Manufacturing & Service Operations Management

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Merchant Commodity Storage and Term-Structure Model Error

Nicola Secomandi, Guoming Lai, François Margot, Alan Scheller-Wolf, Duane J. Seppi

MANUFACTURING & SERVICE OPERATIONS MANAGEMENT
Vol. 17, No. 3, Summer 2015, pp. 302–320
ISSN 1523-4614 (print) | ISSN 1526-5498 (online)
http://dx.doi.org/10.1287/msom.2015.0518

Merchant Commodity Storage and Term-Structure Model Error

Nicola Secomandi
Tepper School of Business, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, ns7@andrew.cmu.edu

Guoming Lai
McCombs School of Business, University of Texas at Austin, Austin, Texas 78712, guoming.lai@mccombs.utexas.edu

François Margot, Alan Scheller-Wolf, Duane J. Seppi
Tepper School of Business, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213
{fmargot@andrew.cmu.edu, awolf@andrew.cmu.edu, ds64@andrew.cmu.edu}

Merchant operations involves valuing and hedging the cash flows of commodity- and energy-conversion assets as real options based on stochastic models that inevitably embed model error. In this paper we quantify how empirically calibrated model errors concerning the futures term structure affect the valuation and hedging of natural gas storage. We find that even small model errors—on the order of 1%–2% of the empirical futures price variance—can have a disproportionate impact on storage valuation and hedging. In particular, theoretically equivalent hedging strategies have very different sensitivities to model error, with one natural strategy exhibiting potentially catastrophic performance in the presence of small model errors. We propose effective approaches to mitigate the negative effect of futures term-structure model error on hedging, also taking into account futures contract illiquidity, and provide theoretical justification for some of these approaches. Beyond commodity storage, our analysis has relevance for other real and financial options that depend on futures term-structure dynamics, as well as for inventory, production, and capacity investment policies that rely on demand-forecast term structures.

Keywords: model error; commodity and energy real options; natural gas storage; futures term structures; delta hedging and mean-variance hedging
History: Received: September 9, 2012; accepted: November 21, 2014. Published online in *Articles in Advance* March 26, 2015.

1. Introduction

Merchants manage commodity- and energy-conversion assets—copper mines, natural gas storage facilities and pipelines, power plants, and oil wells and refineries—as real options on commodity and energy prices. A canonical *merchant operations* problem involves buying or leasing commodity conversion assets and then using an operating policy to maximize their real option value by trading the underlying physical input and output commodities (Secomandi and Seppi 2014).

As in classical operations management, the management of merchant operations entails making operational choices, such as storage, transportation, production, and conversion (processing and refining) decisions. However, the management of merchant operations differs from classical operations management because the key source of uncertainty in merchant operations is price, rather than customer demand. Hence, the goal in merchant operations is to maximize trading margins in the face of price uncertainty rather than to minimize the cost of providing adequate service to customers. Operational decisions are critical in merchant operations because these choices directly affect the merchant's trading margins.

Merchants also engage in financial hedging of their operational cash flows to mitigate the effects of mismatches between *financing* payments and *operational* cash flows in the presence of credit frictions (see, e.g., §5.4 of Tirole 2006 and references therein). This hedging activity is analogous to how financial institutions manage similar risks in their financial derivative trading operations (Hull 2010). It is a complex multifunctional task faced every day in merchant operations.

The key components of merchant operations are a valuation model, an operating policy, and risk sensitivities for dynamic financial hedging. Each of these components relies on an assumed model of the risk-neutral dynamics of the underlying state variables (Smith and McCardle 1999, Birge 2000, Seppi 2002). Since, in practice, the risk-neutral dynamics are not known with certainty, model error can cause merchants to misvalue and potentially mismanage commodity-conversion assets and can disrupt risk management (hedging) strategies.

Real options have been studied extensively in the theoretical and practice-based literatures (Dixit and Pindyck 1994, Smith and McCardle 1999, Clewlow and Strickland 2000, Eydeland and Wolyniec 2003, Geman 2005, Leppard 2005), but model error has not been analyzed in the context of merchant operations. Our goals in this paper are to (a) quantify the impact of model error on real option valuation and hedging in merchant operations and (b) propose practical strategies to mitigate the exposure of merchant operations to model error.

We focus on commodity storage assets, which allow merchants to trade a commodity over time; specifically, we consider natural gas storage lease contracts (Maragos 2002). Natural gas is an important commodity, representing 25% of total U.S. energy consumption in 2009 (according to the U.S. Energy Information Adminstration (EIA) 2010), and it is projected to become even more important in the future (EIA 2011a). Natural gas storage capacity amounts to roughly 17% of annual U.S. natural gas demand (EIA 2011b, c). With natural gas at a price of \$4.9/mmBtu (as of February 2014), the value of natural gas peak inventory is about \$20 billion. Natural gas storage valuation has attracted a substantial amount of research (Chen and Forsyth 2007, Boogert and de Jong 2008, Thompson et al. 2009, Carmona and Ludkovski 2010, Lai et al. 2010, Secomandi 2010b, Bjerksund et al. 2011, Boogert and de Jong 2011/2012, Thompson 2012, Wu et al. 2012).

As described in the practice-based literature (Maragos 2002, pp. 440 and 449–453; Gray and Khandelwal 2004a, b), natural gas storage merchants use multifactor models for the term structure of commodity futures prices (Clewlow and Strickland 2000, Chap. 8; Eydeland and Wolyniec 2003, Chap. 5 and pp. 351–367; Geman 2005, Chap. 3). Empirically, the dynamics of commodity and energy futures price term structures are dominated by a few large *common* factors. A typical number of factors is three, which have qualitative interpretations of level (shifting), slope (tilting), and curvature (bending) effects (Cortazar and Schwartz 1994; Clewlow and Strickland 2000, Chap. 8; Blanco et al. 2002; Tolmasky and Hindanov 2002; Geman and Nguyen 2005; Borovkova and Geman 2008; Frestad 2008). However, fewer factors (Manoliu and Tompaidis 2002, Borovkova and Geman 2006, Suenaga et al. 2008, Wu et al. 2012) or more factors (Eydeland and Wolyniec 2003, pp. 351–367; Gray and Khandelwal 2004a, b; Bjerksund et al. 2011; Thompson 2012, 2013) are also used. In particular, Eydeland and Wolyniec (2003, pp. 351–367) and Gray and Khandelwal (2004a, b) discuss modeling the full term structure of the natural gas futures curve using as many common factors as there are futures delivery dates over the term of a storage contract. This approach is related to string and Brace-Gatarek-Musiela models used to value fixed-income options (Kennedy 1994, Brace et al. 1997, Longstaff et al. 2001). Moreover, Frestad (2008) and Suenaga et al. (2008) model *maturity-specific* (idiosyncratic) noise in addition to common factors.

There is an infinite variety of possible term-structure model errors, so we look to empirical data to impose a realistic structure on the types of model errors we analyze. In other words, we study the impact of empirically realistic model errors that merchant storage traders are *likely* to face. Our starting point is the estimation of a flexible family of multifactor futures curve models (Cortazar and Schwartz 1994, Blanco et al. 2002). For New York Mercantile Exchange (NYMEX) natural gas futures prices, as expected, we find that the three most important common factors explain about 98%–99% of the observed daily variance of futures price log returns. However, even with such empirical evidence, merchant traders are still uncertain about the true model for natural gas futures price term-structure dynamics. For example, does the treatment of the last 1%–2% of the observed futures price log-return variance matter for storage valuation and hedging? We thus center our model error analysis on various futures term-structure models that differ in their treatment of the three most important common factors and of this small residual variance, organizing the resulting model errors in a novel framework.

Our quantitative analysis uses the rolling intrinsic (RI) operating policy, which is both widely used in practice and near optimal (see Lai et al. 2010 and references therein), to value realistic natural gas storage contracts. We analyze versions of practical hedging methods based on both cash flow replication through delta hedging (Hull 2010, Chap. 6; Luenberger 2014, §12.5) and minimization of hedged cash flow variance (Luenberger 2014, §12.10). In particular, for delta hedging we consider *bucket hedging* (BH), which trades futures contracts with delivery dates corresponding to each date on which operational trading has a cash flow (Driessen et al. 2003), and *factor hedging* (FH), which just trades as many futures contracts as the number of factors in the assumed futures price model (Cortazar and Schwartz 1994; Clewlow and Strickland 2000, §9.5). FH is appealing because, in contrast to BH, it is feasible when there are fewer futures contracts available than cash flow dates for the option. Both of these methods take the storage deltas as inputs. We estimate these deltas by extending the pathwise approach (see §7.2 of Glasserman 2004 and references therein) to real options with inventory. (Secomandi and Wang 2012 extend this approach to real options without inventory.) We also obtain a novel expression for the FH positions.

Our analysis shows that model error concerning even small amounts of empirical variance can have a disproportionate impact on storage valuation and hedging. The inclusion/exclusion/treatment of the last 1%–2% of the empirical futures price log-return variance can change storage valuations by ±14%, an amount that is material for merchants when deciding how to bid when acquiring storage capacity. The reason for this numerical disparity is that storage valuation depends strongly on the ability to trade on transitory price shocks, even if they are statistically small in comparison to statistically large but more persistent shocks. The impact of model error on hedging performance—which we measure as the reduction in the variance of the operational cash flows—can be even more dramatic for both simulated and historical prices. The performance of hedges that in the absence of model error are theoretically equivalent can be orders of magnitude different given even small model errors, pertaining, again, to just 1%–2% of the futures price log-return variance. Specifically, we find that BH performs near optimally and is remarkably robust to model error—an observation for which we provide some theoretical support. By contrast, FH is very sensitive to which futures contracts are used to hedge and the sizes of the positions taken. Implementing this approach using the nearby, and most liquid, futures contracts, a method that we label *naïve factor hedging* (NFH), performs disastrously in the presence of even small model errors. The differential performance of these hedging strategies is closely tied to the size of the futures positions taken. Large position sizes can magnify hedging errors as a result of even small amounts of futures curve randomness being mistakenly omitted from the futures price model.

We also explain how to mitigate the effect of term-structure model error on hedging. We develop a novel FH variant, *fine-tuned factor hedging* (FTFH), which, by heuristically minimizing the size of the trading positions, performs near optimally. However, when the storage contract is long-dated, both BH and FTFH may be infeasible because of futures contract illiquidity. We thus propose a new minimum variance hedging (MVH) method, *constrained minimum variance hedging* (CMVH), which trades only liquid contracts (see also Frestad 2012). We find that CMVH performs quite well in the presence of both model error and futures contract illiquidity.

Our research improves our understanding of the effect of realistic futures term-structure model errors on merchant commodity storage management and provides methods to mitigate the negative effect of such errors on hedging. More broadly, our work has relevance for valuing and hedging other commodity- and energy-conversion assets and financial options using potentially misspecified multifactor models. Examples include commodity processing, refining, and transport assets (Secomandi 2010a, Wu and Chen 2010, Boyabatlı et al. 2011, Devalkar et al. 2011, Lai et al. 2011, Secomandi and Wang 2012, Plambeck and Taylor 2013, Thompson 2013); commodity swing and Bermudan options (Jaillet et al. 2004, Detemple 2006); and mortgages and interest rate caps, floors, and swaptions (Fan et al. 2001, Longstaff et al. 2001, Driessen et al. 2003, Veronesi 2010). Model error is also possible when applying the martingale model of demand-forecast evolution (Graves et al. 1986, Heath and Jackson 1994) to determine inventory/production management and capacity investment policies under demand uncertainty, for which financial hedging can be relevant (Birge 2000, Van Mieghem 2003, Berling and Rosling 2005, Gaur and Seshadri 2005, Caldentey and Haugh 2006, Ding et al. 2007, Goel and Gutierrez 2011, Kouvelis et al. 2013, Secomandi and Kekre 2014).

We present a common multifactor model of futures price dynamics in §2. We discuss the valuation and hedging of natural gas storage under such a model in §§3 and 4, respectively. We calibrate this futures price model in §5. In §6 we use this calibration to introduce our framework for studying futures term-structure model error. We propose approaches to mitigate the negative effect of this model error on hedging in §7. We report on our numerical analysis of futures term-structure model error in §8. Section 9 concludes. Online Appendix A (available as supplemental material at http://dx.doi.org/10.1287/msom.2015.0518) summarizes the acronyms used in this paper. Online Appendix B extends the pathwise delta estimation analysis to the case of exact storage valuation. Online Appendix C provides comprehensive plots from our calibration. Online Appendix D presents the estimated standard errors in our numerical hedging analysis. Online Appendix E includes proofs.

2. A Family of Futures Price Evolution Models

Merchants often use multifactor models to describe the risk-neutral dynamics of the term structure of commodity futures prices. Consider the first $N > 1$ contracts in the natural gas futures curve over the term $[T_0, T_{N-1}]$ of a storage contract, with T_0 set equal to 0. We denote by $F(t, T_m)$ the futures price at time $t \in [0, T_m]$ with maturity date T_m for all m in the set $\mathcal{N}$ of monthly maturity labels $\{0, \ldots, N-1\}$. The spot price at time T_m is $F(T_m, T_m)$. Given a number of common factors $K \in \{1, \ldots, N-1\}$, the following stochastic differential equation describes the merchant's assumed risk-neutral dynamics of $F(t, T_m)$ over the time interval $[0, T_m]$:

$$\frac{dF(t, T_m)}{F(t, T_m)} = \sum_{k \in \mathcal{K}} \sigma_{m,k}(t)\, dZ_k(t), \tag{1}$$

where $\mathcal{K}$ is the set $\{1, \ldots, K\}$ of factor labels, $\sigma_{m,k}(t)$ is the loading coefficient of the futures price with maturity date T_m at time t on factor k, and $dZ_k(t)$ is a standard Brownian motion increment corresponding to factor k. The standard Brownian motion increments for the different factors are all uncorrelated. Seasonality in spot price levels is captured by the initial futures prices themselves. Model (1) also allows for possible seasonality in the futures log-return covariance matrix, because the factor loading coefficients can depend on calendar time t.

Factor models of futures price dynamics are mathematically equivalent to factor models of spot price dynamics (Cortazar and Schwartz 1994; Secomandi and Seppi 2014, Chap. 4). For example, the two-factor spot price model of Schwartz and Smith (2000) and §§5.3.2.2 and 6.7 of Pilipovic (2007) is equivalent to a two-factor futures curve model with parallel shift loadings on one factor and decaying loadings on a second (correlated) factor (see Ross 1997, Schwartz 1997, Jaillet et al. 2004, Casassus and Collin-Dufresne 2005 for related models). We conduct our analysis using a futures curve specification both because of its empirical flexibility and because, in practice, futures are widely used for hedging (Clewlow and Strickland 2000, §8.5 and Chap. 9).

Our specification of model (1) restricts each coefficient $\sigma_{m,k}(\cdot)$ to be a constant $\sigma_{m,k,n}$ over each time interval $[T_n, T_{n+1})$, with $n \in \mathcal{N}$ and m in the set $\mathcal{N}_{n+1}$ of remaining maturity labels $\{n+1, \ldots, N-1\}$. That is, distinct factor loadings $\sigma_{m,k,n}$ are obtained for each futures maturity for each trading month. This seasonal specification captures the important role of weather in natural gas supply and demand. Given n and $m \in \mathcal{N}_{n+1}$, under this specification, the merchant's assumed dynamics of the futures price $F(t, T_m)$ during the time interval $[T_n, T_{n+1})$ are

$$\frac{dF(t, T_m)}{F(t, T_m)} = \sum_{k \in \mathcal{K}} \sigma_{m,k,n}\, dZ_k(t). \tag{2}$$

3. Valuation

This section is based in part on Lai et al. (2010, §2 and §§3.2 and 3.3). Natural gas storage contracts give merchants the right to inject, store, and withdraw natural gas at a storage facility over a finite time horizon, subject to capacity and inventory constraints. Capacity flow constraints specify the maximal amount of natural gas, measured in million British thermal units, that a merchant can inject or withdraw per unit of time. Inventory constraints specify the minimal and maximal amounts of natural gas inventory that the merchant can hold at any given point in time. A contract also specifies injection and withdrawal charges, as well as fuel losses.

We assume that the contracted storage facility is located near a liquid wholesale spot market where merchants execute their physical trading. This assumption is realistic for North America (which has roughly 100 geographically dispersed wholesale markets for natural gas) and for the United Kingdom (the National Balancing Point), Belgium (the Zeebrugge Hub), Germany (the Norsea Gas Terminal in Emden), and the Netherlands (the Title Transfer Facility). We also assume that a natural gas futures market is associated with the wholesale physical market. This assumption is again realistic for North America, where NYMEX and the Intercontinental Exchange (ICE) trade futures contracts with delivery at the Henry Hub in Erath, Louisiana and basis swaps for about 40 other locations (financially settled forward locational price differences relative to the Henry Hub). In the United Kingdom, ICE trades natural futures associated with the National Balancing Point. The European Energy Exchange also trades natural gas futures.

We are interested in valuing the cash flows associated with making physical natural gas trading decisions on a monthly basis. The set of maturity labels, $\mathcal{N}$, is thus the stage set. This is realistic because the spot market in the United States is most liquid during the monthly "bid week" in which blocks of gas for the ensuing month are traded. The timing of the physical trading/inventory decisions is assumed to coincide with the futures maturity dates, which define the stage set. Let a denote an action. A purchase-and-injection is a negative action (i.e., it generates a negative cash flow), a withdrawal-and-sale is a positive action, and the do-nothing action is zero. The cash flow for a physical transaction made at date T_n happens at time T_n. Natural gas purchased (respectively, sold) at time T_n is available in (respectively, removed from) storage by time T_{n+1}. A nonzero action thus represents a commodity flow in between two successive stages.

The minimal and maximal inventory levels are 0 and $\bar{x} \in \Re_+$, respectively, making the set of feasible inventory levels $[0, \bar{x}] =: \mathcal{X}$. The injection and withdrawal capacities per stage are $C^I < 0$ and $C^W > 0$, respectively. The feasible injection action, withdrawal action, and action sets, with feasible inventory x, are $\mathcal{A}^I(x) := [C^I \vee (x - \bar{x}), 0]$, $\mathcal{A}^W(x) := [0, x \wedge C^W]$, and $\mathcal{A}(x) := \mathcal{A}^I(x) \cup \mathcal{A}^W(x)$, respectively (where $\cdot \wedge \cdot \equiv \min\{\cdot, \cdot\}$ and $\cdot \vee \cdot \equiv \max\{\cdot, \cdot\}$). We model in-kind fuel losses using the coefficients $\phi^W \in (0,1]$ and $\phi^I \geq 1$ for withdrawals and injections, respectively. The marginal withdrawal and injection costs are c^W and c^I, respectively. Thus, given an action a and the spot price s, the per-stage cash flow function $p(a,s)$ is $(\phi^I s + c^I)a$ if $a \in \Re_-$, 0 if $a = 0$, and $(\phi^W s - c^W)a$ if $a \in \Re_+$.

The storage contract is managed using a feasible physical trading policy, π. Using the more compact notation $F_{n,m}$ for the futures price $F(T_n, T_m)$, let $\mathbf{F}_n := (F_{n,m}, m \in \mathcal{N}_n)$ $n \in \mathcal{N}$ denote the futures curve at time T_n inclusive of the spot price $s_n \equiv F_{n,n}$, with the definition $\mathbf{F}_N := 0$. Let $A_n^\pi(x, \mathbf{F}_n)$ be the decision rule giving the amount of gas sold or bought under a policy π at stage n in state $(x, \mathbf{F}_n)$, and let x_n^π be the inventory level reached at stage n by policy π. Denote risk-neutral expectation by $\mathbb{E}$ and the per-stage risk-free discount factor by the (for simplicity) constant δ. The value function corresponding to policy π in state $(x, \mathbf{F}_n)$ at stage n under an assumed model (2) is

$$V_n^\pi(x, \mathbf{F}_n) := \sum_{m=n}^{N-1} \delta^{m-n} \mathbb{E}[p(A_m^\pi(x_m^\pi, \mathbf{F}_m), s_m) \mid x, \mathbf{F}_n]. \quad (3)$$

The optimal storage valuation problem can be formulated as a Markov decision process, but this model is, in general, intractable because of the curse of dimensionality, as decisions are conditioned on the entire futures curve $\mathbf{F}_n$ at each stage (see Online Appendix B for a stochastic dynamic programming formulation for this model). As a result, merchants often manage natural gas storage optionality using simpler heuristic operating policies. Our theoretical analysis of hedging in §4 considers general feasible operating policies. Our numerical analysis in §8 specifically restricts attention to the RI heuristic operating policy, which is both widely used among practitioners (Maragos 2002) and has been shown by Lai et al. (2010) to be nearly optimal (see also Thompson 2012, Wu et al. 2012). This policy is based on sequential reoptimization of the model that computes the intrinsic value of storage (i.e., the deterministic equivalent of the stochastic dynamic program in Online Appendix B). The value of the RI policy in a given stage and state can be estimated by a Monte Carlo simulation of the evolution of the futures curve and the resulting inventory levels and cash flows. We refer the reader to §§3.2 and 3.3 of Lai et al. (2010) for details on this policy. We expect that our model error analysis in §8 would carry over qualitatively to other operating policies.

4. Hedging

Under price model (2), hedging occurs, in theory, in continuous time; therefore, we map the discrete-time function $V_n^\pi(\cdot,\cdot)$ in (3) into a continuous-time function. For every maturity label $n \in \mathcal{N} \backslash \{0\}$ and date $t \in [T_{n-1}, T_n]$, we denote as $\check{\mathbf{F}}(t)$ the $(N-n)$-dimensional vector of prices $(F(t, T_m), m \in \mathcal{N}_n)$, where $\mathcal{N}_n$ is the set $\{n, \ldots, N-1\}$ ($\check{\mathbf{F}}(T_{n-1}) \neq \mathbf{F}_{n-1}$ because s_{n-1} is not part of $\check{\mathbf{F}}(T_{n-1})$, whereas $\check{\mathbf{F}}(T_n) = \mathbf{F}_n$ because s_n is part of $\check{\mathbf{F}}(T_n)$); we then define the continuous-time continuation function at time t in state $(x, \check{\mathbf{F}}(t)) \in \mathcal{X} \times \Re_+^{N-n}$ for a given feasible operating policy π as

$$U^\pi(t, x, \check{\mathbf{F}}(t)) := \bar{\delta}(t, T_n) \mathbb{E}[V_n^\pi(x, \mathbf{F}_n) \mid \check{\mathbf{F}}(t)],$$

where $\bar{\delta}(t, T_n)$ is the risk-free discount factor from time T_n back to time t ($\bar{\delta}(T_{n-1}, T_n)$ corresponds to the constant one-period risk-free discount factor δ); the dependence of $U^\pi(t, x, \check{\mathbf{F}}(t))$ and $\check{\mathbf{F}}(t)$ on n is implicit in our notation given t. The function $U^\pi(t, x, \check{\mathbf{F}}(t))$ is the time t value of having x inventory units in storage at time T_n given the vector $\check{\mathbf{F}}(t)$ of futures prices at time t when using policy π to manage storage under the merchant's assumed price model (2).

For hedging, we need to know how the value of storage changes in response to changes in the futures curve. Given price model (2), by Ito's lemma (see, e.g., Glasserman 2004, Theorem B.1.1), the dynamics of $U^\pi(t, x, \check{\mathbf{F}}(t))$ over $(T_{n-1}, T_n]$ are

$$dU^\pi(t, x, \check{\mathbf{F}}(t)) = D(t)\, dt + \sum_{m \in \mathcal{N}_n} \Delta_m^\pi(x, t, \check{\mathbf{F}}(t))\, dF(t, T_m), \quad (4)$$

where the drift $D(t)$ is

$$\frac{\partial U^\pi(t, x, \check{\mathbf{F}}(t))}{\partial t} + \frac{1}{2} \sum_{m \in \mathcal{N}_n} \sum_{l \in \mathcal{N}_n} \frac{\partial^2 U^\pi(t, x, \check{\mathbf{F}}(t))}{\partial F(t, T_m) \partial F(t, T_l)} \cdot \sum_{k \in \mathcal{K}} \sigma_{m,k,n-1} F(t, T_m) \sigma_{l,k,n-1} F(t, T_l),$$

and the storage *delta* with respect to the time t futures prices with maturity T_m, with $m \in \mathcal{N}_n$, is

$$\Delta_m^\pi(t, x, \check{\mathbf{F}}(t)) := \frac{\partial U^\pi(t, x, \check{\mathbf{F}}(t))}{\partial F(t, T_m)}. \quad (5)$$

When $t = T_{n-1}$, the inventory x is held fixed, because the deltas are computed right after the *commercial* implementation of a feasible action at time T_{n-1}. Otherwise, when $t \in (T_{n-1}, T_n)$, inventory is fixed based on the decision made at time T_{n-1}.

To ease exposition, we mostly simplify the suffix $(t, x, \check{\mathbf{F}}(t))$ to (t). Denote by $q_m^\pi(t)$ the replicating position in the futures contract with maturity T_m, $m \in \mathcal{N}_n$, at time $t \in [T_{n-1}, T_n)$ when hedging $U^\pi(t)$. The dynamics of the value $\Pi^\pi(t)$ of a portfolio that is long physical storage and continuously *shorts* these futures positions are $d\Pi^\pi(t) = dU^\pi(t) - \sum_{m \in \mathcal{N}_n} q_m^\pi(t)\, dF(t, T_m)$, which, using (4), under the merchant's assumed specification (2), can be expressed as

$$d\Pi^{\pi}(t) = D(t)\ dt + \sum_{k\in\mathcal{K}}\left[\sum_{m\in\mathcal{N}_n}\Delta_m^{\pi}(t)F(t,T_m)\sigma_{m,k,n-1} - \sum_{m\in\mathcal{N}_n} q_m^{\pi}(t)F(t,T_m)\sigma_{m,k,n-1}\right]dZ_k(t). \quad (6)$$

It follows from (6) that the stochastic variability is hedged—in other words, the futures positions replicate the assumed dynamics of the value of the storage asset—when

$$\sum_{m\in\mathcal{N}_n} q_m^{\pi}(t)F(t,T_m)\sigma_{m,k,n-1} = \sum_{m\in\mathcal{N}_n}\Delta_m^{\pi}(t)F(t,T_m)\sigma_{m,k,n-1} \quad \forall\, k\in\mathcal{K}. \quad (7)$$

This is a system of K linear equations (one for each factor) with $N-n$ unknowns (the futures positions $q_m^{\pi}(t)$'s). When the number of factors driving the futures curve is equal to or greater than the number of futures contracts (i.e., $K \geq N-n$), setting $q_m^{\pi}(t) = \Delta_m^{\pi}(t)$ for all $m\in\mathcal{N}_n$ solves the system of linear equations (7). Otherwise (i.e., $K < N-n$), this system of linear equations is underdetermined and multiple replicating positions exist. We consider two particular solutions.

Bucket Hedging. BH sets the replicating positions equal to the deltas, i.e., $q_m^{\pi}(t) = \Delta_m^{\pi}(t)$ for all $m\in\mathcal{N}_n$, which provides a solution to the system of linear equations (7). BH does not depend directly on the factor loadings $\sigma_{m,k,n-1}$'s; rather, the futures curve factor structure only affects BH indirectly through the deltas. As a result, the bucket hedges for different futures price models with different numbers of factors differ only to the extent that the corresponding deltas differ.

Factor Hedging and Naïve Factor Hedging. FH only takes positions in as many futures contracts as the number of futures curve factors, K. Consequently, FH relies strongly on the assumed futures factor structure. In principle, any K contracts can be used as long as their factor loadings are linearly independent. We include in set $\mathcal{H}_n$ the maturity labels of the K futures contracts used for hedging, and we collect the remaining unused maturity labels in set $\bar{\mathcal{H}}_n \equiv \mathcal{N}_n\backslash\mathcal{H}_n$. We set $q_m^{\pi}(t) := 0$ for all $m\in\bar{\mathcal{H}}_n$ and then include in the column vector $\mathbf{q}^{\pi,\mathcal{H}_n}(t)$ the positions of the contracts with maturity labels in set $\mathcal{H}_n$ that solve the remaining system of K linear equations in K unknowns from (7). The solution, given in Proposition 1, uses the following notation: at time $t\in[T_{n-1},T_n)$, $\operatorname{diag}(\check{\mathbf{F}}^{\mathcal{H}_n}(t))$ and $\operatorname{diag}(\check{\mathbf{F}}^{\bar{\mathcal{H}}_n}(t))$ are diagonal matrices corresponding to the time t futures prices with maturity labels in sets $\mathcal{H}_n$ and $\bar{\mathcal{H}}_n$; $\Delta^{\pi,\mathcal{H}_n}(t)$ and $\Delta^{\pi,\bar{\mathcal{H}}_n}(t)$ are column vectors of deltas, which together form $\Delta^{\pi}(t)$, corresponding to maturity labels in sets $\mathcal{H}_n$ and $\bar{\mathcal{H}}_n$; and B_{n-1} and E_{n-1} are $K\times K$ and $(N-n-K)\times K$ submatrices of the factor loading coefficients $\sigma_{m,k,n-1}$'s for maturity labels in sets $\mathcal{H}_n$ and $\bar{\mathcal{H}}_n$. We denote transposition by a superscript T.

Proposition 1 (FH Positions). *Suppose that B_{n-1} is invertible. Pick $n\in\mathcal{N}$ such that $N-n>K$ and $t\in[T_{n-1},T_n)$. The replicating FH positions in futures with maturity labels in set $\mathcal{H}_n$ are*

$$\mathbf{q}^{\pi,\mathcal{H}_n}(t) = \Delta^{\pi,\mathcal{H}_n}(t) + \operatorname{diag}^{-1}(\check{\mathbf{F}}^{\mathcal{H}_n}(t))(B_{n-1}^{\mathsf{T}})^{-1}E_{n-1}^{\mathsf{T}} \cdot \operatorname{diag}(\check{\mathbf{F}}^{\bar{\mathcal{H}}_n}(t))\Delta^{\pi,\bar{\mathcal{H}}_n}(t). \quad (8)$$

Each position in a traded futures contract is the sum of the delta for that futures contract and a linear combination of the deltas corresponding to the untraded futures contracts. In contrast to BH, FH depends on the futures curve factor structure directly via the submatrices of factor loadings B_{n-1} and E_{n-1} as well as indirectly via the deltas.

Different implementations of FH use different subsets of futures to delta hedge. NFH is one simple implementation. At time $t\in[T_{n-1},T_n)$, NFH takes positions in the K shortest maturity futures contracts (i.e., it sets $\mathcal{H}_n$ equal to $\{n,\ldots,n+K-1\}$). As documented in §5, these are typically the most liquid contracts. However, as discussed in §7.1, the performance of NFH may be fragile in the presence of model error. We present alternative hedges in §7.1 to rectify this problem.

The storage deltas are critical inputs to both BH and FH. Typically, they must be estimated numerically. Assumption 1 gives a set of sufficient conditions that allows us to extend to storage the pathwise approach (see, e.g., §7.2 in Glasserman 2004) for unbiased Monte Carlo delta estimation.

Assumption 1 (Lipschitz Continuity and Derivative Characterization). (a) *In every stage $n\in\mathcal{N}$, for each given inventory $x\in\mathcal{X}$, the function $V_n^{\pi}(x,\mathbf{F}_n)$ is Lipschitz continuous in the futures curve $\mathbf{F}_n\in\Re_+^{N-n}$; i.e., there exists $\mathsf{L}_n^{\pi}(x)\in\Re_+$ such that $|V_n^{\pi}(x,\mathbf{F}_n^2) - V_n^{\pi}(x,\mathbf{F}_n^1)| \leq \mathsf{L}_n^{\pi}(x)\sum_{m=n}^{N}|F_{n,m}^2 - F_{n,m}^1|$ for all $\mathbf{F}_n^1,\mathbf{F}_n^2\in\Re_+^{N-n}$.* (b) *Moreover, at every futures curve $\bar{\mathbf{F}}_n\in\Re_+^{N-n}$, where $V_n^{\pi}(x,\mathbf{F}_n)$ is differentiable with respect to each futures price in $\mathbf{F}_n$, the decision rule $A_n^{\pi}(x,\bar{\mathbf{F}}_n)$ given the policy $\boldsymbol{\pi}$ has a unique action denoted by $a_n^{\pi}(x,\bar{\mathbf{F}}_n)$, and for all $m\in\mathcal{N}_n$*

$$\left.\frac{\partial V_n^{\pi}(x_n,\mathbf{F}_n)}{\partial F_{n,m}}\right|_{\mathbf{F}_n=\bar{\mathbf{F}}_n} = \left.\frac{\partial p(a_n^{\pi}(x_n,\bar{\mathbf{F}}_n),s_n)}{\partial F_{n,m}}\right|_{s_n=\bar{F}_{n,n}} + \left.\frac{\partial\delta\mathbb{E}[V_{n+1}^{\pi}(x_n - a_n^{\pi}(x_n,\bar{\mathbf{F}}_n),\mathbf{F}_{n+1})\mid\mathbf{F}_n]}{\partial F_{n,m}}\right|_{\mathbf{F}_n=\bar{\mathbf{F}}_n}.$$

The second part of this assumption roughly states that the feasible policy π has constant actions in neighborhoods of differentiability of its value function with respect to each element of the futures curve. As shown in Lemma 3 in Online Appendix E, under a mild assumption, an optimal operating policy satisfies Assumption 1. We denote by $1\{\mathscr{E}\}$ the indicator function of event $\mathscr{E}$.

Proposition 2 (Pathwise Deltas). *Under Assumption 1, for every $n \in \mathcal{N} \setminus \{0\}$, it holds that*

$$\begin{aligned}\Delta_m^\pi(t, x_n, \check{\mathbf{F}}(t)) = \frac{\bar{\delta}(t, T_m)}{F(t, T_m)} \mathbb{E}\big[\big(\phi^I 1\{A_m^\pi(x_m^\pi, \mathbf{F}_m) < 0\} \\ + \phi^W 1\{A_m^\pi(x_m^\pi, \mathbf{F}_m) > 0\}\big) \\ \cdot s_m A_m^\pi(x_m^\pi, \mathbf{F}_m) \mid x_n, \check{\mathbf{F}}(t)\big] \quad (9)\end{aligned}$$

for all $m \in \mathcal{N}_n$, $x_n \in \mathscr{X}$, $t \in [T_{n-1}, T_n)$, and $\check{\mathbf{F}}(t) \in \Re_+^{N-n}$.

Expression (9) can be used to estimate deltas by Monte Carlo simulation simultaneously with the valuation of a policy π. In §8 we use this expression with the RI policy, even though the corresponding value function can violate the Lipschitz condition in Assumption 1. This calculation may yield biased delta estimates but can be implemented efficiently. Our computational results in §8 suggest, however, that any bias in our calculated deltas is small. This outcome is expected given that Proposition 2 holds for an optimal policy (see Proposition 9 in Online Appendix B), and Lai et al. (2010) find that the RI policy is near optimal (see also Thompson 2012, Wu et al. 2012). This approach seems better than calculating deltas by resimulation, which is both biased and computationally expensive (Glasserman 2004, §7.1).

The term on the right-hand side of (9) can be interpreted as being proportional to a "weighted average" of the action taken by a feasible policy in a given stage and state, as now explained. The ratio $s_m/F(t, T_m)$ is a nonnegative random variable with mean one (because $\mathbb{E}[s_m \mid F(t, T_m)] = F(t, T_m)$). This ratio weighs the amount of commodity traded by policy π in state $(x_m^\pi, \mathbf{F}_m)$ at stage n; i.e., $(\phi^I 1\{A_m^\pi(x_m^\pi, \mathbf{F}_m) < 0\} + \phi^W 1\{A_m^\pi(x_m^\pi, \mathbf{F}_m) > 0\}) A_m^\pi(x_m^\pi, \mathbf{F}_m)$. The expression on the right-hand side of (9) is thus a weighted risk-neutral expectation of the amount of commodity traded by policy π in this stage and state discounted by $\bar{\delta}(t, T_m)$.

Proposition 3 bounds the storage deltas. This result informs us about the size of the BH trading positions, assuming BH uses exact deltas, as pointed out in §7.1, and lets us relate the size of the RI policy deltas that we compute in §§8.2.2 and 8.2.3 to those of a policy that does satisfy Assumption 1—in particular, an optimal policy (under a mild assumption satisfied in our numerical analysis).

Proposition 3 (Bounds on Deltas). *Under Assumption 1, for every $n \in \mathcal{N} \setminus \{0\}$, it holds that $\bar{\delta}(t, T_m)\phi^I C^I \le \Delta_m^\pi(t, x, \check{\mathbf{F}}(t)) \le \bar{\delta}(t, T_m)\phi^W C^W$ for all $t \in [T_{n-1}, T_n)$, $m \in \mathcal{N}_n$, $x \in \mathscr{X}$, and $\check{\mathbf{F}}(t) \in \Re_+^{N-n}$.*

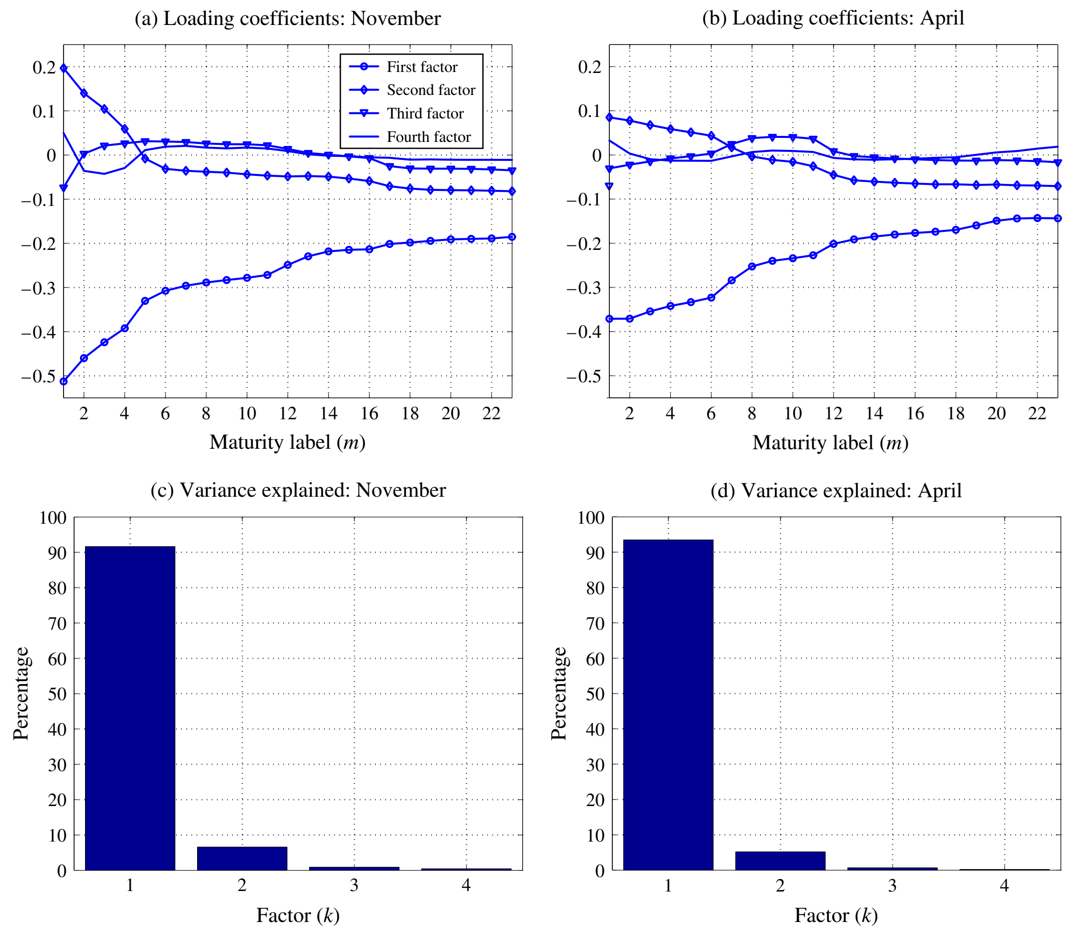
5. Calibration

Consider a merchant trader who wants to implement storage valuation and hedging based on an assumed futures price term-structure model as in (2). Consistent with industry practice (Clewlow and Strickland 2000, §8.6; Blanco et al. 2002), we estimate the constant intramonth factor loadings of the futures price evolution model (2) using 12 separate monthly principle component analyses (PCAs). We implement this estimation using daily closing NYMEX natural gas futures prices from January 1997 to December 2012.

We distinguish between months in the heating season (November–March) and the rest of the year. Panels (a) and (b) of Figure 1 display the estimated November and April loading coefficients, respectively, for 23 monthly maturities for the first four factors (Online Appendix C includes the corresponding charts for all the monthly PCAs). Typically, the factor loading coefficients decrease rapidly, with the magnitude of the coefficients of a given factor being less than half of the one of the previous factor. The first factor changes the slope of the natural gas futures term structure. In particular, the first factor induces positively correlated shocks along the futures term structure (since the coefficients are of the same sign), but it moves the short end of the term structure more than the long end (since the coefficients decline in magnitude in time to delivery). The second factor also shocks the slope of the futures term structure, but the changes at the long and short ends of the curve are negatively correlated (since the coefficients change sign). The third factor shocks the curvature of the futures term structure (since the short-term and long-term futures coefficients have the opposite sign of the intermediate-term coefficients). The fourth factor is typical of the remaining PCA factors. They induce small irregular "squiggles" along the futures term structure (since the signs of the coefficients change multiple times). Thus, whereas shocks to the larger factors are somewhat persistent, shocks to the smaller factors are highly transitory. Seasonality in the shapes of the loading coefficients also appears to change across different months.

Panels (c) and (d) of Figure 1 display the percentages of the total variance of the observed price log returns explained by the first four factors for the November and April PCAs, respectively (the corresponding charts for all the monthly PCAs are available in Online Appendix C). The first factor captures roughly 88.7%–93.5% of this variance, the second factor 5.0%–7.6%, the third factor 0.6%–3.0%, and the

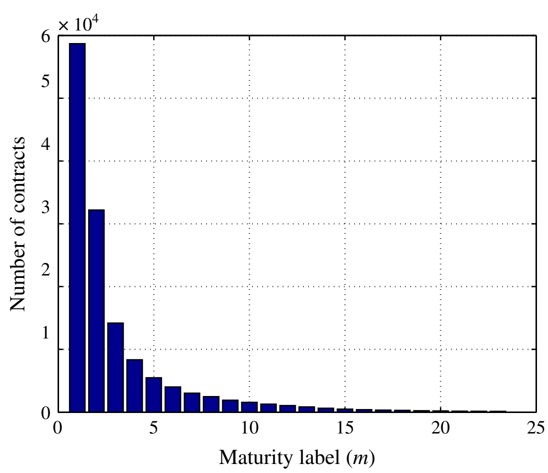
Figure 1 (Color online) The Estimated Loading Coefficients of and the Percentage of the Total Futures Price Log-Return Variance Explained by the First Four Factors in the November and April PCAs over the Period from January 1997 to December 2012



fourth factor always less than 1%. The first three factors explain just over 98% of the total variance in the observed futures price log returns in each month.

Our calibration ignores liquidity concerns because it is based on closing futures prices. Figure 2 displays the average daily trading volume of the natural gas futures contracts for the first 23 maturities from January 1997 to December 2012. As is apparent from this figure, longer-maturity futures suffer from lower liquidity, which could result in higher transaction costs (wider bid–ask spreads). Although we do not consider this issue in our calibration, we do model the impact of limited liquidity on hedging in §6.

Figure 2 (Color online) The Average Daily Trading Volume of the First 23 Natural Gas Futures Contracts over the Period from January 1997 Through December 2012



6. Term-Structure Model Error

The PCA estimation results in §5 are the starting point for our analysis of empirically likely model error. Merchant storage traders use PCA estimations to calibrate their *assumed model* of futures curve dynamics on which they then base their valuation and hedging decisions. However, they know PCA estimation contains a variety of likely model errors. First, there is *parameter error* in the estimated factor loading coefficients. Second, it may be unclear how to treat the estimated PCA factors themselves. If estimated futures curve randomness is excluded from the assumed model (2), then the model may have *missing factors* (if bona fide factors are excluded) or *missing noise* (if the excluded randomness is really due to maturity-specific idiosyncratic shocks rather than common factors). If, instead, estimated factors that are not bona fide are mistakenly included, the assumed model (2) includes *spurious factor* error. Third, the estimated PCA may be structurally inconsistent with the true process because of omitted jumps, stochastic volatility, or incorrect functional specifications for the factor loadings.

These model errors all have potentially adverse effects on storage valuation and hedging. We disentangle the potential effects of different term-structure model errors on storage valuation and hedging in a controlled setting. In particular, we consider pairs of a hypothetically true model and a model assumed by a merchant that are both plausible given the PCA estimates, and then we measure how wrong the merchant can be. There are, of course, many possible such model pairs.

As possible hypothetically true models, we consider two empirically plausible families of futures price models given the PCA estimation in §5. The first family consists of *common factor models* in which different numbers of the PCA-estimated common factors are considered to be bona fide and the rest of the estimated factors are viewed as spurious and are hence ignored. We let $K^\diamond$ denote the number of such bona fide factors, which we include in set $\mathcal{K}^\diamond$. Specifically, under this family the futures price $F(t, T_m)$ evolves during the time interval $[T_n, T_{n+1})$ according to (2) but with set $\mathcal{K}$ replaced by set $\mathcal{K}^\diamond$. The second family consists of *common factor plus noise models* that include both $K^\diamond$ bona fide common factors and maturity-specific (idiosyncratic) noise factors that capture the empirical variability associated with the $(N-1)-K^\diamond$ excluded common factors. Specifically, under this family the dynamics of the futures price $F(t, T_m)$ during the time interval $[T_n, T_{n+1})$ are

$$\frac{dF(t, T_m)}{F(t, T_m)} = \sum_{k\in\mathcal{K}^\diamond} \sigma_{m,k,n} dZ_k(t) + \check{\sigma}_{m,n} d\check{Z}_m(t), \qquad (10)$$

where $\check{\sigma}_{m,n}$ is the idiosyncratic volatility for maturity T_m and $d\check{Z}_m(t)$ is an uncorrelated standard Brownian motion increment specific to this maturity. Given the PCA estimates of the factor loadings $\sigma_{m,k,n}$'s for the hypothetical common factors in (10), we calibrate the volatility coefficients $\check{\sigma}_{m,n}$ to match the maturity T_m futures price log-return variance unexplained by the included common factors. In a common factor model, the futures price log-return variances are less than (or equal to) the empirical variances if $K^\diamond < N-1$ (or $K^\diamond = N-1$). In a common factor plus noise model, the modeled and empirical variances are equal, but the covariance structures differ.

As possible models assumed by the merchant, we suppose concretely that the merchant uses a common factor model (2) with some number K of common factors, where K may be less than the number of estimated PCA factors (i.e., the merchant believes that some estimated factors are spurious) and is potentially different from the unknown (to the trader) hypothetically true number of common factors. The assumed factor loadings for these K factors are, again, the same as the empirical PCA estimates.

Given the PCA factor sizes, see, e.g., panels (c) and (d) of Figure 1, our numerical analysis focuses on a natural form of model uncertainty concerning the statistically small residual common factors after the first three large common factors are included. We specifically posit that a merchant uses an *assumed* common factor model with K factors while the actual futures prices are generated by one of the following three hypothetically true reference models: (a) a $K^\diamond = 3$ common factor model (that captures more than 98% of the empirical variance of the futures price log returns), (b) a $K^\diamond = N-1$ common factor model (that takes the full empirical variance as true), and (c) a $K^\diamond = 3$ common factor plus noise model (that attributes the 1%–2% PCA residual variance to uncorrelated noise).

The framework in Figure 3 organizes the resulting term-structure model errors for different combinations of omitted or spurious factors and missing noise. For example, using an assumed common factor model with K equal to 1 when the hypothetically true model has $K^\diamond$ equal to 3 results in model error because of two missing large common factors, whereas assuming K equal to 4 when the hypothetically true model has $K^\diamond$ equal to 3 and includes noise results in model error because of the inclusion of one spurious common factor as well as missing noise.

Using this model error framework, we conduct a series of *what-if* Monte Carlo simulations to assess the likely potential impact of different model errors on natural gas storage valuation and hedging. That is, we estimate the value of storage under various assumed common factor models and the three hypothetically true models in our framework, and we compare the resulting storage valuations. (We obtain the storage valuation for each given model using Monte

Figure 3 **Term-Structure Model Error Framework**

K	$K^\diamond = 3$	$K^\diamond = N-1$	$K^\diamond = 3$ plus noise
1	Missing large factors	Missing large and small factors	Missing large factors and noise
2			
3	No error	Missing small factors	Missing noise
4	Spurious factors		Spurious factors and missing noise
$\vdots$			
$N-1$		No error	

Carlo samples specific to this model, because valuation is done *before* using a policy on a price sample path.) We also apply various hedging strategies based on different assumed common factor models and measure their performance on futures curve paths generated by sampling from the hypothetically true models (because hedging is done along a price sample path). We perform this valuation and hedging model error analysis in §§8.1 and 8.2.2, respectively. We also measure the combined effect of actual model errors on storage hedging in §8.2.3 using historical futures curve paths.

The low liquidity for futures with long-dated maturities discussed in §5 can limit the availability of these contracts for hedging, thereby creating an important friction for hedging in practice. In our analysis we allow for the possibility that illiquid futures cannot be traded, and we refer to this friction as the *missing-contracts* friction. This friction interacts with model error because some hedging strategies may become infeasible when they require trading in illiquid contracts. In §7.3 we propose a hedging strategy to manage the missing-contracts friction.

7. Coping with Term-Structure Model Error in Hedging

In this section we discuss practical approaches to mitigate the negative effect of term-structure model error on hedging, also taking into account the missing-contracts friction. We propose FTFH in §7.1, analyze the relationship between BH and MVH in §7.2, and present CMVH in §7.3.

7.1. Fine-Tuned Factor Hedging

The size of the futures trading positions is, in theory, irrelevant in the absence of model error. Size becomes important, however, if large trading positions amplify unmodeled randomness in the realized price changes. Hence, we expect hedges with smaller positions to be more robust to such model errors. Controlling the size of the hedges is relevant for FH. The BH positions are fixed by the storage deltas, which, if determined exactly, are bounded as in Proposition 3.

NFH makes no attempt to control the size of the trading positions and thus may take extreme positions. We propose FTFH to control the size of the FH trading positions. Since determining the set of contracts with the smallest possible FH positions requires solving a potentially difficult minimum norm constrained optimization problem at each trading time t and in each state $(x, \check{\mathbf{F}}(t))$, FTFH aims heuristically to take small futures trading positions. Based on Proposition 1, if we express the inverse of the matrix B_{n-1} in (8) as $\zeta_{n-1}^{\mathsf{T}}/|B_{n-1}|$, where ζ_{n-1} is the matrix of cofactors of B_{n-1} (if $K = 1$, then the cofactor matrix is just the scalar 1), then we can rewrite (8) as

$$\mathbf{q}^{\pi,\mathscr{H}_n}(t) = \Delta^{\pi,\mathscr{H}_n}(t) + \operatorname{diag}^{-1}(\check{\mathbf{F}}^{\mathscr{H}_n}(t))\frac{\zeta_{n-1}}{|B_{n-1}|}E_{n-1}^{\mathsf{T}} \cdot \operatorname{diag}(\check{\mathbf{F}}^{\bar{\mathscr{H}}_n}(t))\Delta^{\pi,\bar{\mathscr{H}}_n}(t). \tag{11}$$

FTFH selects the set of K contracts to trade that yield the largest determinant $|B_{n-1}|$. This method is computationally efficient because each matrix B_{n-1}, which we refer to as *position generating matrix*, is chosen once up front (at time 0), since this choice only depends on the monthly factor loadings.

7.2. Relationship Between Bucket Hedging and Minimum Variance Hedging

Both BH and FH rely on replication (see §4). However, in the presence of term-structure model error, an alternative is a minimum variance hedge. We show here that BH is a minimum approximate-variance method. This result thus suggests that BH can be useful even in the presence of term-structure model error.

Fix date $t \in [T_{n-1}, T_n)$. Given $\check{\mathbf{F}}(t)$, the variance of the sum of the change between times t and $t+\Delta t$ in the *true* value of storage $U^{\pi,\diamond}(\cdot)$ (rather than the assumed value of storage $U^{\pi}(\cdot)$), and the time $t+\Delta t$ cash flow from *shorting* at time t the futures positions in the column vector $\mathbf{q}(t)$ is

$$\mathbb{VAR}\big(U^{\pi,\diamond}(t+\Delta t) - U^{\pi,\diamond}(t) - (\mathbf{q}(t))^{\mathsf{T}}(\check{\mathbf{F}}^{\diamond}(t+\Delta t) - \check{\mathbf{F}}(t)) \mid \check{\mathbf{F}}(t)\big), \tag{12}$$

where $\check{\mathbf{F}}^{\diamond}(t+\Delta t)$ is the *true* random futures curve at time $t+\Delta t$. Assuming a hypothetically true futures curve evolution model of the type discussed in §6 and letting the vector of true model deltas be $\Delta^{\pi,\diamond}(t)$, a first-order Taylor series approximation for the random change in the true storage value over $[t, t+\Delta t]$ is $(\Delta^{\pi,\diamond}(t))^{\mathsf{T}}(\check{\mathbf{F}}^{\diamond}(t+\Delta t) - \check{\mathbf{F}}(t))$. The variance in (12) can be approximated as $[\mathbf{q}(t) - \Delta^{\pi,\diamond}(t)]^{\mathsf{T}}\mathbb{COV}(\check{\mathbf{F}}^{\diamond}(t+\Delta t) \mid \check{\mathbf{F}}(t))[\mathbf{q}(t) - \Delta^{\pi,\diamond}(t)]$, where $\mathbb{COV}(\check{\mathbf{F}}^{\diamond}(t+\Delta t) \mid \check{\mathbf{F}}(t))$ is the covariance matrix of $\check{\mathbf{F}}^{\diamond}(t+\Delta t)$ conditional on $\check{\mathbf{F}}(t)$. The resulting MVH model is

$$\min_{\mathbf{q}(t)}[\mathbf{q}(t) - \Delta^{\pi,\diamond}(t)]^{\mathsf{T}}\mathbb{COV}(\check{\mathbf{F}}^{\diamond}(t+\Delta t) \mid \check{\mathbf{F}}(t)) \cdot [\mathbf{q}(t) - \Delta^{\pi,\diamond}(t)]. \tag{13}$$

There are two difficulties with this model: both the vector of deltas $\Delta^{\pi,\diamond}(t)$ and the covariance matrix $\mathbb{COV}(\check{\mathbf{F}}^{\diamond}(t+\Delta t) \mid \check{\mathbf{F}}(t))$ are unknown to the merchant. Despite the second difficulty, Proposition 4 shows that BH is an optimal MVH method when the unknown deltas $\Delta^{\pi,\diamond}(t)$ are approximated with the deltas obtained under an assumed K common factor model, $\Delta^{\pi}(t)$.

Proposition 4 (BH Optimality). *BH optimally solves the version of optimization model* (13) *obtained by approximating $\Delta^{\pi,\diamond}(t)$ with $\Delta^{\pi}(t)$.*

7.3. Constrained Minimum Variance Hedging

With long-tenor storage contracts, the missing-contracts friction renders BH impractical and also invalidates FTFH if some of the futures contracts chosen by this method are long-dated and, hence, illiquid. We thus explore a *constrained* version of the minimum variance model (13) using a given set of liquid futures contracts. We model the missing-contracts friction by supposing that on date $t \in [T_{n-1}, T_n)$ only $L < N - n$ futures contracts are tradable, and hence, the positions in all of the remaining $N - n - L$ contracts must be zero. We let $\mathcal{Q}_n^L$ denote the set of all futures position vectors that satisfy this condition.

The CMVH positions depend on the chosen futures price covariance matrix. We proceed by using the covariance matrix $\mathbb{COV}^{N-1}(\check{\mathbf{F}}(t+\Delta t) \mid \check{\mathbf{F}}(t))$ of the estimated full-dimensional specification of price model (2), i.e., the specification of this price model with $N-1$ common factors on date t. (Other choices are possible, such as the common factor plus noise model in our model error framework; using the covariance matrix obtained from this model yielded results similar to the ones discussed in §§8.2.2 and 8.2.3.) In particular, the entry in position (l, m) of $\mathbb{COV}^{N-1}(\check{\mathbf{F}}(t+\Delta t) \mid \check{\mathbf{F}}(t))$ is

$$F(t, T_l)F(t, T_m)\left[\exp\left(\Delta t \sum_{k=1}^{N-1} \sigma_{l,k,n-1}\sigma_{m,k,n-1}\right) - 1\right]. \quad (14)$$

The CMVH positions solve

$$\min_{\mathbf{q}(t)\in\mathcal{Q}_n^L} [\mathbf{q}(t) - \Delta^{\pi}(t)]^{\mathsf{T}} \mathbb{COV}^{N-1}(\check{\mathbf{F}}(t+\Delta t) \mid \check{\mathbf{F}}(t)) \cdot [\mathbf{q}(t) - \Delta^{\pi}(t)]. \quad (15)$$

Proposition 5 characterizes the optimal solution to (15). Define ξ_{n-1}^{N-1} as the matrix that includes in position (l, m) the term in square brackets in (14). We include in set $\mathcal{L}_n$ the labels of the maturities of the L liquid futures that can be used for hedging and in $\bar{\mathcal{L}}_n$ the labels of the remaining futures. The matrices $\xi_{n-1}^{N-1,\mathcal{L}_n}$ and $\xi_{n-1}^{N-1,\bar{\mathcal{L}}_n}$ are the L rows and L columns and the L rows and the $N-n-L$ columns of the matrix ξ_{n-1}^{N-1} corresponding to the sets $\mathcal{L}_n$ and $\bar{\mathcal{L}}_n$, respectively. The vectors $\check{\mathbf{F}}^{\mathcal{L}_n}(t)$ and $\check{\mathbf{F}}^{\bar{\mathcal{L}}_n}(t)$ and $\Delta^{\pi,\mathcal{L}_n}(t)$ and $\Delta^{\pi,\bar{\mathcal{L}}_n}(t)$ include the prices and deltas corresponding to futures with maturity labels in sets $\mathcal{L}_n$ and $\bar{\mathcal{L}}_n$, respectively. The column vector $\mathbf{q}^{\pi,\mathcal{L}_n}(t)$ contains the CMVH trading positions in futures with maturity labels in set $\mathcal{L}_n$.

Proposition 5 (CMVH Positions). *If the matrix $\xi_{n-1}^{N-1,\mathcal{L}_n}$ is positive definite, the CMVH futures positions for maturity labels in set $\mathcal{L}_n$ are*

$$\mathbf{q}^{\pi,\mathcal{L}_n}(t) = \Delta^{\pi,\mathcal{L}_n}(t) + \mathrm{diag}^{-1}(\check{\mathbf{F}}^{\mathcal{L}_n}(t)) \left(\xi_{n-1}^{N-1,\mathcal{L}_n}\right)^{-1} \cdot \xi_{n-1}^{N-1,\bar{\mathcal{L}}_n} \mathrm{diag}(\check{\mathbf{F}}^{\bar{\mathcal{L}}_n}(t)) \Delta^{\pi,\bar{\mathcal{L}}_n}(t). \quad (16)$$

The CMVH positions in (16) superficially resemble the positions in (8) for NFH when CMVH and NFH trade the same futures contracts. In general, however, the two solutions are different even when $\mathcal{L}_n = \mathcal{K}_n$. Proposition 6 shows that these solutions are approximately equal in a specific case. Let $\mathbb{COV}(\check{\mathbf{F}}(t+\Delta t) \mid \check{\mathbf{F}}(t))$ be the analogue of $\mathbb{COV}^{N-1}(\check{\mathbf{F}}(t+\Delta t) \mid \check{\mathbf{F}}(t))$ for a K factor model.

Proposition 6 (Approximate Equivalence). *Suppose that the set $\mathcal{L}_n$ of maturity labels for the tradable contracts includes the labels of the K shortest maturity contracts ($\mathcal{L}_n = \mathcal{K}_n$), the matrix $\mathbb{COV}^{N-1}(\check{\mathbf{F}}(t+\Delta t) \mid \check{\mathbf{F}}(t))$ in (15) is replaced by the matrix $\mathbb{COV}(\check{\mathbf{F}}(t+\Delta t) \mid \check{\mathbf{F}}(t))$, and the NFH position generating matrix is invertible. If in this version of model (15) each element of the matrix $\mathbb{COV}(\check{\mathbf{F}}(t+\Delta t) \mid \check{\mathbf{F}}(t))$ is approximated by its first-order Taylor expansion around Δt equal to 0, then the resulting CMVH positions are equal to the NFH positions.*

We expect the CMVH positions in (16) to perform better than the positions in (8) for NFH in the presence of omitted factors or noise. In this situation, the NFH positions can be extreme, as previously mentioned, because NFH attempts to fully eliminate assumed variability while being oblivious to omitted variability. By contrast, when the rank of the assumed covariance matrix is greater than the number of tradable futures contracts, our implementation of CMVH always takes into account the impact of futures positions on some amount of unavoidable residual futures randomness, which will tend to keep the futures position sizes from being too large.

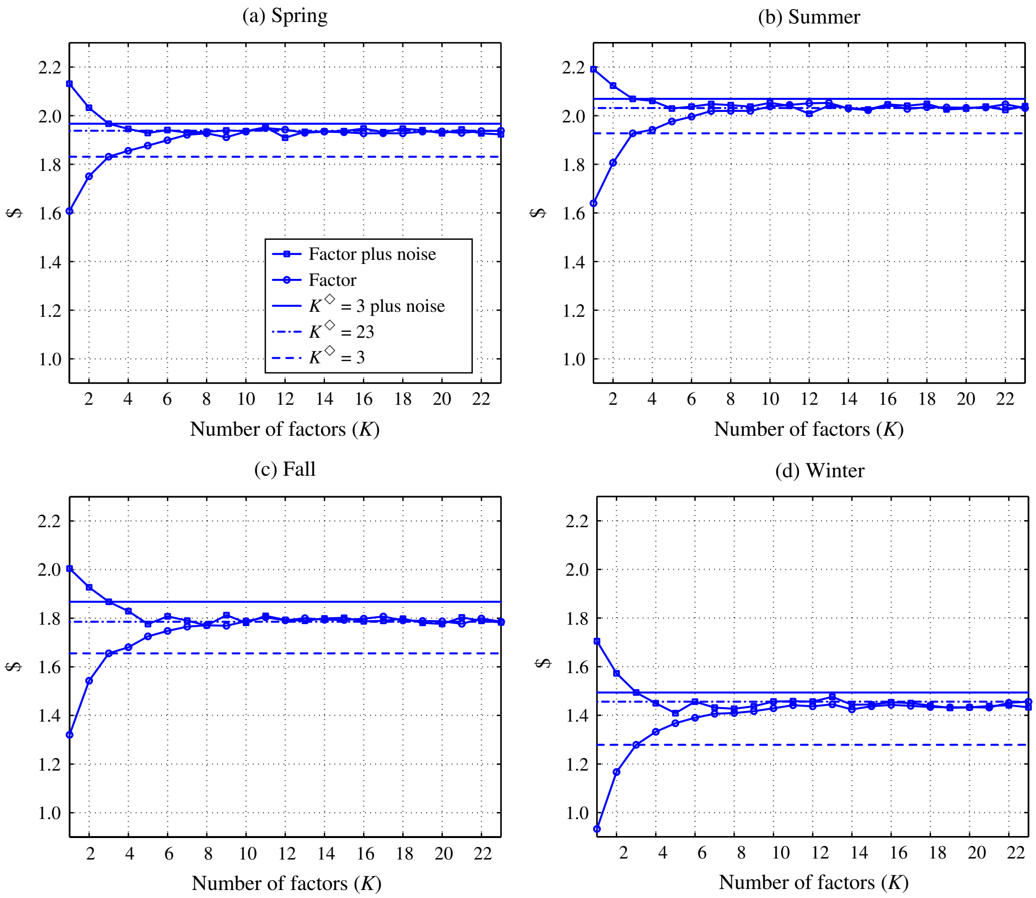
8. Numerical Analysis of Term-Structure Model Error

In this section we quantify the impact of futures term-structure model error on storage valuation (§8.1) and hedging (§8.2), also assessing the performance of the hedging methods proposed in §7.

8.1. Valuation Results

The specific gas storage contracts we consider are related to the 24-month gas storage valuation analysis of Lai et al. (2010). We normalize the maximum storage space, $\bar{x}$, to 1 mmBtu and set the initial inventory, x_0, equal to 0; the injection capacity C^I and withdrawal capacity C^W equal to -0.3 and 0.6 mmBtu per month, respectively (i.e., 3.33 months to fill up and 1.67 months to empty); the injection marginal cost c^I and withdrawal marginal cost c^W equal to \$0.02/mmBtu and \$0.01/mmBtu, respectively; and the injection fuel loss ϕ^I and withdrawal fuel loss ϕ^W equal to 1.01 and 0.99, respectively. These parameters are realistic for natural gas storage contracts (Lai et al. 2010).

Figure 4 (Color online) The Value of Storage Under Different Term-Structure Models on Different Valuation Dates

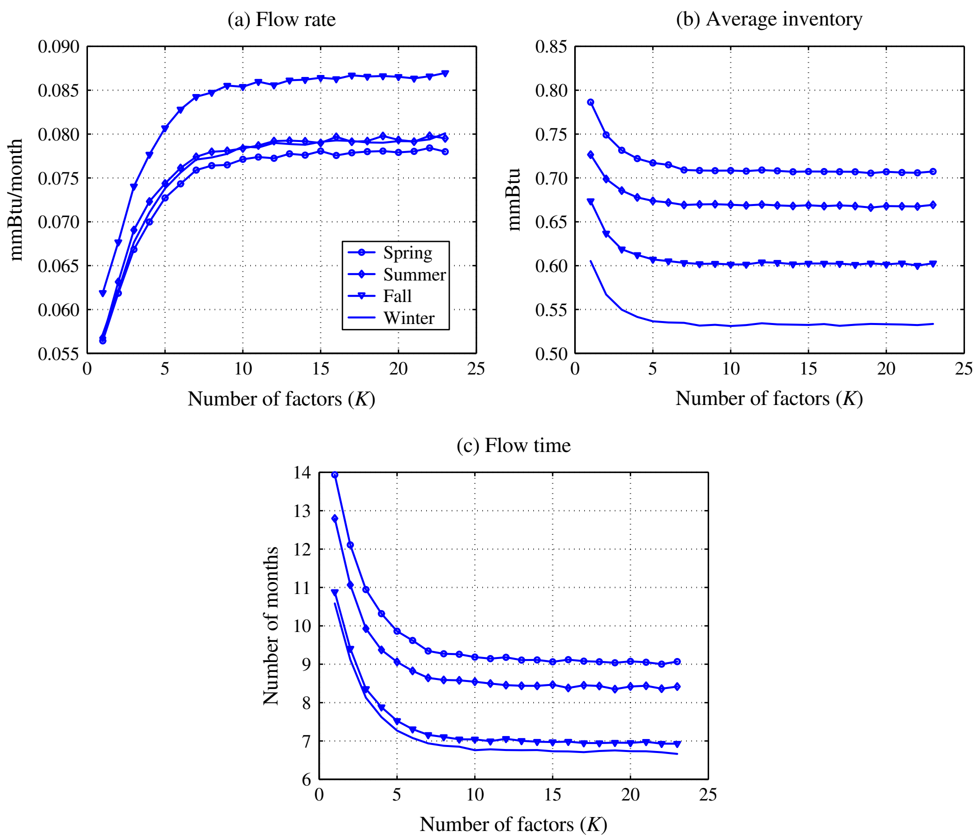


Note. The horizontal lines correspond to the three hypothetically true models in our model error framework in Figure 3.

We first investigate the valuation of storage using the RI policy, $V_0^{\mathrm{RI}}(0, \mathbf{F}_0)$, given different term-structure models. We consider four valuation dates corresponding to the first trading day for different seasons of the year 2012—March 1 (spring), June 1 (summer), September 4 (fall), and December 3 (winter)—with annualized short-term risk-free Treasury rates equal to the then prevailing rates: 0.18%, 0.17%, 0.16%, and 0.18%, respectively. Spring and winter are part of the heating season. Summer and fall are not. We assume that the storage contract starts on the valuation date (i.e., there are $N = 24$ dates with cash flows). Figure 4 shows the range of possible storage valuations using different term-structure models. The Monte Carlo standard errors for these valuations are between \$0.013 and \$0.016 (based on 10,000 simulated futures price paths). Compared with the other seasons, the storage values starting in winter are lower because winter is a high-price season when inventory is typically sold, but here the initial inventory is assumed to be zero. As expected given our PCA results, the value of storage based on common factor models initially increases with the number of factors but starts leveling off once the most important statistical factors are included. Also as expected, the storage valuations based on the family of common factor plus noise models are (a) initially higher than the common factor model valuations with the same number of common factors (since there is more transitory price variability) and (b) decreasing as the most important common factors are added (and idiosyncratic variability is replaced with correlated variability) and (c) eventually largely flattening out.

We can provide an operational explanation for the largely increasing value of storage in Figure 4 as the number of assumed factors in price model (2) increases. Specifically, for each assumed common factor model we compute the average amount of natural gas sold per stage (i.e., the flow rate) across all the 24 stages and the states encountered during the Monte Carlo simulation used for valuation as well as the average inventory across these stages and states. The ratio of average inventory to flow rate is the average number of months that a given unit of purchased natural gas spends in inventory before being sold back to the market (i.e., the flow time from Little's law). Panels (a)–(c) of Figure 5 show the estimated

Figure 5 (Color online) The Operational Effect of Increasing the Number of Factors in Price Model (2)



flow rate, average inventory, and flow time, respectively. The flow rate increases and the average inventory decreases, and thus the flow time shortens when increasing the number of factors, K. (The increase in the flow rate is due to the increased transitory price variability to trade on with more factors.) These patterns taper off at about the same number of factors when the storage value levels off in Figure 4, as a result of the declining incremental magnitude of this transitory variability. The consequent increased storage value is brought about by trading more natural gas, holding less inventory, and, hence, increasing the frequency of trading.

Given uncertainty about the true futures term-structure model, merchants are uncertain about which valuation is correct. Using our model error framework (see Figure 3), Table 1 displays the ratios of the storage valuations corresponding to a subset of assumed models with different numbers of common factors, K, relative to the valuations of the three hypothetically true models. A ratio larger (smaller) than 1 means that the assumed model overvalues (undervalues) storage relative to the hypothetically true valuation. As expected, omitting large common factors (e.g., $K=1$ and $K^\diamond=3$) leads to substantial undervaluation. More surprisingly, however, the model errors associated with the last 1%–2% of the empirical futures price log-return variance cause the storage valuations to change by roughly between −14% and +14% depending on whether this residual empirical variance is due to missing noise ($K=3$ and $K^\diamond=3$ plus noise in winter) or spurious factors ($K=23$ and $K^\diamond=3$ in winter). Missing small factors can also lead to substantially different valuations (e.g., $K=3$ and $K^\diamond=23$ in winter). By contrast, model error caused by missing very small factors has a somewhat lesser impact on storage valuation ($K\geq 5$ and $K^\diamond=23$ across seasons).

Table 1 The Impact of Model Error on Storage Valuation: Ratios of the Storage Values Obtained with Different Assumed Factor Models Relative to Three Hypothetically True Term-Structure Models

	$K^\diamond=3$				$K^\diamond=23$				$K^\diamond=3$ plus noise			
K	Sp	Su	Fa	Wi	Sp	Su	Fa	Wi	Sp	Su	Fa	Wi
1	0.88	0.85	0.80	0.73	0.83	0.81	0.74	0.64	0.82	0.79	0.71	0.62
2	0.96	0.94	0.93	0.91	0.90	0.89	0.86	0.80	0.89	0.87	0.83	0.78
3	1.00	1.00	1.00	1.00	0.94	0.95	0.93	0.88	0.93	0.93	0.89	0.86
5	1.03	1.03	1.04	1.07	0.97	0.97	0.97	0.94	0.95	0.95	0.92	0.92
10	1.06	1.06	1.08	1.12	1.00	1.00	1.00	0.98	0.98	0.98	0.96	0.96
15	1.06	1.05	1.08	1.12	1.00	1.00	1.00	0.99	0.98	0.98	0.96	0.96
20	1.06	1.05	1.08	1.12	1.00	1.00	1.00	0.98	0.98	0.98	0.96	0.96
23	1.06	1.05	1.08	1.14	1.00	1.00	1.00	1.00	0.99	0.98	0.96	0.97

Note. Sp, spring; Su, summer; Fa, fall; Wi, winter.

Moreover, adding these very small factors to a factor model effectively adds some noise to a factor model ($K \geq 5$ and $K^{\diamond} = 23$ versus $K \geq 5$ and $K^{\diamond} = 3$ plus noise across seasons). In other words, adding spurious factors to a factor model can be interpreted as a way of adding some noise to this model. However, the true model remains unknown and may be simpler than such a model.

8.2. Hedging Results

We first explain how we measure hedging effectiveness in §8.2.1. We then present our results based on simulated and historical natural gas prices in §§8.2.2 and 8.2.3, respectively.

8.2.1. Measuring Hedging Effectiveness. As natural gas storage is not traded in a liquid market, market prices for natural gas storage are not available to measure hedging effectiveness. Thus, we focus on how effective futures hedging is in reducing the variance of the physical cash flows generated by the RI policy along futures curve sample paths. Given an initial futures curve at time 0, let ω denote a sample path of N futures curve realizations at times T_1 through T_{N-1}. We use futures price paths both from Monte Carlo simulation according to our model error framework (Figure 3) and from historical data.

To simplify bookkeeping, we start the storage contract in stage 1 with no inventory and value it at time 0, so that the initial value of storage is $U^{\mathrm{RI}}(0, 0, \check{\mathbf{F}}(0))$. There are operational (physical) and futures trading cash flows on 23 monthly stages. Denote by $P(\omega)$ the sum $\sum_{n \in \mathcal{N} \setminus \{0\}} \delta^n p(a_n^{\mathrm{RI}}(x_n^{\mathrm{RI}}, \mathbf{F}_n), s_n; \omega)$ of the time 0 discounted values of the physical trading cash flows from time T_1 through time T_{N-1} along path ω, where $a_n^{\mathrm{RI}}(x_n^{\mathrm{RI}}, \mathbf{F}_n)$ is the action taken by the RI policy in state $(x_n^{\mathrm{RI}}, \mathbf{F}_n)$ at stage n. Let $\Psi(\omega)$ be the sum $\sum_{n \in \mathcal{N} \setminus \{0\}} \delta^n \psi_n(\mathbf{q}(T_{n-1}); \omega)$ of the time 0 discounted values of the futures trading cash flows from time T_1 through time T_{N-1}, where $\psi_n(\mathbf{q}(T_{n-1}); \omega)$ is the cash flow $\sum_{m \in \mathcal{N}_n} q_m(T_{n-1})(F_{n,m} - F_{n-1,m})(\omega)$ generated at time T_n along path ω by the futures trading positions $\mathbf{q}(T_{n-1})$. Our analysis is for monthly futures position rebalancing and mark-to-market, but doing this more frequently just involves increasing the simulation granularity.

From a merchant's perspective, the reduction in physical cash flow variance caused by hedging is a natural metric of hedging effectiveness. We denote by $\mathbb{VAR}(P \mid x_1, \mathbf{F}_0)$ and $\mathbb{VAR}(P - \Psi \mid x_1, \mathbf{F}_0)$ the cross-path variances of the total physical cash flows, P, and residual total cash flows net of hedging, $P - \Psi$, respectively, given x_1 and $\mathbf{F}_0$, under the model that generates the prices. Our hedging effectiveness metric (HEM) is the reduction in $\mathbb{VAR}(P \mid x_1, \mathbf{F}_0)$ caused by hedging:

$$\mathrm{HEM} := 100 \cdot \left[1 - \frac{\mathbb{VAR}(P - \Psi \mid x_1, \mathbf{F}_0)}{\mathbb{VAR}(P \mid x_1, \mathbf{F}_0)}\right].$$

For each pair of an assumed factor model and a hedging method that we consider, we use Monte Carlo simulation under a hypothetical true futures price model to estimate $\mathbb{VAR}(P \mid x_1, \mathbf{F}_0)$ and $\mathbb{VAR}(P - \Psi \mid x_1, \mathbf{F}_0)$, and hence HEM. HEM equals 100 with a perfect hedge. In this case, $P(\omega) - \Psi(\omega)$ is equal to $U^{\mathrm{RI}}(0, 0, \check{\mathbf{F}}(0))$, a constant given x_1 $(= 0)$ and $\mathbf{F}_0$, for almost every sample path ω.

Empirical backtesting of hedging using historical price data involves following a hedging strategy over a series of subintervals, each of which represents one observation in the sample. A complication is that the initial futures curve $\mathbf{F}_0$ and hence the initial valuation $U^{\mathrm{RI}}(0, 0, \check{\mathbf{F}}(0))$ vary across subintervals (i.e., observations). We thus modify HEM using the sample variances of $P - \Psi - U^{\mathrm{RI}}(0, 0, \check{\mathbf{F}}(0))$ and $P - U^{\mathrm{RI}}(0, 0, \check{\mathbf{F}}(0))$ to control for the different initial values across observations. We denote by BHEM the resulting historical backtesting version of HEM.

8.2.2. Hedging Results with Simulated Prices. Given paths of simulated prices from each of the three hypothetically true models in our model error framework (Figure 3), we implement and evaluate different hedging strategies based on a variety of assumed common factor futures price models and their associated deltas. The loading coefficients of factors in the assumed and hypothetically true models are taken to be the empirically estimated PCA coefficients (i.e., there is no parameter error). We then compare the resulting HEM estimates to assess the sensitivities of different hedging strategies to model error.

For brevity, we focus on the summer instance and consider a subset of the possible values of the number of assumed common factors, K. We take the NYMEX closing futures prices on June 1, 2012 as the initial time 0 futures curve. The financial trading positions are rebalanced once per month. We use a Monte Carlo sample size of 200 paths to estimate HEM for each assumed-model/hypothetically-true-model/hedge-strategy triple that we consider. Rebalancing requires nested simulations, under an assumed model, to compute deltas, for which we use 10,000 nested Monte Carlo samples. Obtaining the cash flows associated with a given hedging strategy along a futures curve sample path is computationally intensive—requiring about seven CPU minutes, depending on the price models and hedging method employed. We use the gcc version 4.8.2 20131017 (Red Hat 4.8.2-1) compiler on a 64-bit PowerEdge R515 with 12 AMD Opteron 4176 2.4 GHz processors, each with 64 GB of memory running Linux Fedora 19 (the stated CPU times correspond to using a single processor). We use the same random numbers when generating price samples across different models that share common factors.

Table 2 The Impact of Model Error on Hedging with Simulated Prices: HEM Estimates for NFH, FTFH, and BH

	$K^\diamond = 3$			$K^\diamond = 23$			$K^\diamond = 3$ plus noise		
K	NFH	FTFH	BH	NFH	FTFH	BH	NFH	FTFH	BH
1	87.79	87.40	98.14	87.73	87.73	98.31	87.09	87.09	98.61
2	49.96	96.56	98.80	−246.73	96.51	98.94	−89.71	96.61	99.19
3	97.55	98.87	98.87	−1,011.47	98.43	99.10	−1,172.10	98.61	99.32
5	73.25	98.82	98.80	−5,838.15	98.88	99.09	−32,304.47	98.89	99.33
10	98.00	98.80	98.77	−2,522.85	99.08	99.08	−31,325.74	99.20	99.32
15	98.30	98.79	98.76	−1,885.37	99.07	99.08	−25,540.89	99.25	99.32
20	98.69	98.78	98.76	−227.40	99.08	99.08	−4,224.62	99.32	99.32
23	98.76	98.78	98.76	99.08	99.08	99.08	99.32	99.32	99.32

Table 2 reports the HEM estimates for NFH, FTFH, and BH. The standard errors of these estimates are available in Online Appendix D. These hedging strategies are all near optimal (HEM close to 100) in the absence of model error ($K^\diamond = K = 3$ or 23; in the latter case, these strategies all coincide). These HEM estimates are not equal to 100 because of simulation, discrete rebalancing, and delta computation errors. The precisions (standard errors) of these HEM estimates are comparable. With spurious factors in the assumed model ($K^\diamond = 3 < K$), NFH, FTFH, and BH all perform well (estimated HEM $\geq$ 98.00), with one exception (estimated HEM = 73.25 with $K = 5$ for NFH). This good performance is expected, since overhedging against additional zero-probability futures curve changes has no impact on the effectiveness of hedges against assumed futures curve changes that actually can occur in the hypothetically true process. All the corresponding HEM estimates have similar precisions. The exception for NFH with $K = 5$ is due to the amplification by the large trading positions of NFH (discussed below) of delta estimation and discrete rebalancing errors, which also makes the resulting HEM estimate appreciably less precise than the other HEM estimates (for the spurious factor case).

The performance of different hedging strategies diverges sharply in the presence of missing factors or missing noise. With omitted important factors ($K^\diamond = 3 > K$), NFH with $K = 1$ assumed factor performs reasonably well (estimated HEM = 87.79), but its performance actually degrades with $K = 2$ assumed factors (estimated HEM = 49.96). With more omitted factors ($K^\diamond = 23 > K$), NFH with $K = 1$ assumed factor again exhibits reasonable performance (estimated HEM = 87.73) but with two or more assumed factors performs disastrously: the HEM estimates are large and negative. In other words, NFH dramatically *increases* residual cash flow variance relative to the unhedged cash flow variance. It is striking how large the effect on NFH is, since the omitted factors with $K \geq 3$ only contribute at most 1%–2% to the futures price log-return variance. The effect on NFH of omitting noise and either missing the second factor or having spurious factors ($K^\diamond = 3$ plus noise) is similarly bad, with the exceptions of the special cases $K = 1$ and 23.

The disastrous performance of NFH is a consequence of this strategy taking extremely large trading positions that magnify small model errors. These positions also make the resulting HEM estimates extremely imprecise compared with the case of smaller trading positions, which we interpret as a manifestation of the inadequacy of NFH rather than of the HEM estimator. Table 3 reports the average cumulative long and short futures positions taken by NFH along our simulated price paths for different hypothetical true models. The positions for values of K between 2 and 20 are very large. These very large positions are induced by the small determinants of the position generating matrices in the denominator of the second term on the right-hand side of expression (11), which occurs because the factor loadings of short-dated futures prices are rather similar. With no model error ($K^\diamond = K = 3$), these positions, even if very large, hedge the value of storage well because they are consistent with price changes from the hypothetically true model (although they might be impractical to trade when large). With underhedging caused by omitted factors or noise, however, these positions are structurally inconsistent with the realized price changes, and their large size magnifies the omitted randomness. This inconsistency generates severe discrepancies between the realized physical trading cash

Table 3 Average Cumulative Futures Positions Taken by NFH

	$K^\diamond = 3$		$K^\diamond = 23$		$K^\diamond = 3$ plus noise	
K	Long	Short	Long	Short	Long	Short
1	12.32	0.37	12.33	0.38	12.52	0.35
2	145.45	135.84	146.87	137.45	146.31	136.43
3	483.96	464.89	486.25	467.22	484.10	464.99
5	2,332.76	2,308.64	2,351.07	2,326.65	2,333.68	2,308.26
10	2,683.42	2,665.49	2,693.93	2,676.05	2,685.47	2,667.65
15	2,336.51	2,321.98	2,328.71	2,314.14	2,378.49	2,363.09
20	927.05	908.81	924.48	906.49	964.38	946.99
23	25.49	8.02	25.84	8.56	27.41	10.24

flows and the hedging cash flows. The good performance of NFH when K is equal to 1 is explained by the matrix inversion in (8), which is equivalent to (11), reducing to a division by a scalar (which does not generate huge trading positions); when K is equal to 23, it is the result of NFH reducing to BH, which is not based on (8).

The poor performance of NFH motivated us to develop a hedging strategy with small futures positions; FTFH heuristically achieves this goal (see §7.1). Although we expect FTFH to perform better than NFH with underhedging because of omitted factors or noise, how much was unclear a priori. Our results indicate that the improvement is dramatic. Specifically, the estimated HEMs of FTFH in Table 2 are 96.51 or higher and dominate the estimated HEMs for NFH in all cases irrespective of the type of model error whenever $K > 1$ (the two strategies behave similarly with $K = 1$). In fact, FTFH achieves near-optimal performance (i.e., HEM estimates close to 100 and with high precisions) when the number of assumed factors K is sufficiently large. The improvement in performance of FTFH relative to NFH is due to the smaller trading position sizes of FTFH: the average cumulative long (short) trading positions vary from 12.41 to 28.41 (0.35 to 9.97) across the considered cases. This finding underscores the impact of carefully selecting the particular traded contracts when implementing FH.

The HEM estimates for BH in Table 2 vary between 98.14 and 99.33 and are highly precise. The similar performance of BH across these different cases is remarkable. BH even outperforms FTFH. Consistent with the bounds in Proposition 3, the smallest and largest realized BH trading positions (i.e., deltas) across all sample paths, trading dates, and contracts traded are −0.30 and 0.59 irrespective of both the assumed number of factors used to compute the deltas and the hypothetically true model from which prices are sampled. The average cumulative long positions vary in between 24.90 and 27.48, and short positions vary in between 8.02 and 10.98. Hence, BH is an effective approach to reduce the impact of model error on hedging. This numerical finding is consistent with Proposition 4, which shows that BH is optimal in a minimum approximate-variance sense.

We investigate the impact of model error and the missing-contracts friction on hedging by analyzing the performance of CMVH implemented when trading is limited to 3, 5, and 10 front-end futures contracts ($L = 3$, 5, and 10). Table 4 shows that the performance of CMVH is good. The standard errors of the HEM estimates in this table are available in Online Appendix D. The improvement from using 5 versus 3 contracts is appreciable, and the improvement from using 10 versus 5 contracts is still nontrivial. CMVH is outperformed by BH, because, as shown in Proposition 4, BH is the optimal *unconstrained* approximate MVH approach. However, CMVH can still be near optimal in most cases when $L = 10$. The precisions of the HEM estimates in Table 4 increase with the number of traded contracts, L, given the number of assumed common factors, K, used to compute the deltas and the hypothetically true model. The CMVH trading positions are not large, despite being limited to just the front-end contracts. The ranges of the average cumulative trading positions for L equal to 3, 5, and 10 are 20.70–22.09, 25.59–27.22, and 30.17–32.00 for long positions and 8.71–9.74, 11.87–13.58, and 13.43–16.48 for short positions. Moreover, CMVH uses the same contracts as NFH when $K = L = 3$, 5, and 10. Thus, the good CMVH performance highlights the impact not just of *which* contracts are traded (as with FTFH) but also of the specific positions themselves (as compared with NFH). We further conjecture that the CMVH strategy would perform even better if some long-dated contracts had sufficient liquidity, so that they could be traded in addition to the short-dated contracts.

Table 4 **The Impact of Model Error and Missing Contracts on Hedging with Simulated Prices: HEM Estimates for CMVH Implemented Using the First $L = 3$, 5, and 10 Front-End Futures**

	$K^\diamond = 3$			$K^\diamond = 23$			$K^\diamond = 3$ plus noise		
	L			L			L		
K	3	5	10	3	5	10	3	5	10
1	93.22	95.26	98.11	90.80	93.74	97.33	91.00	93.73	96.74
2	93.74	95.92	98.63	91.65	94.50	97.91	91.91	94.55	97.56
3	93.89	96.05	98.67	91.92	94.64	97.94	92.29	94.78	97.80
5	93.94	96.09	98.64	92.07	94.73	97.93	92.48	94.87	97.86
10	93.94	96.08	98.62	92.09	94.72	97.90	92.55	94.89	97.87
15	93.94	96.08	98.62	92.08	94.71	97.89	92.57	94.90	97.88
20	93.94	96.08	98.62	92.09	94.71	97.88	92.57	94.89	97.89
23	93.94	96.08	98.62	92.09	94.71	97.89	92.57	94.89	97.88

8.2.3. Hedging Results with Historical Prices. We conclude with backtested hedging results for historical futures price data. Since the true futures price model is unknown, our backtesting analysis combines all of the possible model errors, including an unknown number of factors, parameter estimation errors, and misspecification resulting from possible stochastic volatility and jumps. Because history contains a single path, we split the data into rolling two-year subperiods to obtain multiple subpaths. We perform 15 in-sample experiments (for 1997–1998, 1998–1999, ..., 2011–2012) and 13 out-of-sample experiments (for 1999–2000 through 2011–2012, where the futures price model (2) is estimated over the preceding two years).

The hedging performance results using actual prices in Table 5 largely confirm our results with simulated prices (standard errors are available in Online

Table 5 **The Impact of Model Error on Hedging with Actual Prices: BHEM Estimates for NFH, FTFH, BH, and CMVH**

				CMVH		
K	NFH	FTFH	BH	$L=3$	$L=5$	$L=10$
			In-sample			
1	79.90	79.91	99.33	82.60	91.12	98.27
2	−2,300.25	99.00	99.36	84.69	92.36	98.44
3	−6,726.48	99.38	99.42	87.29	94.06	98.83
5	−1,302.06	99.12	99.39	88.26	94.63	98.91
10	−1,327.40	99.33	99.40	88.81	95.03	99.00
15	14.59	99.41	99.40	88.83	95.04	99.00
20	97.93	99.40	99.39	88.83	95.04	99.00
23	99.39	99.39	99.39	88.84	95.04	99.00
			Out-of-sample			
1	73.52	71.58	99.60	81.60	89.77	99.41
2	−2,195.80	95.11	98.86	85.40	92.84	99.74
3	−159,004.31	97.42	98.51	86.33	93.50	99.59
5	−7,643.56	98.10	98.28	86.62	93.64	99.42
10	−1,040.73	98.22	98.18	86.82	93.78	99.34
15	−500.70	98.15	98.15	86.82	93.78	99.32
20	94.39	98.15	98.16	86.84	93.79	99.32
23	98.15	98.15	98.15	86.82	93.78	99.32

Appendix D): NFH can perform disastrously (except when K is 1, 20, or 23, where the last case corresponds to BH), FTFH is effective once the most important statistical factors are included, and BH is still more effective. The BHEM estimates for FTFH and BH are more precise than the BHEM estimates for NFH (except for $K=1$ when comparing FTFH and NFH). With a few exceptions, the BHEM estimates are less precise out-of-sample than in-sample. When only the first three contracts are traded, CMVH performs reasonably well. Its performance improves substantially when the first 5 contracts are available, and it becomes near optimal using the first 10 contracts, in which case it is even slightly better than the performance of BH in the out-of-sample case, except when K equals 1. The precisions of the BHEM estimates improve considerably when the number of futures contracts available to trade increases, keeping the number of factors fixed. Compared with the in-sample case, in the out-of-sample case, the BHEM estimates are generally less precise when L is equal to 3 or 5 and more precise when L is equal to 10. (We attribute the slightly better out-of-sample than in-sample performance of CMVH with 10 traded contracts to the inferior precision of the reported in-sample BHEM estimates compared with the out-of-sample case.) Thus, CMVH appears to be an effective method in practice despite the presence of various model errors and the missing-contracts friction.

9. Conclusions

In this paper we show that various types of small empirically calibrated term-structure model errors can substantially impact the valuation and hedging of merchant commodity storage. We propose several near-optimal approaches to mitigate the negative effects of these errors on hedging, also considering liquidity constraints. Beyond our application to the merchant storage of natural gas, our research has relevance in other merchant operations contexts when valuing and hedging real and financial options that are contingent on futures price term-structure dynamics, or when deriving inventory/production management and capacity investment policies that depend on demand-forecast term structures and for which financial hedging can also be relevant. Further research might focus on (a) considering other hedging approaches, e.g., gamma and vega hedging; (b) investigating the effect of limited liquidity on the calibration of futures curve term-structure models; (c) studying the impact of model error when using operating policies other than the RI policy; (d) considering different assumed futures price term-structure models, e.g., common factor plus noise models or stochastic volatility models; and (e) extending our model error analysis to incorporate transaction costs (bid–ask spreads) in physical and financial trading and their impact on operating and hedging policies.

Supplemental Material

Supplemental material to this paper is available at http://dx.doi.org/10.1287/msom.2015.0518.

Acknowledgments

The authors thank the entire review team whose constructive feedback led to an improved version of this paper. F. Margot was supported by the Office of Naval Research [Grant N00014-09-1-0033] and National Science Foundation [Grant CMMI 1129163]. N. Secomandi was supported by the National Science Foundation [Grant CMMI 1129163].

References

Berling P, Rosling K (2005) The effects of financial risk on inventory policy. *Management Sci.* 51(12):1804–1815.

Birge JR (2000) Option methods for incorporating risk into linear capacity planning models. *Manufacturing Service Oper. Management* 2(1):19–31.

Bjerksund P, Stensland G, Vagstad F (2011) Gas storage valuation: Price modeling v. optimization methods. *Energy J.* 32(1):203–227.

Blanco C, Soronow D, Stefiszyn P (2002) Multi-factor models for forward curve analysis: An introduction to principal component analysis. *Commodities Now* (June):76–78.

Boogert A, de Jong C (2008) Gas storage valuation using a Monte Carlo method. *J. Derivatives* 15(3):81–98.

Boogert A, de Jong C (2011/2012) Gas storage valuation using a multifactor price process. *J. Energy Markets* 4(4):29–52.

Borovkova S, Geman H (2006) Seasonal and stochastic effects in commodity forward curves. *Rev. Derivatives Res.* 9(2):167–186.

Borovkova S, Geman H (2008) Forward curve modeling in commodity markets. Geman H, ed. *Risk Management in Commodity Markets: From Shipping to Agriculturals and Energy* (John Wiley & Sons, Chichester, UK), 9–32.

Boyabatlı O, Kleindorfer PR, Koontz SR (2011) Integrating long-term and short-term contracting in beef supply chains. *Management Sci.* 57(10):1771–1787.

Brace A, Gatarek D, Musiela M (1997) The market model of interest rate dynamics. *Math. Finance* 7(2):127–155.

Caldentey R, Haugh M (2006) Optimal control and hedging of operations in the presence of financial markets. *Math. Oper. Res.* 31(2):285–304.

Carmona R, Ludkovski M (2010) Valuation of energy storage: An optimal switching approach. *Quant. Finance* 10(4):359–374.

Casassus J, Collin-Dufresne P (2005) Stochastic convenience yields implied from interest rates and commodity futures. *J. Finance* 60(5):2283–2331.

Chen Z, Forsyth PA (2007) A semi-Lagrangian approach for natural gas storage valuation and optimal operation. *SIAM J. Sci. Comput.* 30(1):339–368.

Clewlow L, Strickland C (2000) *Energy Derivatives: Pricing and Risk Management* (Lacima Publications, London).

Cortazar G, Schwartz ES (1994) The valuation of commodity-contingent claims. *J. Derivatives* 1(4):27–39.

Detemple J (2006) *American-Style Derivatives: Valuation and Computation* (Taylor & Francis, London).

Devalkar SK, Anupindi R, Sinha A (2011) Integrated optimization of procurement, processing, and trade of commodities. *Oper. Res.* 59(6):1369–1381.

Ding Q, Dong L, Kouvelis P (2007) On the integration of production and financial hedging decisions in global markets. *Oper. Res.* 55(3):470–489.

Dixit AK, Pindyck RS (1994) *Investment Under Uncertainty* (Princeton University Press, Princeton, NJ).

Driessen J, Klaassen P, Melenberg B (2003) The performance of multi-factor term structure models for pricing and hedging caps and swaptions. *J. Financial Quant. Anal.* 38(3):635–678.

Eydeland A, Wolyniec K (2003) *Energy and Power Risk Management: New Developments in Modeling, Pricing, and Hedging* (John Wiley & Sons, Hoboken, NJ).

Fan R, Gupta A, Ritchken P (2001) On pricing and hedging in the swaption market: How many factors, really? Working paper, Case Western Reserve University, Cleveland.

Frestad D (2008) Common and unique factors influencing daily swap returns in the Nordic electricity market, 1997–2005. *Energy Econom.* 30(3):1081–1097.

Frestad D (2012) Liquidity and dirty hedging in the Nordic electricity market. *Energy Econom.* 34(5):1341–1355.

Gaur V, Seshadri S (2005) Hedging inventory risk through market instruments. *Manufacturing Service Oper. Management* 7(2):103–120.

Geman H (2005) *Commodities and Commodity Derivatives: Modeling and Pricing for Agriculturals, Metals and Energy* (John Wiley & Sons, Chichester, UK).

Geman H, Nguyen V-N (2005) Soybean inventory and forward curve dynamics. *Management Sci.* 51(7):1076–1091.

Glasserman P (2004) *Monte Carlo Methods in Financial Engineering* (Springer, New York).

Goel A, Gutierrez GJ (2011) Multiechelon procurement and distribution policies for traded commodities. *Management Sci.* 57(12):2228–2244.

Graves S, Meal H, Dasu S, Qiu Y (1986) Two-stage production planning in a dynamic environment. Axsäter S, Schneeweiss C, Silver E, eds. *Multi-Stage Production Planning and Inventory Control*, Lecture Notes in Economics and Mathematical Systems, Vol. 266 (Springer-Verlag, Berlin), 9–43.

Gray J, Khandelwal P (2004a) Realistic natural gas storage models II: Trading strategies. *Commodities Now* (September):1–5.

Gray J, Khandelwal P (2004b) Towards a realistic gas storage model. *Commodities Now* (June):1–4.

Heath D, Jackson PL (1994) Modeling the evolution of demand forecasts with application to safety stock analysis in production/distribution systems. *IIE Trans.* 26(3):17–30.

Hull JC (2010) *Risk Management and Financial Institutions*, 2nd ed. (Prentice-Hall, Boston).

Jaillet P, Ronn EI, Tompaidis S (2004) Valuation of commodity-based swing options. *Management Sci.* 50(7):909–921.

Kennedy DP (1994) The term structure of interest rates as a Gaussian random field. *Math. Finance* 4(3):247–258.

Kouvelis P, Li R, Ding Q (2013) Managing storable commodity risks: The role of inventory and financial hedge. *Manufacturing Service Oper. Management* 15(3):507–521.

Lai G, Margot F, Secomandi N (2010) An approximate dynamic programming approach to benchmark practice-based heuristics for natural gas storage valuation. *Oper. Res.* 58(3):564–582.

Lai G, Wang MX, Kekre S, Scheller-Wolf A, Secomandi N (2011) Valuation of storage at a liquefied natural gas terminal. *Oper. Res.* 59(3):602–616.

Leppard S (2005) *Energy Risk Management: A Non-Technical Introduction to Energy Derivatives* (Risk Books, London).

Longstaff FA, Santa-Clara P, Schwartz ES (2001) The relative valuation of caps and swaptions: Theory and empirical evidence. *J. Finance* 56(6):2067–2109.

Luenberger DG (2014) *Investment Science*, 2nd ed. (Oxford University Press, New York).

Manoliu M, Tompaidis S (2002) Energy futures prices: Term structure models with Kalman filter estimation. *Appl. Math. Finance* 9(1):21–43.

Maragos S (2002) Valuation of the operational flexibility of natural gas storage reservoirs. Ronn E, ed. *Real Options and Energy Management Using Options Methodology to Enhance Capital Budgeting Decisions* (Risk Publications, London), 431–456.

Pilipovic D (2007) *Energy Risk: Valuing and Managing Energy Derivatives*, 2nd ed. (McGraw-Hill, New York).

Plambeck EL, Taylor TA (2013) On the value of input efficiency, capacity efficiency, and the flexibility to rebalance them. *Manufacturing Service Oper. Management* 15(4):630–639.

Ross SA (1997) Hedging long run commitments: Exercises in incomplete market pricing. *Econom. Notes* 26(2):385–420.

Schwartz ES (1997) The stochastic behavior of commodity prices: Implications for valuation and hedging. *J. Finance* 52(3): 923–973.

Schwartz ES, Smith JE (2000) Short-term variations and long-term dynamics in commodity prices. *Management Sci.* 46(7):893–911.

Secomandi N (2010a) On the pricing of natural gas pipeline capacity. *Manufacturing Service Oper. Management* 12(3):393–408.

Secomandi N (2010b) Optimal commodity trading with a capacitated storage asset. *Management Sci.* 56(3):449–467.

Secomandi N, Kekre S (2014) Optimal energy procurement in spot and forward markets. *Manufacturing Service Oper. Management* 16(2):270–282.

Secomandi N, Seppi DJ (2014) Real option and merchant operations of energy and other commodities. *Foundations Trends Tech., Inform. Oper. Management* 6(3–4):161–331.

Secomandi N, Wang MX (2012) A computational approach to the real option management of network contracts for natural gas pipeline transport capacity. *Manufacturing Service Oper. Management* 14(3):441–454.

Seppi DJ (2002) Risk-neutral stochastic processes for commodity derivative pricing: An introduction and survey. Ronn E, ed. *Real Options and Energy Management Using Options Methodology to Enhance Capital Budgeting Decisions* (Risk Publications, London), 3–60.

Smith JE, McCardle KF (1999) Options in the real world: Lessons learned in evaluating oil and gas investments. *Oper. Res.* 47(1):1–15.

Suenaga H, Smith A, Williams J (2008) Volatility dynamics of NYMEX natural gas futures prices. *J. Futures Markets* 28(5): 438–463.

Thompson M (2012) Natural gas storage valuation, optimization, market and credit risk management. Working paper, Queen's School of Business, Queen's University, Kingston, ON, Canada.

Thompson M (2013) Optimal economic dispatch and risk management of thermal power plants in deregulated markets. *Oper. Res.* 61(4):791–809.

Thompson M, Davison M, Rasmussen H (2009) Natural gas storage valuation and optimization: A real options application. *Naval Res. Logist.* 56(3):226–238.

Tirole J (2006) *The Theory of Corporate Finance* (Princeton University Press, Princeton, NJ).

Tolmasky C, Hindanov D (2002) Principal components analysis for correlated curves and seasonal commodities: The case of the petroleum market. *J. Futures Markets* 22(11):1019–1035.

U.S. Energy Information Administration (EIA) (2010) Annual energy review 2009. Report DOE/EIA-0384(2009), U.S. Department of Energy, Washington, DC.

U.S. Energy Information Administration (EIA) (2011a) Annual energy outlook 2011 with projections to 2035. Report DOE/EIA-0383(2011), U.S. Department of Energy, Washington, DC.

U.S. Energy Information Administration (EIA) (2011b) Natural gas consumption by end use. Accessed March 3, 2015, http://www.eia.gov/dnav/ng/ng_cons_sum_dcu_nus_a.htm.

U.S. Energy Information Administration (EIA) (2011c) Underground natural gas working storage capacity. Accessed March 3, 2015, http://www.eia.gov/naturalgas/storagecapacity/.

Van Mieghem JA (2003) Capacity management, investment and hedging: Review and recent developments. *Manufacturing Service Oper. Management* 5(4):269–302.

Veronesi P (2010) *Fixed Income Securities: Valuation, Risk, and Risk Management* (John Wiley & Sons, Hoboken, NJ).

Wu OQ, Chen H (2010) Optimal control and equilibrium behavior of production-inventory systems. *Management Sci.* 56(8):1362–1379.

Wu OQ, Wang DD, Qin Z (2012) Seasonal energy storage operations with limited flexibility: The price-adjusted rolling intrinsic policy. *Manufacturing Service Oper. Management* 14(3): 455–471.