0.7575	17.895
$\mathbf{c}_5$	0.0046	4.532	0.1845	4.971	0.7322	18.403
$\mathbf{c}_6$	0.0033	3.273	0.2052	5.471	0.7131	15.341

where $\omega_i > 0$ and $\alpha_i, \beta_i \geq 0$ are fixed parameters and $c_{i,t}$ is the value of the ith principal component at time t. The parameters for the six GARCH models were fitted using standard MLE techniques and are given in Table 4. We note that all t-statistics are significant.

E.2. Parameter values for top 50 stocks in the S&P.

Table 5
Parameter values for top 50 stocks in the S&P.

Security ticker	Initial price	Average daily volume (million)	Annual volatility	Permanent linear price-impact coefficient $\times 10^9$	Temporary linear price-impact coefficient $\times 10^6$
AAPL	407.33	22.85	21.53%	178.3009	8.9150
ABT	52.47	24.65	21.55%	21.2826	1.0641
AIG	22.63	7.44	47.10%	30.4169	1.5208
AMGN	57.39	4.99	20.66%	114.9911	5.7496
AMZN	236.84	6.15	33.77%	385.3714	19.2686
AXP	45.93	10.36	31.70%	44.3491	2.2175
BAC	6.51	263.47	51.04%	0.2469	0.0123
BRK/B	74.03	6.44	24.55%	115.0245	5.7512
C	28.40	68.25	52.89%	4.1609	0.2080
CAT	81.80	10.63	37.09%	76.9853	3.8493
CMCSA	23.16	15.95	29.09%	14.5161	0.7258
COP	67.69	11.22	24.62%	60.3056	3.0153
CSCO	17.16	44.23	26.29%	3.8795	0.1940
CVS	34.44	8.04	19.53%	42.8311	2.1416
CVX	98.08	8.61	26.07%	113.9225	5.6961
DIS	32.95	10.10	25.96%	32.6170	1.6309
EMC	23.26	24.72	31.23%	9.4084	0.4704
GE	16.25	63.55	31.44%	2.5571	0.1279
GOOG	548.10	4.09	24.39%	1340.9765	67.0488
GS	98.05	7.83	41.66%	125.2213	6.2611
HD	34.92	10.67	24.01%	32.7156	1.6358
IBM	185.80	7.32	18.14%	253.9455	12.6973
INTC	23.00	84.04	24.81%	2.7367	0.1368
JNJ	64.10	11.39	18.39%	56.2684	2.8134
JPM	32.71	51.60	42.89%	6.3386	0.3169
KFT	34.67	8.23	16.83%	42.1164	2.1058
KO	67.22	16.75	16.81%	40.1212	2.0061
MCD	89.35	6.36	17.66%	140.5327	7.0266
MMM	76.77	5.19	28.50%	147.8743	7.3937
MO	27.89	11.77	19.24%	23.7031	1.1852
MRK	32.24	13.86	19.48%	23.2673	1.1634
MSFT	27.18	54.55	22.81%	4.9830	0.2491
ORCL	31.35	27.79	26.15%	11.2801	0.5640
OXY	81.25	4.72	35.25%	171.9937	8.5997
PEP	62.45	8.12	16.41%	76.9138	3.8457
PFE	18.85	36.61	23.23%	5.1491	0.2575
PG	64.71	9.26	13.87%	69.9185	3.4959
PM	65.82	8.40	19.55%	78.3178	3.9159
QCOM	52.29	13.83	28.01%	37.8144	1.8907
SLB	67.46	12.53	41.87%	53.8395	2.6920
T	28.83	24.74	16.46%	11.6527	0.5826
UNH	47.21	9.23	35.99%	51.1342	2.5567
UPS	68.32	4.67	23.28%	146.3456	7.3173
USB	24.15	18.35	37.35%	13.1607	0.6580
UTX	74.31	4.79	27.04%	155.0823	7.7541
V	90.99	3.97	25.88%	229.3355	11.4668
VZ	36.64	13.44	17.03%	27.2555	1.3628
WFC	26.38	48.78	39.73%	5.4085	0.2704
WMT	55.15	12.59	16.54%	43.8191	2.1910
XOM	76.87	21.67	20.25%	35.4807	1.7740

Acknowledgments. We thank Costis Maglaras for providing us with the 5-minute price data on the S&P 500 securities that we used in section 4 and Andrew Lo for providing the calibrated model parameters for the numerical experiments in Appendix D. We are also grateful to Costis Maglaras and Garud Iyengar for helpful comments and discussions. All errors are our own.

REFERENCES

[1] C. Alexander, *Principal component models for generating large Garch covariance matrices*, Econ. Notes, 31 (2002), pp. 337–359.

[2] A. Alfonsi, A. Fruth, and A. Schied, *Optimal execution strategies in limit order books with general shape function*, Quant. Finance, 10 (2010), pp. 143–157.

[3] R. Almgren, *Optimal trading with stochastic liquidity and volatility*, SIAM J. Financial Math., 3 (2012), pp. 163–181.

[4] R. Almgren and N. Chriss, *Optimal execution of portfolio transactions*, J. Risk, 3 (2001), pp. 5–39.

[5] R. Almgren and N. Chriss, *Value under liquidation*, Risk Magazine, 12 (1999), pp. 61–63.

[6] L. Andersen and M. Broadie, *Primal-dual simulation algorithm for pricing multidimensional American options*, Manag. Sci., 50 (2004), pp. 1222–1234.

[7] C. Bender, *Primal and dual pricing of multiple exercise options in continuous time*, SIAM J. Financial Math., 2 (2011), pp. 562–586.

[8] C. Bender, J. Schoenmakers, and J. Zhang, *Dual representations for general multiple stopping problems*, Math. Finance, to appear.

[9] D. Bertsimas, P. Hummel, and A.W. Lo, *Optimal control of execution costs for portfolios*, Comput. Sci. Eng., 1 (1999), pp. 40–53.

[10] D. Bertsimas and A.W. Lo, *Optimal control of execution costs*, J. Financ. Market, 1 (1998), pp. 1–50.

[11] S.P. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, Cambridge, UK, 2004.

[12] D.B. Brown and J.E. Smith, *Dynamic portfolio optimization with transaction costs: Heuristics and dual bounds*, Manag. Sci., 57 (2011), pp. 1752–1770.

[13] D.B. Brown, J.E. Smith, and P. Sun, *Information relaxations and duality in stochastic dynamic programs*, Oper. Res., 58 (2010), pp. 785–801.

[14] S. Buti, B. Rindi, and I.M. Werner, *Dark Pool Trading Strategies*, Technical report, 2011.

[15] S. Chandramouli and M. Haugh, *A unified approach to multiple stopping and duality*, Oper. Res. Lett., 40 (2012), pp. 258–264.

[16] C. O' Cinneide, B. Scherer, and X. Xu, *Pooling trades in a quantitative investment process*, J. Portfolio Manage., 32 (2006), pp. 33–43.

[17] R. Cont and A. Kukanov, *Optimal Order Placement in Limit Order Markets*, Technical report, 2013.

[18] R. Cont, S. Stoikov, and R. Talreja, *A stochastic model for order book dynamics*, Oper. Res., 58 (2010), pp. 549–563.

[19] M.H.A. Davis and I. Karatzas, *A deterministic approach to optimal stopping*, in Probability, Statistics and Optimisation, F.P. Kelly, ed., John Wiley & Sons, New York, Chichester, 1994, pp. 455–466.

[20] M.H.A. Davis and M. Zervos, *A new proof of the discrete-time LQG optimal control theorem*, IEEE Trans. Automat. Control, 40 (1995), pp. 1450–1453.

[21] V.V. Desai, V.F. Farias, and C.C. Moallemi, *Bounds for Markov decision processes*, in Reinforcement Learning and Approximate Dynamic Programming for Feedback Control, F.L. Lewis and D. Liu, eds., IEEE Press, New York, 2012, pp. 452–473.

[22] J. Gatheral, *No-dynamic-arbitrage and market impact*, Quant. Finance, 10 (2010), pp. 749–759.

[23] J. Gatheral and A. Schied, *Optimal trade execution under geometric Brownian motion in the Almgren and Chriss framework*, Int. J. Theor. Appl. Finance, 14 (2011), pp. 353–368.

[24] P. Glasserman, *Monte Carlo Methods in Financial Engineering*, Appl. Math. (N. Y.) 53, Springer-Verlag, New York, 2004.

[25] O. Guéant, *Permanent Price Impact Can Be Nonlinear*, Technical report, 2013.

[26] M.B. Haugh and L. Kogan, *Pricing American options: A duality approach*, Oper. Res., 52 (2004), pp. 258–270.
[27] M.B. Haugh and A.E.B. Lim, *Linear-quadratic control and information relaxations*, Oper. Res. Lett., 40 (2012), pp. 521–528.
[28] G. Huberman and W. Stanzl, *Price manipulation and quasi-arbitrage*, Econometrica, 72 (2004), pp. 1247–1275.
[29] G. Huberman and W. Stanzl, *Optimal liquidity trading*, Rev. Finance, 9 (2005), pp. 165–200.
[30] R. Huitema, *Optimal Portfolio Execution Using Market and Limit Orders*, Technical report, 2012.
[31] F. Jamshidian, *The duality of optimal exercise and domineering claims: A Doob-Meyer decomposition approach to the Snell envelope*, Stochastics, 79 (2007), pp. 27–60.
[32] P. Kratz and T. Schoneborn, *Portfolio liquidation in dark pools in continuous time*, Math. Finance, to appear.
[33] G. Lai, F. Margot, and N. Secomandi, *An approximate dynamic programming approach to benchmark practice-based heuristics for natural gas storage valuation*, Oper. Res., 58 (2010), pp. 564–582.
[34] G. Lai, M. X. Wang, S. Kekre, A. Scheller-Wolf, and N. Secomandi, *Valuation of storage at a liquefied natural gas storage terminal*, Oper. Res., 59 (2011), pp. 602–616.
[35] A.E.B. Lim and P. Wimonkittiwat, *Dynamic portfolio choice with market impact costs*, in Proceedings of the 50th IEEE Conference on Decision and Control and European Control Conference, 2011.
[36] N. Meinshausen and B.M. Hambly, *Monte Carlo methods for the valuation of multiple-exercise options*, Math. Finance, 14 (2004), pp. 557–583.
[37] C. Moallemi and M. Sağlam, *Dynamic Portfolio Choice with Linear Rebalancing Rules*, Technical report, 2011.
[38] A. Obizhaeva and J. Wang, *Optimal trading strategy and supply/demand dynamics*, J. Financ. Market, 16 (2013), pp. 1–32.
[39] S. Predoiu, G. Shaikhet, and S. Shreve, *Optimal execution in a general one-sided limit-order book*, SIAM J. Financial Math., 2 (2011), pp. 183–212.
[40] L.C.G. Rogers, *Monte Carlo valuation of American options*, Math. Finance, 12 (2002), pp. 271–286.
[41] L.C.G. Rogers, *Pathwise stochastic optimal control*, SIAM J. Control Optim., 46 (2007), pp. 1116–1132.
[42] A. Schied, T. Schoneborn, and M. Tehranci, *Optimal basket liquidation for Cara investors is deterministic*, Appl. Math. Finance, 17 (2010), pp. 471–489.
[43] J. Schoenmakers, *A pure martingale dual for multiple stopping*, Finance Stoch., 16 (2012), pp. 319–334.
[44] R. A. Stubbs and D. Vandenbussche, *Multi-portfolio Optimization and Fairness in Allocation of Trades*, Technical report, 2009.
[45] B. Tóth, Y. Lempérière, C. Deremble, J. de Lataillade, J. Kockelkoren, and J.P. Bouchaud, *Anomalous price impact and the critical nature of liquidity in financial markets*, Phys. Rev. X, 1 (2011), 021006.
[46] G. Tsoukalas, J. Wang, and K. Giesecke, *Dynamic Portfolio Execution with Correlated Supply/Demand Curves*, Technical report.