Manufacturing & Service Operations Management

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

Capacity Investment in Renewable Energy Technology with Supply Intermittency: Data Granularity Matters!

Shanshan Hu, Gilvan C. Souza, Mark E. Ferguson, Wenbin Wang

M&SOM MANUFACTURING & SERVICE OPERATIONS MANAGEMENT
Vol. 17, No. 4, Fall 2015, pp. 480–494
ISSN 1523-4614 (print) | ISSN 1526-5498 (online)

informs
http://dx.doi.org/10.1287/msom.2015.0536

Capacity Investment in Renewable Energy Technology with Supply Intermittency: Data Granularity Matters!

Shanshan Hu, Gilvan C. Souza
Department of Operations and Decision Technologies, Kelley School of Business, Indiana University, Bloomington, Indiana 47405
{hush@indiana.edu, gsouza@indiana.edu}

Mark E. Ferguson
Management Science Department, Moore School of Business, University of South Carolina, Columbia, South Carolina 29208,
mark.ferguson@moore.sc.edu

Wenbin Wang
Department of Operations Management, School of International Business Administration,
Shanghai University of Finance and Economics, 200433 Shanghai, China, wang.wenbin@shufe.edu.cn

We study an organization's one-time capacity investment in a renewable energy-producing technology with supply intermittency and net metering compensation. The renewable technology can be coupled with conventional technologies to form a capacity portfolio that is used to meet stochastic demand for energy. The technologies have different initial investments and operating costs, and the operating costs follow different stochastic processes. We show how to reduce this problem to a single-period decision problem and how to estimate the joint distribution of the stochastic factors using historical data. Importantly, we show that data granularity for renewable yield and electricity demand at a fine level, such as hourly, matters: Without energy storage, coarse data that does not reflect the intermittency of renewable generation may lead to an overinvestment in renewable capacity. We obtain solutions that are simple to compute, intuitive, and provide managers with a framework for evaluating the trade-offs of investing in renewable and conventional technologies. We illustrate our model using two case studies: one for investing in a solar rooftop system for a bank branch and another for investing in a solar thermal system for water heating in a hotel, along with a conventional natural gas heating system.

Keywords: supply intermittency; renewable energy; capacity investment; sustainability
History: Received: June 15, 2014; accepted: February 13, 2015. Published online in *Articles in Advance* June 5, 2015.

1. Introduction

Rising and volatile energy prices, in addition to carbon footprint considerations, are motivating many organizations to explore renewable energy-generating technologies, such as solar and wind power, to meet a portion of their energy needs. For example, Walmart has solar installations in 123 of its stores in California, out of 282 stores in total for the state (Nisperos 2014). In another example, Wells Fargo is investing in rooftop solar panels to generate electricity for some of its bank branches across the United States (Ovchinnikov and Hvaleva 2013). In this paper, we study the optimal capacity investment decision in a renewable technology at a single location, such as a rooftop solar system for generating electricity or heating water. The financial viability of a renewable energy investment depends on factors such as future energy prices, the availability of government incentives for renewable technologies, and the interaction between the firm's energy demand and the random renewable yield caused by supply intermittency. We discuss each of these below.

First, consider the volatility of energy prices. Figure 1 displays the monthly retail prices of electricity and natural gas for commercial users in the state of California between 2010 and 2014 (http://www.eia.gov). Electricity prices in this market exhibit a clear monthly seasonality as well as a steady upward trend, whereas the natural gas prices have no obvious pattern but display considerable volatility. Future monetary savings of replacing a conventional technology with a renewable one, such as a rooftop solar system to replace buying electricity from the grid, should take these price variabilities into account.

Second, recent incentives provided by federal and state governments, such as investment rebates and net energy metering (NEM) programs, increase the viability of renewable technologies. In a NEM program,

Figure 1 **(Color online) Monthly Commercial Electric and Gas Retail Prices in California**

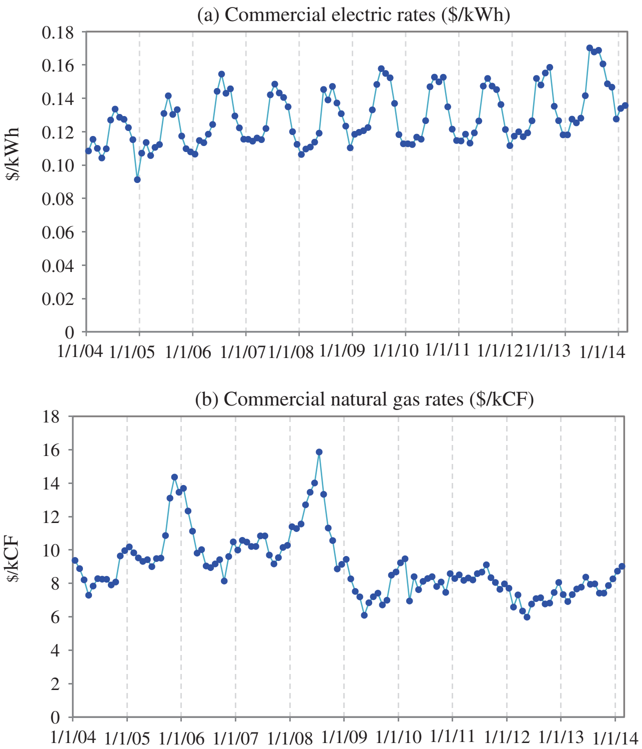
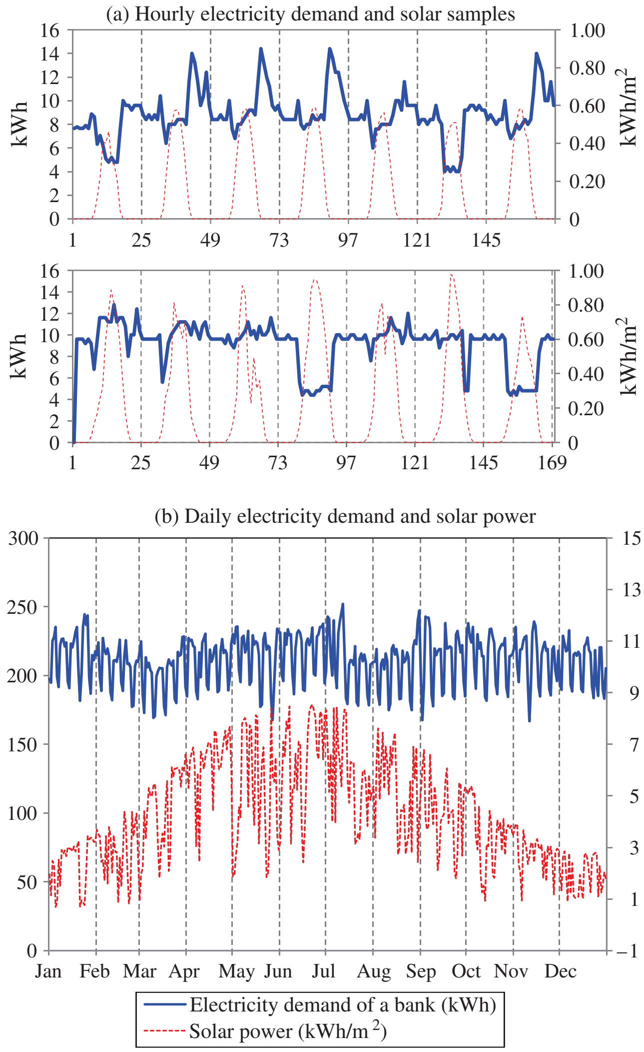


Figure 2 **(Color online) Electricity Demand at a Wells Fargo Branch in Los Angeles, California, and Solar Radiation**



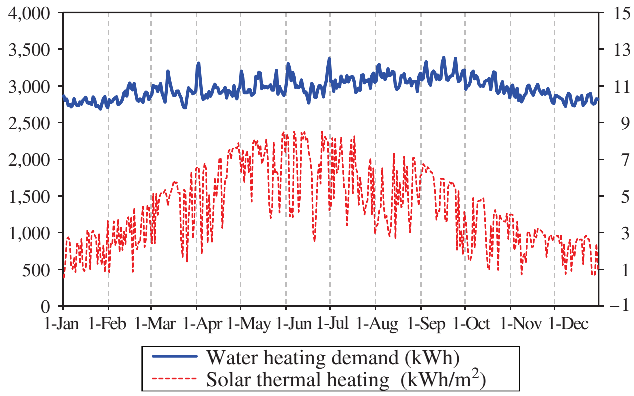
Notes. In panel (a), we plot the hourly electricity consumption and solar radiation data at a Wells Fargo branch for a typical winter week (top chart) (January 1, 2013, through January 7, 2013) and a typical summer week (bottom chart) (July 1, 2013, through July 7, 2013). In panel (b), we plot the daily electricity demand and solar power.

a facility that generates renewable energy on site can sell surplus electricity, in excess of demand, back to the grid. Because some states offer NEM whereas others do not, the presence or absence of NEM causes the optimal capacity decisions for renewable technologies to vary significantly between states. Moreover, with NEM, installers of renewable energy technologies can reduce their energy costs when their energy demand exceeds their supply as well as generate revenue when their supply exceeds their demand.

Third, renewable technologies present supply intermittency, which has consequences for serving energy demand. As an example, Figure 2 displays both hourly and daily solar radiation along with electricity consumption for a Wells Fargo branch in Los Angeles, California. The two charts in panel (a) display hourly solar radiation and electricity demand for two typical weeks in the winter (top) and summer (bottom) seasons. The chart in panel (b) displays the daily electricity demand and solar energy accumulated through each day for the entire year, where the data is displayed in daily (as opposed to hourly) buckets for ease of visualization. It is clear from these figures that there is both daily and monthly seasonality in both the solar radiation yield and the energy demand. It is also clear that solar radiation does not always peak at the same time periods as the demand for electricity does. Although it is technically possible to store solar energy in batteries for later use, the current energy storage technology is rarely cost effective for small installations such as the rooftop systems used in Wells Fargo bank branches and at Walmart's retail stores.

The example of Figure 2 would involve investment in a single technology—solar panels—with any unmet demand being delivered by the grid, which has an abundant supply. In other applications, the firm may require a portfolio of technologies, both renewable and conventional, to meet its energy demand. For example, consider the demand for hot water in a hotel in San Francisco, California, illustrated in Figure 3. The conventional technology being considered is natural gas-powered water heaters, but the hotel can also invest in a solar thermal system for heating water to replace a portion of its gas heater capacity.

Figure 3 (Color online) Water Heating Demand at a San Francisco Hotel and Solar Thermal Yield



Sources. http://en.openei.org/datasets/files/961/pub/ and http://www.nrel.gov/rredc (accessed June 14, 2014).

The solar thermal system can store solar energy captured during a sunny period as hot water for several hours. In this application, NEM rates do not matter because there is no electricity generation. The capacity investment decision in this case—how much capacity to invest for both solar thermal and natural gas systems—continues to depend on the daily interaction between solar yield and hot water demand. In addition, this decision also depends on the natural gas prices and the respective upfront investments per unit of installed capacity.

To factor in supply intermittency in renewable investment decisions, practitioners often use the concept of an average efficiency, that is, the average random yield through the year (Ovchinnikov and Hvaleva 2013). For example, the average efficiency of a photovoltaic (PV) solar system in Los Angeles is 18%. This means that for every kilowatt (kW) of installed capacity (in peak conditions), the system delivers 0.18 kW of power on average. Performing a capacity investment using average efficiency ignores the granular interplay of random yield and demand, as evidenced in Figure 2.

Our main contribution is to show that data granularity for renewable yield and energy demand at a fine level, such as hourly, matters for capacity investment decisions: Without energy storage, coarse data that does not reflect the intermittency of renewable generation may lead to a significant overinvestment in renewable capacity. Consequently, the average renewable efficiency approach currently used in practice may lead to a significantly lower benefit compared to the optimal investment that fully incorporates intermittency. We also indirectly demonstrate the value of energy storage technology by aggregating solar radiation and demand into longer time intervals (e.g., 12 hours). A secondary contribution is that we provide a decision support tool that addresses the challenges involved in determining the optimal one-time capacity investment levels for a portfolio of technologies that includes renewable energy options. Our tool factors in randomness and seasonality in energy demand and its interplay with the stochastic renewable yield. It also allows stochastic operating costs for conventional technologies (such as natural gas) and a stochastic penalty cost (e.g., buying electricity from the grid). In particular, our tool addresses a key challenge in this decision, namely different data granularities: The renewable yield may be provided in minute or hourly intervals, the electricity demand is obtained at 15-minute intervals, and the retail electricity prices are provided in monthly intervals. Considering the one-time nature of the investment, we show how to reduce the multiperiod cost structure into a single-period one that factors in time-dependent distributions and financial discounting. Finally, as a practical contribution we present two case studies that can be used in advanced classes in sustainable energy investments and/or business analytics.

2. Literature Review

Because we model random yield in the solar capacity, our work is related to the large literature on production and procurement with random yield. Yano and Lee (1995) provide a review of the early research regarding lot sizing with random yield, and more recent research is reviewed by Grosfeld-Nir and Gerchak (2004). Similar to our reduced single-period problem, many papers consider a newsvendor-type model, such as Karlin (1958), Noori and Keller (1986), and Henig and Gerchak (1990), among others. Compared with these papers, we contribute to the random yield literature by incorporating both cost and demand uncertainties and by allowing a portfolio of competing technologies to meet the demand.

Our recommendation of a supply portfolio (most optimal solutions consist of some renewable and some nonrenewable capacities) relates to the procurement problem with multiple supply sources. Significant efforts have been made to generalize the single-sourcing problem with random yield into a multi-sourcing scenario. However, as pointed out by Yano and Lee (1995, p. 329), even in the single-period problem with two suppliers, the problem is "extremely complex and hence it is difficult to obtain structural results." For instance, in a newsvendor setting with two unreliable suppliers, Parlar and Wang (1993) are only able to show the concavity of the expected profit function. In a generalized newsvendor model, Federgruen and Yang (2009) develop a highly efficient computational procedure to find the optimal set of suppliers and corresponding order quantities that minimize the total expected cost. Dada et al. (2007)

derive several structural results about the optimal sourcing policy in the presence of multiple suppliers that are unreliable because of random yield (similarly to us) or uncertain capacity (absent in our model). It is interesting to notice that their insights are different from ours even in their random-yield case: In selecting suppliers, the lowest unit-ordering cost supplier is selected first, whereas in our model a lower acquisition (unit investment) cost cannot guarantee the technology is selected in the optimal solution. This difference is attributed to the critical assumption about how random yield influences variable cost: Dada et al. (2007) assume that the buyer only pays for the good parts; thus, the payment depends on the yield (which applies to procurement contexts), whereas in our model the investment cost is independent of the actual yield realizations. Instead, the investment cost is solely determined by the installed capacity level (which applies to capacity investment contexts).

Capacity investment in multiple competing technologies has been studied in the literature on the joint decisions of capacity investment and peak load pricing for electricity providers (see Crew et al. 1995 for a literature survey of the area). In a seminal work, Crew and Kleindorfer (1976) study the problem of optimal pricing and capacity planning when the available plant types (technologies) have different investment and generating costs, and when demand is stochastic and price dependent. Their modeling framework is later extended by Chao (1983) and Kleindorfer and Fernando (1993), who model supply uncertainty caused by plant outage. Masters (2004) provides a graphical decision-making tool that allows one to find the optimal mix of technologies based on variable load demand, and fixed and variable generating costs for each technology; this is, in essence, a newsvendor problem with multiple products, which we discuss in the case of multiple technologies and a given renewable capacity in §5. Our paper differs from this stream of research in that we consider yield rate uncertainty in the renewable technology (such as solar and wind) by explicitly including its supply intermittency at a very granular level. Because the yield rate is primarily influenced by weather in the generation site, the yield rate is identical (but random) for all renewable generation units. In contrast, the supply uncertainty modeled in Chao (1983) and Kleindorfer and Fernando (1993) is meant to represent plant outages; thus, the authors assume that, within the same generating technology, supply uncertainty from an individual generating unit is independent of supply uncertainty from the aggregate capacity. This independence assumption, although technically convenient, does not apply to the case of renewable power intermittency.

Our paper also belongs to the growing literature on sustainable operations (e.g., Kleindorfer et al. 2005), particularly recent research on capacity investment in sustainable technologies. A stream of this research studies the impact of government policy on the adoption of sustainable technologies at a macro level (e.g., Krass et al. 2013, Chamama et al. 2014, Cohen et al. 2015, Kok et al. 2014, Ovchinnikov and Raz 2014). Krass et al. (2013) show that investment in sustainable technologies is nonmonotone in the magnitude of an environmental tax, but a fixed cost subsidy corrects this effect. Kok et al. (2014) show that a flat electricity pricing policy leads to a higher investment level in renewable technologies for utility firms when compared to a peak pricing policy. Ovchinnikov and Raz (2014) compare subsidies and rebates as policy mechanisms for the adoption of public interest goods, such as electric vehicles. Chamama et al. (2014) analyze how industry production of sustainable technologies changes over time under fixed and flexible subsidy policies. Cohen et al. (2015) study the impact of demand uncertainty on the optimal design of consumer subsidies for sustainable technologies. Another stream of research considers a given policy and provides a finer analysis of capacity investment decisions at the firm level, as we do (e.g., Wang et al. 2013, Drake et al. 2012). Motivated by Coca-Cola Enterprise's management of its delivery truck fleet, Wang et al. (2013) study a dynamic capacity investment problem with two competing technologies, where the diesel-electric hybrid vehicle requires higher upfront investment but lower operating cost than the conventional diesel truck. The fuel costs for both types of trucks are stochastic and have a perfect correlation in that both trucks use diesel as the fuel. Drake et al. (2012) study the optimal capacity investment for two technologies with different emission intensity, clean/expensive and dirty/cheap. Similar to our paper, Drake et al. (2012) also implement a newsvendor modeling framework and consider uncertain emission allowance pricing, which drives two stochastic cost processes with perfect correlation. In contrast to these two papers, we allow multiple technologies with different cost processes that may not be perfectly correlated. More importantly, we consider a random yield rate for the renewable technology to capture intermittency.

Renewable power intermittency has received significant attention by operation management researchers, but most papers (e.g., Kim and Powell 2011, Wu and Kapuscinski 2013, Zhou et al. 2014) consider decisions at the operational level, after the capacity is already installed. Two recent papers, Ambec and Crampes (2012) and Aflaki and Netessine (2012), investigate capacity investment in two technologies: renewable/intermittent and conventional/reliable. The first paper considers the problem under deterministic demand, whereas the second models stochastic demand. Because of this, the paper by Aflaki and Netessine (2012) is probably the closest to ours in that they consider both demand and supply uncertainties. In their case, however, supply uncertainty is assumed to have a two-point distribution, whereas we consider a general distribution that may be correlated with demand, and the spot price of purchasing power from the grid. Our bivariate distribution of yield rate and demand is built from observed yield and demand observations at a 15-minute or hourly interval, as shown in Figure 2. In §3.3, we show that the penalty from failing to model uncertainty at this more granular level (along with the subsequent correlations) can lead to significantly suboptimal investments in renewable energy capacity.

3. Single Renewable Technology

In this section we consider the simplest setting where the firm plans a one-time capacity investment k in a single renewable energy technology, such as wind or solar power, which is used to serve stochastic energy demand X_t in each period t of a planning horizon comprising T periods, corresponding to the capacity's lifespan. This setting applies to the case of Wells Fargo evaluating a PV system to generate electricity for one of its branches. A typical period length would be 15 minutes or one hour. Throughout this paper, we use the convention that lowercase (uppercase) symbols denote deterministic (stochastic) variables and parameters, except for the deterministic parameter T. There is a physical constraint on the capacity investment $\bar{k}$, for example, the size of a roof for a solar panel installation. Investment cost per unit of capacity is v, and so the total investment cost is vk, for $k \in [0, \bar{k}]$. Besides the initial investment cost, vk includes the expected net present value of total insurance and maintenance costs (including parts replacement and repurposing in case of failure) through the lifespan T, where the discount factor per period is δ. We later comment on how the model changes if there is a fixed installation cost, independent of k.

The effective capacity in any period is random to account for supply intermittency; denote the yield rate in period t by $\Lambda_t \in [0, 1]$, and so the effective capacity in period t is $\Lambda_t k$. There is a unit operating cost w per period, which could be negligible. In some cases, however, the government provides a production tax credit for generating energy from renewables, and so w is negative to reflect an operating income (credit) instead of cost. We also model NEM compensation: When solar or wind power are connected to the grid, power generated in excess of demand in a given period t is returned to the grid, and the consumer is credited at a rate of M_t per unit (say, kWh). We assume, reasonably, that the total profit generated from each capacity unit throughout its lifespan cannot be higher than the unit investment cost. Otherwise, the investment is decoupled from the demand, and the firm's optimal decision is to invest in an as large as possible amount of renewable capacity. Formally, $v > \sum_{t=1}^{T} \delta^{t-1} \mathsf{E}[\Lambda_t(M_t - w)]$. If there is unmet demand in any period, the firm can source energy from the spot market at a random cost P_t. (Alternatively, if demand is lost, P_t denotes the per-unit penalty cost.) We assume that $P_t >_{\text{a.s.}} M_t$, where a.s. stands for almost surely.

Denote $\mathbf{Y}_t = \{X_t, \Lambda_t, M_t, P_t\}$ as the vector of stochastic processes, which can be dependent. Total operating cost to meet demand at period t is

$$C(k; \mathbf{Y}_t) = w\Lambda_t k + P_t(X_t - \Lambda_t k)^+ - M_t(\Lambda_t k - X_t)^+. \quad (1)$$

The firm's cost-minimization problem is then

$$\min_{k \le \bar{k}} \mathscr{C}(k) = \sum_{t=1}^{T} \delta^{t-1} \mathsf{E}_{\mathbf{Y}_t}[C(k; \mathbf{Y}_t)] + vk. \quad (2)$$

If there is a fixed capacity installation cost m (in addition to the variable capacity cost vk), then the firm should compare $\mathscr{C}(k^*) + m$, where k^* is the optimal solution to (2), and $\mathscr{C}(0)$. This is a result of the convexity of the objective function. If $\mathscr{C}(0) < \mathscr{C}(k^*) + m$, then it is not optimal to invest in renewable capacity. Otherwise, the optimal renewable capacity investment is k^*.

3.1. Conversion to a Single-Period Problem

We now transform the multiperiod cost function into an equivalent single-period function by appropriately modifying the probability distributions. The main idea is to construct a new random vector $\mathscr{Y}$ by mixing the different random vectors $\{\mathbf{Y}_t\}_{t=1}^{T}$ with so-called "discounting probabilities" for different periods. Essentially, this is an application of Fubini's theorem (see, e.g., Ash and Doléans-Dade 1999). We can do this because (a) the instantaneous cost function $C(\cdot)$ that operates the *time-dependent* random vector $\mathbf{Y}_t$ is, by itself, *time-independent*; and (b) the objective function (2) is a linear summation of $\mathsf{E}_{\mathbf{Y}_t}[C(k; \mathbf{Y}_t)]$. This transformation would not be possible if the operating cost function (1) depends on the history of previous demands or costs. For example, if there is usage-based capacity deterioration, then the starting capacity in a period would depend on all demand realizations in previous periods. Our model allows, however, time-based capacity deterioration, as we illustrate in §3.3.

The next procedure encapsulates the time-dependence of $\{\mathbf{Y}_t\}_{t=1}^{T}$ and financial discounting into a time-independent *joint* distribution of $\mathscr{Y}$ as the input to the instantaneous cost function $C(\cdot)$.

First, we rewrite expression (2) as

$$\mathscr{C}(k)=\left(\sum_{m=1}^{T}\delta^{m-1}\right)\sum_{t=1}^{T}\frac{\delta^{t-1}}{\sum_{m=1}^{T}\delta^{m-1}}\mathsf{E}_{\mathbf{Y}_t}[C(k;\mathbf{Y}_t)]+vk. \quad (3)$$

Next, we define a discrete random variable Γ, which takes the value of $t\in\{1,2,\ldots,T\}$ with $r_t \doteq \delta^{t-1}/(\sum_{m=1}^{T}\delta^{m-1}) = ((1-\delta)\delta^{t-1})/(1-\delta^T)$. Then, we define a mixture of random vectors,

$$\mathscr{Y}=\sum_{t=1}^{T}\mathbf{1}_{\{\Gamma=t\}}\mathbf{Y}_t,$$

so that $\mathscr{Y}$ is a random sample of $\mathbf{Y}_1,\ldots,\mathbf{Y}_T$ and $\mathbf{Y}_t$ is selected with "probability" r_t. It follows that $\sum_{t=1}^{T} r_t \mathsf{E}_{\mathbf{Y}_t}[C(k;\mathbf{Y}_t)] = \sum_{t=1}^{T}\Pr\{\mathscr{Y}=\mathbf{Y}_t\}\mathsf{E}_{\mathbf{Y}_t}[C(k;\mathbf{Y}_t)] = \mathsf{E}_{\mathscr{Y}}[C(k;\mathscr{Y})]$. We may now rewrite the objective function (3) as

$$\mathscr{C}(k)=\frac{1-\delta^T}{1-\delta}\left(\mathsf{E}_{\mathscr{Y}}[C(k;\mathscr{Y})]+\frac{1-\delta}{1-\delta^T}vk\right).$$

Define $a=((1-\delta)/(1-\delta^T))v$ as the per-period allocation of the investment cost v (similar to the accounting practice of depreciating fixed assets with equal shares). We label a as the acquisition cost for the technology. Dropping the scaling factor $(1-\delta^T)/(1-\delta)$, we are looking for k that minimizes the following single-period objective function:

$$\mathscr{C}(k)=\mathsf{E}_{\mathscr{Y}}[C(k;\mathscr{Y})]+ak. \quad (4)$$

We show how to estimate the joint distribution of $\mathscr{Y}$ in practice using the Wells Fargo application in §3.3.

3.2. Structure of the Optimal Solution

We use the equivalent single-period formulation (4) and drop the time indices. Now, for the random variables X, Λ, M, and P, the marginal cumulative distribution functions are denoted by $F_X(\cdot)$, $F_\Lambda(\cdot)$, $F_M(\cdot)$, and $F_P(\cdot)$, with means μ_x, μ_λ, μ_m, and μ_p, respectively. These marginal distributions can be obtained from the joint distribution of $\mathscr{Y}=(X,\Lambda,M,P)$; we provide more details in the Wells Fargo application in §3.3.

Using (1), the objective function (4) can now be written as

$$\mathscr{C}(k) = (a+w\mu_\Lambda)k+\mathsf{E}[P(X-\Lambda k)^+] - \mathsf{E}[M(\Lambda k - X)^+]. \quad (5)$$

The objective function is convex in k. The optimal capacity decision of the renewable energy technology is as follows.

Proposition 1. *It is optimal to invest in the renewable technology if and only if* $\mathsf{E}[P\Lambda] > a+w\mu_\Lambda$. *If this condition is satisfied, then the optimal capacity* k^* *is given by the unique solution to*

$$\mathsf{E}[\Lambda\cdot M\cdot \mathbf{1}_{\{X<\Lambda k^*\}}]+\mathsf{E}[\Lambda\cdot P\cdot\mathbf{1}_{\{X\geq\Lambda k^*\}}]=a+w\mu_\Lambda \quad (6)$$

or $\bar{k}$, *whichever is smaller.*

The optimality condition (6) can be viewed as a generalized newsvendor solution with random unit retail price and random salvage value and may be reasoned by a marginal analysis: In the simplest newsvendor model, a retailer builds inventory Q to meet random demand X, with unit retail price π, unit acquisition cost c, and unit salvage value of s; the optimal Q^* satisfies

$$sF_X(Q^*)+\pi(1-F_X(Q^*))=c,$$

where the left-hand side is the expected marginal revenue of an extra unit of inventory and the right-hand side is the marginal acquisition cost of that unit. Equation (6) has the same interpretation after incorporating the yield uncertainty and its dependence on the spot market price and NEM rate. Note that, when solving Equation (6), a numerical search is necessary as k^* is embedded implicitly on the left-hand side of the equation.

3.3. Application 1: Solar Photovoltaic (PV) System in a Bank Branch

We illustrate the use of our model to optimize Wells Fargo's investment in a solar PV system given the demand and solar radiation data shown in Figure 2.

3.3.1. Data Set. The planning horizon is 30 years, representing the lifespan of a PV system. The operating cost w for a solar PV system is negligible. There are four stochastic processes as inputs to the model: demand $\{X_t\}$, solar yield $\{\Lambda_t\}$, electricity prices $\{P_t\}$ from the grid, and NEM compensation rates $\{M_t\}$. Combining multiple data sources, we establish the following data set. First, Wells Fargo provided us with one of their Los Angeles, California, bank branch's electricity meter readings for 2013 in 15-minute intervals; this is actual demand as the branch did not have a PV system in 2013. It is expected that electricity demand for this branch will be stationary during the planning horizon. Second, we obtained minute-by-minute solar radiation data for Los Angeles, available from April 2010 onward at http://www.nrel.gov/midc/lmu (accessed January 13, 2015). Third, we obtained the monthly retail electricity rates for commercial users in California as shown in Figure 1(a). Finally, the NEM compensation rate is regulated and relatively stable, so we assumed that its ratio to the grid electricity price is maintained at the current value of 0.33 for California (PG&E 2014). Hence, the NEM compensation rates increase annually, along with the annual average electricity prices.

The effective capacity of a solar panel degrades geometrically over its lifespan at an annual rate of 0.5%. Hence, in our reduced single-period problem, the solar yield rate is represented by the product of two independent random variables $\Lambda = L\cdot G$, where

L is the solar radiation rate (with no annual trend), and G has 30 discrete realizations [1,0.9950, 0.9900, 0.9851, ..., 0.8647], with their probabilities proportional to the corresponding discounting factors for 30 years.

The annual discount rate for Wells Fargo is 3.5%, which results in an annual discount factor of $\delta = 0.965$. In our calculations, we discount on a yearly basis so that all costs within the same year are discounted by the same factor. It is a straightforward extension to apply financial discounting on a shorter time interval.

3.3.2. Construction of $\mathcal{Y}$. We next construct the distribution of $\mathcal{Y}$. As discussed in §3.1, the idea is to collapse all *time-dependent* stochastic processes into *time-independent* probability measures, proportionally to the discount factors. Because we discount the costs annually, the cost for each hour within a year is treated equally for discounting purposes.

We first explain how to construct the electricity price process $\{P_t\}$ based on our 10-year historical monthly data. Three independent random variables P_A, P_B, and P_C are constructed to capture the trend, monthly seasonality, and random shocks of the electricity prices in the planning horizon. First, we estimate the annual trend by regressing the annual averages of historical prices over time (10 observations) and then use the regression equation to project the price trend for the next 30 years. The 30 projected values are used as the possible realizations of P_A, and their probability masses are $\mathbf{r} = \{1/(\sum_{t=0}^{29} \delta^t), \delta/(\sum_{t=0}^{29} \delta^t), \ldots, \delta^{29}/(\sum_{t=0}^{29} \delta^t)\}$, corresponding to r_t defined below Equation (3). To construct the monthly seasonality component P_B, we remove the trend from the historical observations so that the residuals only contain monthly seasonality and the random shocks. We estimate the monthly seasonality by averaging all residuals for the same month in the past 10 years. The 12 averages obtained are the realizations of P_B, and each has a probability mass of 1/12. If we need to discount the cost on a monthly basis, we use a vector of 12 probabilities that are proportional to the monthly discounting rates, $\{1/(\sum_{j=0}^{11} \delta_m^j), \delta_m/(\sum_{j=0}^{11} \delta_m^j), \ldots, \delta_m^{11}/(\sum_{j=0}^{11} \delta_m^j)\}$, where the monthly discounting rate is $\delta_m = \sqrt[12]{\delta} = 0.997$. However, it is easy to verify that all 12 values are very close to 1/12. Finally, for constructing P_C, we further remove the seasonality component from the first-round residuals, and the remaining values are observations for P_C. We then create a discrete probability distribution for P_C by generating a histogram for the P_C observations. The above procedure guarantees that P_C has a mean of zero.

If needed, the marginal distribution of $P = P_A + P_B + P_C$ can be easily obtained by sorting all realizations of P (the sum of three independent realizations) and the corresponding probabilities (the product of three probabilities). In fact, the decomposition helps to simplify the analysis: Because P_C is independent of all other random variables (X, Λ, M), it enters the optimality Equation (6) only through its mean, which is zero in our case. Hence, when numerically solving Equation (6), we can ignore P_C. It is also worth noting that the above decomposition can be replaced with a multivariate linear regression.

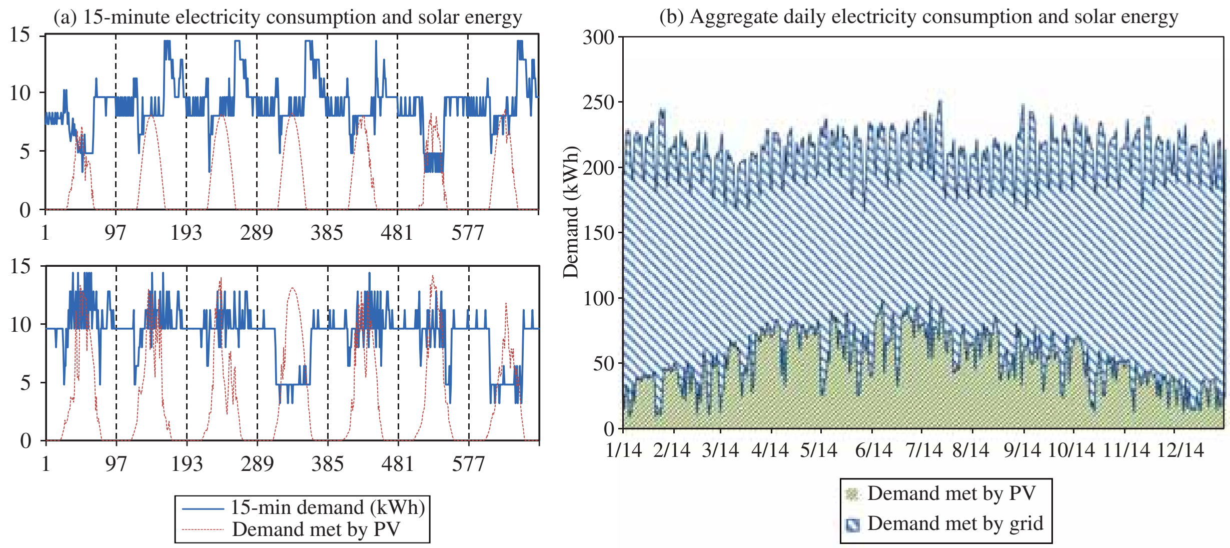
Next we construct the remaining random variables of $\mathcal{Y}$. Two other processes have an annual trend: the NEM compensation rate $\{M_t\}$, which increases with the price of electricity, and the solar yield $\{\Lambda_t\}$, which decreases because of PV degradation. For the NEM compensation rate, we set $M = 0.33P_A$ as explained before. The annual degradation for solar yield is captured by the random variable G. Note that the three trending elements P_A, M, and G are *fully dependent* and uniquely governed by the probability vector $\mathbf{r}$. In other words, they are three functions of a single random variable.

Constructing the demand distribution X and solar radiation L are more involved because they are dependent and have monthly seasonality. Thus, we need to estimate their joint distribution for each month. Define the random vectors $\{X_m, L_m\}$ as the random vector for month $m = 1, 2, \ldots, 12$; each has a probability of 1/12. We use January (when P_B takes the first value and $m = 1$) as an example to explain the procedure. First, we create 20 value bins for solar radiation and 20 value bins for demand, resulting in 400 buckets. (We have performed the computations using a higher number of bins than 20, with very similar results.) We then compute the proportion of observations in January that falls into each of the 400 buckets and use it as an approximate joint probability mass function (pmf). The proportion of the joint observations that fall into each of the 400 buckets approximates the pmf of $\{X_1, L_1\}$. Similarly, we estimate the distribution of $\{X_m, L_m\}$ for $m = 2, \ldots, 12$.

We now have all elements of $\mathcal{Y}$, and it is easy to generate the joint distribution of (P', X, Λ, M), where $P' = P_A + P_B$ with P_C dropped for computing Equation (6). Specifically, we have 30 annual indices for trend, 12 monthly indices for seasonality, and, for each month, 400 buckets of demand and solar radiation. This results in $30 \times 12 \times 400$ combinations, with each probability equal to the product of the three corresponding individual probabilities.

3.3.3. Optimal Solution and Impact of Data Granularity. In this section we show how granularity in demand and renewable yield data impacts the optimal renewable capacity and total cost. We first describe the cost parameters and compute the resulting optimal solution using the most granular data:

Figure 4 **(Color online) Optimal (15-Minute Granularity) Demand Fulfillment for Wells Fargo Branch in Los Angeles**



Notes. Panel (a) displays the 15-minute electricity consumption and energy generated by PV under the optimal capacity, for a typical winter week (top chart) (January 1, 2013, through January 7, 2013) and a typical summer week (bottom chart) (July 1, 2013, through July 7, 2013), corresponding to Figure 2. In panel (b), we plot aggregate daily demand and aggregate daily solar energy.

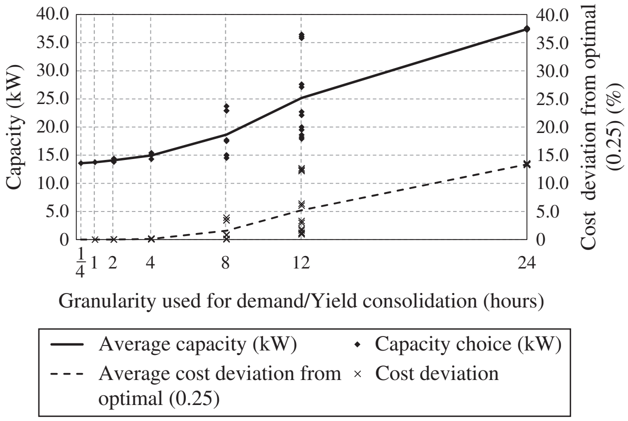
15-minute intervals for demand and solar yield. Then we compute the "optimal" solution for lower levels of granularity in demand and solar yield data by aggregating adjacent 15-minute intervals into one or more hours. Finally, we compare these solutions against the true optimal solution (which uses 15-minute intervals) in terms of both capacity levels and total cost.

The unit investment cost for a new PV system at this branch is $v = \$3{,}178/\text{kW}$. This cost is the sum of upfront investment cost, NPV of lifetime insurance cost, and fixed cost for inverter replacement and extensive maintenance, then deducting a tax incentive of 15% of system cost, tax savings from accelerated depreciation, and the installation cost from the Solar Incentive Program (SIP) in California. The breakdown costs above are from Ovchinnikov and Hvaleva (2013). Multiplying this value by the allocation adjustment ratio $(1-\delta)/(1-\delta^{30})$ results in $a = \$215/\text{kW}$. Solving Equation (6) using Excel Solver, the optimal PV capacity is $k^* = 13.6$ kW. Our entire solution approach is implemented in Excel and is available for downloading (see Example 1 in the supplemental material, available at http://dx.doi.org/10.1287/msom.2015.0536). Given that a typical rooftop panel of 1.56×0.79 m^2 yields 238 watts in peak conditions (Munigowda and de Vericourt 2012), or 5.2 m^2/kW, then a 13.6 kW system would require 70.3 m^2 (757 square feet), which is feasible for a typical bank branch; thus $k^* < \bar{k}$.

Figure 4 plots a hypothetical scenario (sample path) for the year 2014, where demand and solar radiation repeat the same pattern as in 2013. At the 15-minute level of our calculations, the solar PV system provides 100% of the branch's power needs for 22% of the hours. The two charts in panel (a) of Figure 4 show that during many summer hours (bottom), and in a few winter hours (top), the solar output is higher than demand. This fact is disguised in the chart shown in panel (b) of Figure 4, where the solar output never seems to be sufficient to meet the branch's power demand, but this is simply an artifact of daily aggregation for visualization.

We have also computed the optimal solution for different levels of granularity in demand and solar yield, that is, aggregating data in adjacent 15-minute intervals into one or more hours. Consider, for example, a granularity level of two hours. There are two possibilities for aggregating the data, depending on the starting time: (1) 12 A.M.–2 A.M., 2 A.M.–4 A.M., and so forth; and (2) 1 A.M.–3 A.M., 3 A.M.–5 A.M., and so forth. Similarly, a granularity level of h hours (where h is a factor of 24) indicates h aggregation possibilities. The results are shown in Figure 5. The optimal solution does not change significantly if one uses 15 minutes or one hour for demand-yield data granularity, at 13.6 kW and 13.7 kW respectively. As the granularity level lowers from 2 to 12 hours, the spread in the possible solutions increases. For example, the "optimal" solution is either 13.8 kW or 14.4 kW for 2-hour granularity, but ranges between 18.6 kW and 36.5 kW, averaging 25.2 kW, for 12-hour granularity. For 24-hour granularity, however, the solution only ranges between 37.3 kW and 37.6 kW, as one would expect. In general, for a given granularity level, better

Figure 5 Impact of Data Granularity for Demand and Solar Yield on Optimal Solution and Cost



Notes. For each granularity level (say, eight hours), we compute the "optimal" solar capacity for different starting times (12:00 A.M., 1:00 A.M., ..., 7:00 A.M.). Hence, there is more than one capacity choice (♦) for each granularity level (eight points for eight-hour granularity), except for granularity levels of 0.25 and one hour. The solid line plots the average capacity choice for each granularity level. Similarly, there are multiple percentage cost deviations (×) for each granularity level, and the dashed line plots their respective averages.

solutions (i.e., closer to the optimal of 13.6 kW, computed with 15-minute granularity) occur when there is a better match between solar supply and demand during the aggregated time intervals.

We also evaluate the total cost $\mathscr{C}(k)$, using the *correct* demand-yield distribution with 15-minute granularity, that results from a capacity level computed with less granular data as in Figure 5. We then compare this total cost with the optimal total cost and compute the cost deviation from optimal, which is shown in Figure 5. There is essentially no cost penalty for using a capacity solution computed with granularity levels of up to four hours. For 8-hour (12-hour) granularity, the cost penalty ranges from negligible to 3.9% (12.7%), averaging 1.6% (5.2%), depending on the starting time for aggregation. For daily granularity, the penalty cost ranges from 13.3% to 13.5%. In all of these cases, the suboptimal solution results in higher savings from the solar power but a larger increase in the investment cost. Note that if one solves the model with a given granularity level—say, 24 hours—the optimal capacity found using that level of granularity simulates the ability to store energy for that amount of time (say, 24 hours), as we discuss in more detail in Application 2. Here, there is no energy storage, as renewable energy produced by the PV panels in excess of demand is sold back to the grid at a rate equal to 1/3 of the penalty cost.

As another benchmark analysis, we compare our solution with a heuristic that uses the average yield efficiency to approximate the random yield—this would be equivalent to yearly granularity for random yield. To give the heuristic its best shot, we use our

Table 1 Effect of the NEM Rate M on the Optimal Capacity k^* ($\bar{k} = \infty$)

NEM (% of electricity price)	20	30	40	50	60	65	67	70
Optimal capacity (kW)	12.7	13.4	14.5	16.3	20.9	32.9	63.0	∞

decision tool so that all other system uncertainties and interdependencies are properly handled; the only modification is to set the yield rate to be the average yield rate. This would result in a capacity level of 43 kW, which is more than three times that of the optimal solution. In addition, such a system would require a roof size of 223 m^2 (2,400 square feet), which may render the project infeasible for a small branch. In terms of total cost, the heuristic would cost Wells Fargo 18% more than the total lifetime cost of the optimal solution. We emphasize that the suboptimality could be significantly higher if a practitioner approximated other system elements with constant values, such as price uncertainty, price-demand correlation, and so forth. We state our insights formally:

Observation 1. The optimal renewable capacity is (weakly) lower when computed with more granular data. Consequently, a renewable energy project with a given capacity that may not appear to be economically feasible when analyzed using nongranular data (such as daily or yearly average), because of the large upfront investment, may be attractive when analyzed using granular data (such as hourly).

Finally, we examine the effect of the NEM rate on the optimal capacity; the results are summarized in Table 1. When the NEM rate is below 40% of the retail electricity price, the optimal capacity is not very sensitive to the NEM rate. The sensitivity increases significantly when the NEM rate goes above 60% of the retail electricity price; in fact, the optimal solution approaches infinity for NEM rates higher than 70%. In practice, this means that the PV adopter should maximize the solar panel size up to the physical constraint $\bar{k}$.

3.3.4. A Branch at Another Location: Impact of Demand-Yield Correlation. The optimal solution depends to a great extent on the granular interactions between random demand and solar yield, considering that in this example the solar energy cannot be stored. To demonstrate the significance of this point, we obtained demand data from Wells Fargo for another branch located in Bluffton, South Carolina, and calculated the optimal solar PV capacity investment as above. All data and results are given in §C of the online supplement (available at http://dx.doi.org/10.1287/msom.2015.0536), but we provide the main insights here. There are important differences with respect to the branch in Los Angeles. First, solar yield and demand have a more negative correlation in the South Carolina branch compared to the California branch; this is a result of more significant after-hours processing in the South Carolina branch. Second, there is no NEM compensation available in South Carolina. Both of these factors decrease the attractiveness of solar PV as an on-site power generating option for the South Carolina branch. As a result, the optimal capacity size (relative to demand) is significantly smaller in the South Carolina branch. Here, the optimal PV capacity (using hourly granularity) is $k^* = 9.6$ kW.

Given the different demands profiles, we compare the two solutions for the California and South Carolina branches using the fraction of demand met by solar energy in the first year hypothetical sample path as an example (similar to Figure 4). We find that the optimal solar PV capacity is sufficient to meet 100% of the energy needs at the South Carolina branch for only 9.7% of the total energy-demanding hours, compared to 22.0% for the California branch. The large difference in our solutions between these two branches demonstrates that a standard capacity investment decision (e.g., as a percentage of maximum energy demand) can be far from optimal when solar yield, its granular interaction with demand, and the availability of NEM vary significantly across geographic locations.

In fact, we analytically show in §B of the online supplement that the optimal capacity level increases (decreases) in the correlation between the demand and the yield rate at low (high) capacity levels for the special case of a bivariate normal distribution of demand and random yield. This finding leads to our second key insight:

Observation 2. Renewable energy investments are more attractive, ceteris paribus, in areas where there is a higher correlation between demand and the random yield of the renewable technology. Consequently, a renewable energy project with a given capacity that may not be economically feasible in an area with low demand-random yield correlation may be attractive in an area with higher demand-random yield correlation.

In the next section we add two additional features to our model setup: the ability to invest in a conventional energy technology (not just acquiring power from the grid) and the ability to store energy. In the absence of high demand-random yield correlation, energy storage significantly increases the attractiveness of renewables, as we demonstrate in Application 2 next.

4. One Renewable Technology and One Conventional Technology with Storage

We begin by considering the situation where the firm combines a renewable technology with a conventional technology as a secondary source. The conventional technology has unit investment cost v_c, and random unit operating cost S, with mean μ_s, comprising fossil fuel consumption. One such example is the hotel in San Francisco illustrated in Figure 3, which could invest in a solar thermal system along with a conventional gas-powered system for water heating. Denote by c the unit acquisition cost of the conventional technology, that is, the per period allocation of the unit investment cost v_c: $c = ((1-\delta)/(1-\delta^T))v_c$. For simplicity, we assume that the stochastic costs terms satisfy $M <_{\text{a.s.}} S <_{\text{a.s.}} P$. (This assumption may be relaxed by allowing the ordering of M and S to change over time.) Moreover, we assume that the random vector (X, Λ) is independent of the random cost terms (S, M, P).

We present here the intuition and main results, leaving a formal model formulation to the general case of multiple renewable and conventional technologies to §5. Define the random variable $Z(k) := (X - \Lambda k)^+$ as the energy demand for the conventional technology, and denote its cumulative probability as $F_{Z(k)}(x) := \Pr\{(X - \Lambda k)^+ \leq x\}$. For a given capacity of the renewable technology k, the optimal capacity of the conventional technology $q^*(k)$ is given by a newsvendor solution, where the demand is $Z(k)$, the cost of underage is $(\mu_p - \mu_s) - c$, and the cost of overage is c. The optimal capacity decision for the renewable technology is as follows:

Proposition 2. *It is optimal to invest in the renewable technology if and only if $a/\mu_\Lambda + w < c + \mu_s$. If this condition is satisfied, the capacity of the renewable technology k^* is uniquely given by*

$$\mu_m \mathsf{E}[\Lambda F_X(\Lambda k^*)] + \mu_s \mathsf{E}[\Lambda \bar{F}_X(\Lambda k^*)] + (\mu_p - \mu_s)\mathsf{E}[\Lambda \bar{F}_X(\Lambda k^* + q^*(k^*))] = a + w\mu_\Lambda, \quad (7)$$

or $\bar{k}$, whichever is smaller. In (7), the optimal conventional technology capacity $q^*(k)$ is given by $q^*(k) = \bar{F}^{-1}_{Z(k)}(c/(\mu_p - \mu_s))$.

Equation (7) can be viewed as a generalization of (6) as follows. The expected marginal benefit of an extra unit of renewable capacity now comprises the following (a) the expected NEM revenue when effective renewable capacity exceeds demand; (b) the expected cost savings from using the renewable technology instead of the conventional one when demand is higher than the renewable capacity; and (c) the expected savings from using the conventional technology instead of buying from the spot market when demand exceeds total effective capacity. Note that the random variables P and M in (6) have been replaced with their means μ_p and μ_m because of the additional independence assumption between (X, Λ) and (S, M, P).

If there are fixed capacity installation costs m and m_c for the renewable and conventional technologies, respectively (in addition to the variable capacity costs vk and cq), then the firm should perform a scenario analysis to determine which technologies should be included in the portfolio. Specifically, the firm should determine the lowest cost option among $\mathscr{C}(k^*, q^*) + m + m_c$, $\mathscr{C}(0, q^*(0)) + m_c$, $\mathscr{C}(k^*(0), 0) + m$, and $\mathscr{C}(0, 0)$, where $q^*(0)$ and $k^*(0)$ denote the optimal conventional and renewable capacities, respectively, when used alone. The cost function $\mathscr{C}(k, q)$ is provided in §5.

4.1. Application 2: Water Heating System in a Hotel at San Francisco

For our second application, we consider a hotel in San Francisco, California, that invests in water heating capacity to satisfy its demand for hot water, as illustrated in Figure 3. We analyze here the demand for hot water and solar yield at a level of granularity of one or more hours, as the water heated by the solar thermal system that is not immediately used can be stored in the water tank and used later in the day. The level of granularity in the analysis matches the energy storage time, as we discuss below. As a result, we provide an analysis of the value of energy storage for renewable technologies by considering different granularity levels.

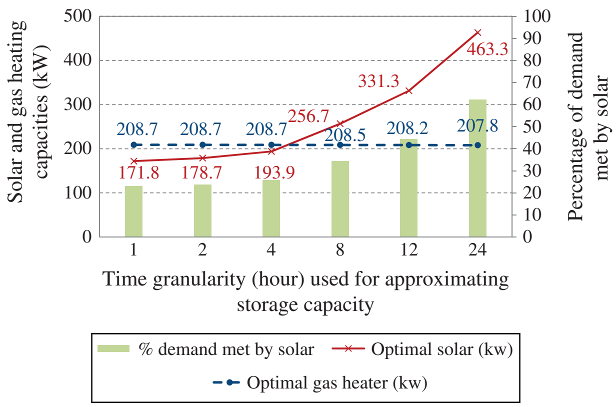
Our analysis is approximate because it assumes that all energy stored during a period (in excess of demand) is immediately lost at the start of the following period. To illustrate, suppose the granularity level (storage capability) is 24 hours. Then, solar energy produced during a 24-hour period can be used to meet any demand during that period but not in the subsequent period. In reality, solar energy stored as hot water in a water tank dissipates continuously. To assess the ability of our model to capture storage, we have also performed the calculations using an alternative simulation model where solar energy in excess of demand in an hour is available to meet demand in the following hour, and there is a given fraction of energy loss in every hour, say 4.2% ($= 1/24$, so that energy in storage is lost in one day). In such a model, energy storage capability is a parameter that limits the total amount of energy that can be stored at any time. The results for the simulation model are similar (details are available from the authors), and so we propose that a simpler way of modeling storage capability is to approximate it to a given granularity level, as we do here.

The hotel considers investments in both a solar thermal system and a conventional natural gas heater. At least some capacity of the natural gas heater is required because of the supply intermittency of solar power, even on a daily basis, and the fact that an outside option does not exist for water heating. The annual discount rate is 3.5%. The investment cost of a natural gas heater is \$193/kW. This cost is estimated based on an Energy Star-certificated commercial water heater that generates 130,000 BTU/h (38.1 kW/h) with 95% efficiency. The price of such an appliance, quoted by www.homedepot.com, is about \$5,380. In addition, the labor cost for installation is 20%, and the fixed maintenance cost is 10% of the acquisition cost. Accordingly, the unit cost is $(5{,}380/0.95) * 1.3/38.1 = 193/\text{kW}$. The lifespan of this heater is 10 years (Energy.gov 2014).

The investment cost of a solar thermal water system is \$614/kW. This cost is broken down as follows: The cost for the solar collector is about \$20 per square feet or \$215 per square meter (EIA 2011, p. 313). The total cost for the collector includes labor, the water storage tank, and the heat exchanger. With an efficiency of 70%, the unit price for a solar thermal system is thus $\$215 * 2/0.7 = \$614/\text{kW}$. The lifespan of this solar system is 25 years (EPIA 2011, p. 25). The operating cost w for a solar thermal system is negligible. As in Application 1, the solar panel has an annual degradation of 0.5%. The penalty cost for unmet demand is $p = \$95/\text{kWh}$, corresponding to a one-night revenue loss for one customer, which is estimated as follows. One person needs 10 gallons of hot water a day. Each gallon of water requires 540 BTU to be heated from 55 °F to 120 °F, so one person requires 1.5825 kWh of energy for water heating needs. At an average revenue of \$150/day per customer, the penalty cost is thus \$150/1.5825, or \$95/kWh.

Since NEM does not apply to the heating case because of the storage capability and the fact that the solar energy is used to heat water, not generate electricity, the uncertainties in this application can be represented as $\mathscr{Y} = (X, \Lambda, S)$, where S is the natural gas price. As shown in Figure 1(b), the historical natural gas prices did not have a noticeable trend or seasonality. Accordingly, we can directly estimate its distribution and also its long-run mean μ_s. Thus, we only need to estimate the joint distribution of (X, Λ). Similarly to Application 1, we separate them into two independent random vectors: (G) as the trend for solar yield because of degradation, and (X, L) as the joint distribution of demand and solar radiation, where $\Lambda = L \cdot G$. Note that, in this application, there is no need to model/estimate monthly seasonality because it does not appear to be present in the gas price process. Accordingly, when estimating the joint distribution of demand and solar radiation (X, L), we create $20 \times 20 = 400$ buckets and estimate the probability for each of them based on the historical joint observations from year 2004, which is the only data set we could find for the hotel's water heating demand. The construction of G is identical to Application 1.

Figure 6 (Color online) Impact of Storage Capability on Optimal Capacity



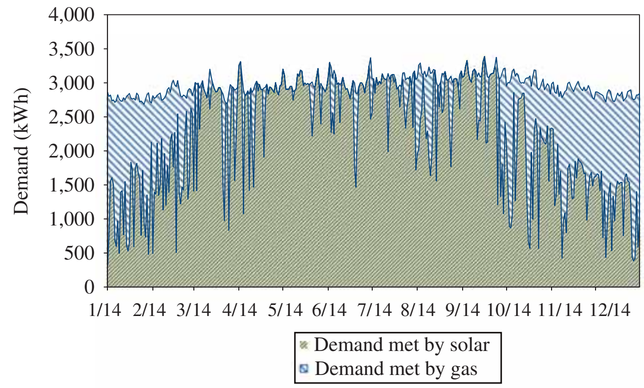
Notes. Similarly to Figure 5, we compute the optimal solar and gas capacities for each storage capability (granularity) level for different starting times. Here, however, we only plot the average capacity level for each storage capability. In addition, we also plot the (average) percentage of demand met by solar power. Differently from Figure 5, the storage capability levels are presented here as categories in the horizontal axis.

Plugging in the distribution of (X, Λ) and other relevant values, we find the optimal capacity levels from (7) using Excel Solver for different levels of storage capability (granularity). For 24-hour storage capability, the optimal capacity level averages 463.3 kW for the solar thermal system (ranging between 458.0 and 474.0 kW, depending on the starting time for aggregation) and 207.8 kW for the natural gas system. Here, 62.3% of water heating demand is met by the solar system. As in Application 1, our entire solution approach is implemented in Excel and is available for downloading (see Example 2 in the supplemental material, available at http://dx.doi.org/10.1287/msom.2015.0536). At the ratio of 5.2 m^2/kW, such a solar system requires 2,409 m^2 (25,932 square feet) of roof space, which is certainly feasible for this hotel.

Figure 6 provides a sensitivity analysis for the optimal capacities at different levels of storage capability (granularity) levels. Intuitively, the optimal solar capacity increases monotonically as the storage capability increases, because of a better match between random supply and demand. Just like in Application 1, there is a significant change in optimal solar capacity as the storage capability increases beyond four hours. The optimal solar capacity averages 193.9 kW for 4-hour storage (ranging between 187.9 and 202.1 kW, depending on the starting point for aggregation), but averages 331.3 kW for 12-hour storage (ranging between 260.8 and 448.3 kW). Similarly, the percentage of demand met by solar averages 25.6% for 4-hour storage and 44.4% for 12-hour storage.

Figure 7 shows the portions of demand fulfilled by the two different technologies in the hypothetical scenario (or sample path) where demand and solar radi-

Figure 7 (Color online) Optimal Water Heating Demand Fulfillment for a Hotel in San Francisco, California



ation in the first year after installation repeat the same pattern from the year 2004, with daily storage capability (granularity). Qualitatively, there is a sharp difference between Applications 1 and 2 in terms of solar power use. In Application 2, a significantly higher portion of the demand is met by the solar heating system: for 40.2% of the days in 2004, solar is the only energy source required to heat the water. There are, of course, many factors that differ between Applications 1 and 2, such as energy and investment costs, but the largest difference is that Application 2 has storage, with the possibility of daily granularity. We summarize this key insight:

Observation 3. Energy storage significantly enhances the attractiveness of renewable technologies. The value of energy storage can be estimated by solving the model at a lower granularity level (for the joint demand-yield distribution), such as daily, and comparing its solution to the solution obtained from solving the problem at a higher granularity level, such as hourly. The lower granularity level (daily) simulates, approximately, the ability to store renewable energy for that amount of time (one day).

5. Generalization to Multiple Technologies

In this section we generalize our model to the case of $N \geq 1$ renewable technologies and $L \geq 1$ conventional technologies. The capacity vectors for the renewable and conventional technologies are indicated by $\mathbf{k} = \{k_1, k_2, \ldots, k_N\}$, and $\mathbf{q} = \{q_1, q_2, \ldots, q_L\}$. We use superscripts $i \leq N$ and $j \leq L$ to denote renewable and conventional technologies i and j, respectively. If there is a supply surplus from the renewable technologies, it is compensated at the NEM rate M_t. We assume that NEM and the operating costs across the technologies are ordered: $M_t <_{\text{a.s.}} S_t^1 <_{\text{a.s.}} \cdots <_{\text{a.s.}} S_t^L <_{\text{a.s.}} P_t$.

At the beginning of any period t, the random vector $\mathbf{Y}_t = \{X_t, \mathbf{\Lambda}_t, \mathbf{S}_t, M_t, P_t\}$ is realized. The firm then dispatches the capacities in order of increasing unit operating costs, as described before, to minimize cost. The firm's cost minimization problem can then be written as

$$\min_{\mathbf{k},\,\mathbf{q}\geq 0} \mathscr{C}(\mathbf{k},\mathbf{q}) = \sum_{t=1}^{T} \delta^{t-1}\,\mathsf{E}_{\mathbf{Y}_t}[C(\mathbf{k};\mathbf{Y}_t)] + \sum_{i=1}^{N} v^i k_i + \sum_{j=1}^{L} v_c^j q_j, \tag{8}$$

where

$$C(\mathbf{k},\mathbf{q};\mathbf{Y}_t) = \sum_{i=1}^{N} w_t^i \Lambda_t^i k_i + \sum_{j=1}^{L} S_t^j \min\left\{q_j, \left(X_t - \sum_{m=1}^{j-1} q_m - \sum_{i=1}^{N} \Lambda_t^i k_i\right)^+\right\} + P_t\left(X_t - \sum_{i=1}^{N}\Lambda_t^i k_i - \sum_{j=1}^{L} q_j\right)^+ - M_t\left(\sum_{i=1}^{N}\Lambda_t^i k_i - X_t\right)^+. \tag{9}$$

The firm's expected discounted cost (8) comprises the sum of its expected operating cost across all periods and capacity acquisition cost across all technologies. The random operating cost in a period $C(\mathbf{k},\mathbf{q};\mathbf{Y}_t)$ comprises (a) the operating cost for all technologies sequentially deployed, (b) the cost of sourcing from the spot market for the demand that is unmet from the firm's capacity portfolio, and (c) the revenue of selling excess renewable energy to the grid for NEM compensation. Our assumption of ordered operating costs allows the problem to be solved efficiently because of the following result.

Lemma 1. *The one-period operating cost $C(\mathbf{k},\mathbf{q};\mathbf{Y})$ is jointly convex in $(\mathbf{k},\mathbf{q})$ for a given $\mathbf{Y}$, and the total discounted cost $\mathscr{C}(\mathbf{k},\mathbf{q})$ is jointly convex in $(\mathbf{k},\mathbf{q})$.*

The proof of Lemma 1 is based on a sample path argument and is provided in §A of the online supplement. Because the objective function (8) is jointly convex, and the set of constraints $\mathbf{k},\mathbf{q}\geq 0$ defines a convex set, the optimal solution $(\mathbf{k}^*,\mathbf{q}^*)$ can be found by the Karesh–Kuhn–Tucker condition. The joint convexity of (9) depends on a fixed dispatching order of using the technologies in any given period, from lowest to highest unit operating cost. If the operating cost depends on capacity utilization, or, in other words, there is a cost associated with unused capacity, such as the case with economies of scale, then the objective function is not convex, and a full numerical search is necessary to find the optimal capacity portfolio. In addition, none of the following analytic results apply.

In the special case where there is a single renewable technology, we can provide further analytic results by transforming the multiperiod cost function into an equivalent single-period function by appropriately modifying the probability distributions, as shown in §3.1. Then, the first-order conditions result in a "stacked newsvendor" solution for the conventional technologies, where $Z(k)$ is as defined in the previous section. This result is similar to the newsvendor model where demand can be met with multiple products, each with different revenue and cost characteristics. This solution is presented graphically by Masters (2004, p. 143) for a given demand distribution; we present it here in the context and notation of our model for completeness, where there is a given renewable capacity k that determines the demand for the conventional technologies $Z(k)$. For notational convenience, we denote the spot market as conventional technology $L+1$, with $c^{L+1}=0$, and $\mu_s^{L+1}=\mu_p$.

Lemma 2 (Masters 2004). *Given the renewable capacity k, the optimal capacities for the conventional technologies $\{q_{i_1}^*, q_{i_2}^*, \ldots, q_{i_n}^*\}$ where $q_{i_m}^*>0$ and $1\leq i_1<i_2<\cdots<i_n\leq L$ are determined as follows:*

(a) *Technology Choice: The first conventional technology is given by*

$$i_1 = \arg\min_{j\in\{1,\ldots,L\}} \{\mu_s^j \bar{F}_{Z(k)}(0)+c^j\}.$$

The remaining technologies are sequentially identified as

$$i_{m+1} = \arg\max_{i_m<j\leq L+1} \left\{\frac{c^{i_m}-c^j}{\mu_s^j-\mu_s^{i_m}}\right\} \quad \text{for } m=1,2,\ldots,n,$$

where we stop when reaching the dummy technology $i_{n+1}=L+1$, which is the spot market.

(b) *Optimal Capacity: The optimal capacities for the selected technologies are sequentially given by*

$$q_{i_m}^*(k) = \bar{F}_{Z(k)}^{-1}\left(\frac{c^{i_m}-c^{i_{m+1}}}{\mu_s^{i_{m+1}}-\mu_s^{i_m}}\right) - \sum_{r=1}^{m-1} q_{i_r}^*(k) \quad \text{for } m=1,2,\ldots,n. \tag{10}$$

Lemma 2 can be interpreted intuitively from a marginal analysis standpoint. Starting from zero capacity, one picks the technology that offers the lowest marginal cost, comprising the unit acquisition cost c^j plus the unit operating cost if there is remaining demand $\mu_s^j \bar{F}_{Z(k)}(0)$. Then, units of capacity are added, one by one, until the point $q_{i_1}^*$ where another technology i_2 provides the same marginal cost, that is, $\mu_s^{i_1}\bar{F}_{Z(k)}(q_{i_1}^*)+c^{i_1}=\mu_s^{i_2}\bar{F}_{Z(k)}(q_{i_1}^*)+c^{i_2}$, resulting in the optimal capacity of the first conventional technology: $\bar{F}_{Z(k)}(q_{i_1}^*)=(c^{i_1}-c^{i_2})/(\mu_s^{i_2}-\mu_s^{i_1})$. This means that the second technology selected, i_2, has the lowest ratio of the decrease in unit acquisition cost to the increase in unit operating cost, $(c^{i_1}-c^{i_2})/(\mu_s^{i_2}-\mu_s^{i_1})$ among the remaining technologies. This process continues until all demand is met. The optimal capacity k^* of the renewable technology is a generalization of (7):

Proposition 3. *It is optimal to invest in the renewable technology if and only if $a/\mu_\Lambda + w < c^j+\mu_s^j$ for all $j=1,\ldots,L$. Moreover, if the renewable technology is included* *in the optimal capacity portfolio, its capacity k^* is uniquely given by*

$$\mu_m \mathsf{E}[\Lambda F_X(\Lambda k^*)] + \mu_s^{i_1} \mathsf{E}[\Lambda \bar{F}_X(\Lambda k)] + \sum_{j=1}^{n} \left\{ (\mu_s^{i_{j+1}} - \mu_s^{i_j}) \mathsf{E}\left[\Lambda \bar{F}_X \left(\Lambda k^* + \sum_{r=1}^{j} q_{i_r}^*(k^*) \right) \right] \right\} = a + w\mu_\Lambda, \tag{11}$$

where the optimal conventional capacities $\{q_{i_r}^*(\cdot)\}_{r=1}^n$, as functions of k, are given per Lemma 2.

6. Concluding Remarks

In this paper we provide a decision support model for energy producing capacity investments in a portfolio of technologies, where one of the technologies is renewable and has a random yield, such as in wind and solar power. We consider a one-time decision, making our model appropriate for situations where the decision regarding the timing of when to add capacity has already been made and any capacity investments are static and irreversible. Given this setting, our model is quite general and incorporates, in addition to renewable supply intermittency, randomness in demand, operating costs, and the penalty cost for not meeting all of the energy demand. We allow for dependency between our stochastic processes, particularly between the energy demand and the random yield, which is a phenomenon observed in practice but not included in previous capacity investment models. To facilitate the application of our model, we provide detailed guidance about reducing time-varying, dependent stochastic processes into a multivariate distribution that can be used in a single-period analysis.

When the firm makes an investment only in one renewable and one conventional technology, we provide a newsvendor-type solution for the optimal capacity levels. We provide two distinct applications of our model with detailed guidance on how to implement our procedure for each one using only a spreadsheet. Our analytical and numerical results establish useful managerial insights for renewable capacity investment at a more general level.

We emphasize one important insight: The optimal capacity level for the renewable energy technology depends significantly on the interplay between the energy demand and the random yield as measured by very small time intervals such as hourly. Put differently, data granularity for renewable yield and electricity demand at a fine level, such as hourly, matters: Without energy storage, coarse data that does not reflect the intermittency of renewable generation will lead to a significant overinvestment in renewable capacity. This is because when the renewable energy cannot be stored, it is more desirable that the energy be generated during high demand periods, and the data must capture this interdependency. This inclusion of interdependencies is especially important for companies like Wells Fargo, who are considering renewable projects across different geographic locations. Instead of implementing a one-size-fits-all solution of installing the same amount of solar capacity at each bank branch, the firm should install larger capacities of solar PV capacity in areas with a positive yield-demand correlation.

From a policy perspective, firms may rely on governmental subsidies to justify investments in renewable energy, as they play a role in determining the optimal investment level through the unit investment cost. Currently, the same federal subsidy is provided to encourage renewable investment across different geographical areas. As illustrated by our model, this would lead to different investment levels in different areas, assuming that the private sector competes and converges to an equilibrium where marginal returns among different projects are the same. From the government's standpoint, this can be very inefficient in terms of a return on carbon emissions reduction per tax dollar spent. Everything else being equal (including total subsidy spending), a lower subsidy rate is needed for areas where there is a high yield-demand correlation, as these areas provide a better match between the supply of solar energy and the demand for energy consumption.

Our model does have limitations. We assume that there is no usage-based capacity deterioration, which implies that our operating cost function in a period is time- and usage-independent, allowing us to transform the problem into a single-period one. In addition, we assume that the marginal operating cost for using a technology is constant in any period, i.e., there are no costs of unused capacity or economies of scale. This assumption results in a convex objective function, allowing us to derive our analytic results and insights. Both of these assumptions apply well to the case of solar energy technology but are approximations in the case of wind technology or other conventional technologies where there are more mechanical components and moving parts.

In addition to addressing the limitations indicated above, future research could provide more insights into the dynamics of using multiple renewable technologies together. Although some of the results of §5 apply to multiple renewable technologies, one could also study specific practical applications. Another avenue is to focus on the design of government policy for renewable capacity investment at the firm level, considering the required data granularity. As discussed in §2, recent research has studied the impact of government policy, such as electricity pricing, subsidies, and rebates on the adoption of sustainable technologies at a macro level. Future research could address the design of policy mechanisms that explicitly address renewable intermittency, such as different NEM rates in different geographic areas.

Supplemental Material

Supplemental material to this paper is available at http://dx.doi.org/10.1287/msom.2015.0536.

Acknowledgments

The authors thank Stephen Graves, an anonymous associate editor, and three referees for their helpful feedback, which significantly improved this paper from earlier versions. The first and second authors acknowledge support from the Summer Research Grant awarded by the Kelley School of Business at Indiana University.

References

Aflaki S, Netessine S (2012) Strategic investment in renewable energy source. Working paper, INSEAD, Fontainebleau, France.

Ambec S, Crampes C (2012) Electricity provision with intermittent sources of energy. *Res. Energy Econom.* 34(3):319–336.

Ash RB, Doléans-Dade C (1999) *Probability and Measure Theory* (Academic Press, San Diego).

Chamama J, Cohen M, Lobel R, Perakis G (2014) Consumer subsidies with a strategic supplier: Commitment vs. flexibility. Working paper, Massachusetts Institute of Technology, Cambridge.

Chao H (1983) Peak load pricing and capacity planning with demand and supply uncertainty. *Bell J. Econom.* 14(1):179–190.

Cohen MC, Lobel R, Perakis G (2015) The impact of demand uncertainty on consumer subsidies for green technology adoption. *Management Sci.*, ePub ahead of print September 14, http://dx.doi.org/10.1287/mnsc.2015.2173.

Crew MA, Kleindorfer PR (1976) Peak load pricing with a diverse technology. *Bell J. Econom.* 7(1):207–231.

Crew MA, Fernando CS, Kleindorfer PR (1995) The theory of peak-load pricing: A survey. *J. Regulatory Econom.* 8(3):215–248.

Dada M, Petruzzi N, Schwarz L (2007) A newsvendor's procurement problem when suppliers are unreliable. *Manufacturing Service Oper. Management* 9(1):9–32.

Drake D, Kleindorfer PR, Van Wassenhove LN (2012) Technology choice and capacity portfolio under emissions regulation. Working paper, Harvard Business School, Boston.

EIA (U.S. Energy Information Administration) (2011) Annual energy review 2011. Accessed September 23, 2014, http://www.eia.gov/totalenergy/data/annual/pdf/aer.pdf.

Energy.gov (2014) Covered product category: Commercial gas water heaters. U.S. Office of Energy Efficiency and Renewable Energy. Accessed September 23, 2014, http://energy.gov/eere/femp/covered-product-category-commercial-gas-water-heaters.

EPIA (European Photovoltaic Industry Association) (2011) Solar photovoltaics competing in the energy sector. Report, EPIA, Brussels. Accessed September 23, 2014, http://www.epia.org/fileadmin/user_upload/Publications/Competing_Full_Report.pdf.

Federgruen A, Yang N (2009) Optimal supply diversification under general supply risks. *Oper. Res.* 57(6):1451–1468.

Grosfeld-Nir A, Gerchak Y (2004) Multiple lotsizing in production to order with random yields: Review of recent advances. *Ann. Oper. Res.* 126(1–4):43–69.

Henig M, Gerchak Y (1990) The structure of periodic review policies in the presence of random yield. *Oper. Res.* 38(4):634–643.

Karlin S (1958) One stage models with uncertainty. Arrow, KJ, ed. *Studies in the Mathematical Theory of Inventory and Production*, Chap. 8 (Stanford University Press, Stanford, CA).

Kim JH, Powell WB (2011) Optimal energy commitments with storage and intermittent supply. *Oper. Res.* 59(6):1347–1360.

Kleindorfer PR, Fernando CS (1993) Peak-load pricing and reliability under uncertainty. *J. Regulatory Econom.* 5(1):5–23.

Kleindorfer PR, Singhal K, Van Wassenhove LN (2005) Sustainable operations management. *Production Oper. Management* 14(4): 482–492.

Kok AG, Shang K, Yucel S (2014) Impact of electricity pricing policy on renewable energy investments and carbon emissions. Working paper, Duke University, Durham, NC.

Krass D, Nedorezov T, Ovchinnikov A (2013) Environmental taxes and the choice of green technology. *Production Oper. Management* 22(5):1035–1055.

Masters G (2004) *Renewable and Efficient Electric Power Systems* (John Wiley & Sons, Hoboken, NJ).

Munigowda M, de Vericourt F (2012) Solairedirect: The quest for solar energy. Case 612-027-1, European Case Clearing House, INSEAD, Fontainebleau, France.

Nisperos N (2014). Walmart to double number of solar projects. *The Sun* (May 27), http://www.sbsun.com/environment-and-nature/20140527/walmart-to-double-number-of-solar-projects.

Noori AH, Keller G (1986) One-period order quantity strategy with uncertain match between the amount received and quantity requisitioned. *INFOR* 24(1):1–11.

Ovchinnikov A, Hvaleva A (2013) Wells Fargo: Solar energy for Los Angeles branches. Darden Business Publishing Case UVA-QA-0800, University of Virginia, Charlottesville.

Ovchinnikov A, Raz G (2014) Coordinating pricing and supply of public interest goods using government rebates and subsidies. Working paper, Queens University, Kingston, ON, Canada.

Parlar M, Wang D (1993) Diversification under yield randomness in two simple inventory models. *Eur. J. Oper. Res.* 66(1):62–54.

PG&E. (2014) Getting credit for surplus energy. Accessed September 23, 2014, http://www.pge.com/en/mybusiness/save/solar/surplus.page.

Wang W, Ferguson ME, Hu S, Souza GC (2013) Dynamic capacity investment with two competing technologies. *Manufacturing Service Oper. Management* 15(4):616–629.

Wu O, Kapuscinski R (2013) Curtailing intermittent generation in electrical systems. *Manufacturing Service Oper. Management* 15(4):578–595.

Yano CA, Lee HL (1995) Lot sizing with random yields: A review. *Oper. Res.* 43(2):311–334.

Zhou Y, Scheller-Wolf AA, Secomandi N, Smith S (2014) Managing wind-based electricity generation in the presence of storage and transmission capacity. Working paper, Carnegie Mellon University, Pittsburgh.