Investments in Renewable and Conventional Energy: The Role of Operational Flexibility

A. Gürhan Kök • Kevin Shang • Şafak Yücel

College of Administrative Sciences and Economics, Koç University, Sariyer, 34450, İstanbul, Turkey
Fuqua School of Business, Duke University, Durham, North Carolina 27708, United States of America
McDonough School of Business, Georgetown University, Washington, D.C. 20057, United States of America
gkok@ku.edu.tr • kevin.shang@duke.edu • safak.yucel@georgetown.edu

There is an ongoing debate among energy experts on how providing a subsidy for one energy source affects the investment in other sources. To explore this issue, we study capacity investments of a utility firm in renewable and conventional sources. Specifically, conventional sources are categorized into two groups as inflexible (e.g., nuclear and coal-fired plants) and flexible (e.g., natural gas), based on operational flexibility, i.e., whether or not the output of a source can be ramped up or down quickly. We model this problem as a two-stage stochastic program in which the firm first determines the capacity investment levels, followed by the dispatch quantities of energy sources so as to minimize the sum of the investment and generation-related costs. We derive the optimal capacity portfolio and characterize the interactions between renewable and conventional sources. We find that operational flexibility plays a key role in these interactions: renewable and inflexible sources are substitutes, whereas renewable and flexible sources are complements. This result suggests that a subsidy for the nuclear or coal-fired power plants leads to a lower investment level in renewables, whereas a subsidy for the natural gas-fired power plants leads to a higher investment in renewables. We validate this insight by using real electricity generation and demand data from the state of Texas. Finally, we show that a carbon tax leads to a lower renewable investment if the inflexible source is carbon-free nuclear energy.

Version: October 17, 2016

Keywords: Energy-related Operations, Capacity Investment, Volume Flexibility.
1 Introduction

Policymakers have introduced various subsidy policies to encourage investments in clean energy sources in order to reduce carbon emissions. For instance, the U.S. government provides a 30% subsidy for the investment costs in solar energy (SEIA 2016) and the state of New York is planning to offer a multibillion-dollar subsidy for the nuclear power plants (Yee 2016). However, there is an ongoing debate on how an increased investment in one type of energy source (due to a subsidy) affects the investment in others. On one hand, in a recent article in Forbes, Kelly-Detwiler (2014) claims that carbon-free nuclear and renewable energy sources are “best friends” because nuclear can provide a steady back-up electricity supply for intermittent renewables. On the other hand, the former chairman of the Federal Energy Regulatory Commission argues that the nuclear source is inflexible (i.e., nuclear power cannot be ramped up or down quickly) to deal with intermittency (Straub and Behr 2009). Contradictory claims are also reported on the interaction between renewable and natural gas-fired power plants. In The New York Times, Kotchen (2012) claims that low natural gas prices are a “trap” for renewables because, in response to lower natural gas cost, a utility firm would invest more in natural gas-fired plants rather than renewables. On the contrary, Keith (2013) calls this claim a “myth” related to the renewables in The Wall Street Journal and argues that natural gas can complement renewables by alleviating the intermittency problem. In this paper, we investigate these interactions between energy sources, with a focus on the capacity investment decisions of utility firms, which constitute the majority of the total energy investment in the U.S.

In recent years, utility firms have significantly invested in renewable sources, such as solar and wind energy, because they provide electricity with negligible generation costs. To cope with the intermittency of the renewable energy sources, utility firms need to invest in conventional sources as well. Conventional sources are categorized into two groups as inflexible and flexible, based on operational flexibility, i.e., whether or not the output of a source can be ramped up or down quickly. A nuclear or coal-fired power plant, for instance, is inflexible because its output cannot be changed rapidly due to technical reasons. On the other hand, a natural gas-fired power plant is flexible (DOE 2011). From the perspective of cost structures, an inflexible source has higher investment but lower generation costs than a flexible source. Considering these characteristics, it is
challenging for a utility firm to determine the right capacity investment portfolio which minimizes its investment and generation costs while maintaining a certain reliability level (i.e., the chance of no blackouts). For example, Smith (2013) has recently identified a “looming energy crisis” for the utility firms in California because they do not have “the right mix of power plants” and are vulnerable to reliability problems due to over-reliance on intermittent renewables. Motivated by these policy discussions, we pose the following research questions. What is the optimal capacity portfolio for a utility firm that aims to minimize its investment and generation costs in the presence of inflexible, renewable, and flexible sources? What is the role of operational flexibility in the interaction between the conventional and renewable sources? How do other policies, such as a carbon tax, affect investments and the probability of blackouts (i.e., reliability) in an electricity system?

We model this problem following the decision process of a utility firm for making capacity investments. More specifically, a utility firm first takes a long-term, strategic capacity decision by investing in different energy sources. The invested capacity level of a source is the maximum output that the utility firm can dispatch from that source during each of the operating periods, which is often set to be five minutes. The decision of dispatching energy supply to match the demand is based on the five-minute-ahead forecasts of the electricity demand and the intermittency of renewable sources. After the uncertainties of demand and supply are realized, a penalty cost is incurred if the electricity demand cannot be fully satisfied. This penalty cost represents consumer’s inconvenience costs and the costs that the utility firm has to purchase excess energy from external sources. Following this practice, we formulate this problem as a two-stage stochastic program with recourse. In the first stage, subject to demand and supply uncertainties, the firm makes a strategic decision by determining the capacity investment in the inflexible, renewable, and flexible sources. In the second stage, the firm determines the amount of electricity dispatched from these energy sources for each operating period based on the forecasts. The objective of the utility firm is to minimize the total expected cost, which is the sum of the initial investment costs, the electricity generation costs, and the penalty costs of supply shortage within finite operating periods.

We solve the utility firm’s investment problem by using backward induction and characterize the optimal dispatch policy: all inflexible capacity is first used, followed by the renewable energy capacity as its generation cost is negligible compared to the flexible source, which is used as the
last resort. Based on this optimal dispatch policy, we determine the optimal investment level for each source. We obtain a multi-dimensional newsvendor-type solution. That is, the utility firm balances the underage cost (e.g., the penalty cost due to supply shortage) with the overage cost (i.e., the investment cost) for each energy source in the demand and intermittency space. In the most practical case where the investment levels of all sources are positive, the critical fractile associated with the flexible source determines the probability of meeting the demand. This indicates that the reliability of the electricity system is determined by the cost parameters of the flexible source and the penalty cost rate. This finding reveals an important policy insight that the reliability is only affected by a subsidy provided for the flexible source but not the subsidies for renewable and inflexible sources.

To identify how a subsidy for one source affects the investment level in others, we examine the interaction between energy sources. In particular, we define two sources as substitutes (complements, respectively) if a decrease in the investment cost of one source leads to a decrease (an increase, respectively) in the investment level of the other. We find that these interactions are determined by the operational flexibility. Specifically, inflexible and renewable sources are substitutes. Intuitively, the inflexible and renewable sources share similar characteristics in the capacity portfolio: both are costly to invest, inexpensive to operate, but uncontrollable (inflexible or intermittent). This result implies that lowering the investment cost of nuclear or coal-fired plants leads to a lower investment in wind or solar energy. On the other hand, under certain conditions, we show that the renewable and flexible sources are complements as they have opposite characteristics: the flexible source is inexpensive to invest but costly to operate, and is controllable. Thus, in response to a price reduction of natural gas, a utility firm increases its investment level in renewables. Lastly, the inflexible and flexible sources are substitutes. This is because the flexible source is complementary to the renewable source which is a substitute for the inflexible source. We validate these findings by using real electricity generation and demand data from the state of Texas in Section 7.

We also consider the impact of a carbon tax on energy investments. Many experts claim that taxing carbon emissions motivates the investment in renewable sources (c.f., EIA 2013a, Tyson 2013, and Walls 2015). Our analysis shows that this claim does not hold when the inflexible source is carbon-free nuclear energy. In this case, the carbon tax only increases the generation cost of the flexible energy source. This results in a reduction of the investment of flexible source, which, in
turn, reduces the investment of the renewable source due to the complementarity effect.

Finally, we investigate how investments are affected by the penalty cost which is incurred by the utility firm if the demand exceeds the supply. The penalty cost can be reduced by introducing a real-time spot market for a network of utility firms (as in the case of Southwest Power Pool). Under certain conditions, we show that a reduction in the penalty cost rate leads to a higher investment in the renewable source and a lower investment in the two conventional sources. This is because a lower penalty rate enables the utility firm to invest more in the (risky) intermittent source with negligible generation cost rather than conventional. Hence, by introducing real-time spot markets, renewable investment can be motivated.

The rest of the paper is organized as follows. Section 2 reviews the related literature. Section 3 introduces our model. Section 4 derives the optimal capacity investment portfolio. Section 5 analyzes the relationships between energy sources under the optimal capacity portfolio. Section 6 studies the effects of carbon tax, generation subsidies, and penalty cost. Section 7 validates our main results by using real data through a case study. Section 8 considers several extensions of our model. Section 9 concludes.

2 Literature Review

2.1 Dual Sourcing

Dual sourcing literature dates back to Barankin (1961) who studies an inventory model with two suppliers: the first supplier features a long lead time and a low procurement cost, whereas the other has a short lead time and a high procurement cost. Our work also relates to this cost and responsiveness trade-off. In a different context, Allon and Van Mieghem (2010) analyze a tailored base-surge policy that replenishes inventory from an offshore supplier to satisfy constant demand and from an onshore supplier to satisfy demand shocks. In addition to the cost and responsiveness trade-off, supplier reliability is studied extensively in the literature. See Yano and Lee (1995) and Minner (2003) for reviews. In this domain, Dada et al. (2007) investigate the procurement decision of a newsvendor who orders from multiple unreliable and capacitated suppliers. They show that the newsvendor selects suppliers based on their cost and determines the order size based on the reliability of selected suppliers. Federgruen and Yang (2008) study a similar setting with fixed costs
of retaining suppliers and propose heuristics to determine order sizes. Tomlin (2009) uses Bayesian learning for the reliability parameter of suppliers and shows that an increase in the reliability forecast increases the attractiveness of a supplier. Wang et al. (2010) compare a dual sourcing strategy and a process improvement strategy in order to mitigate supplier risk. They show that in a random yield model (similar to the intermittency in our case), process improvement can be favored over dual sourcing if the reliability heterogeneity is high among suppliers.

The paper closest to our setting is Sting and Huchzermeier (2012). The authors consider a manufacturer who invests capacity in a responsive, onshore facility and also replenishes from an offshore supplier who is unreliable but less expensive. After demand and supply uncertainties are realized, the manufacturer orders from its responsive capacity to satisfy the demand. They characterize the optimal production policy and show that the service level is determined by the critical fractile of the responsive capacity. We extend these results in that we consider three sources: the two reliable sources, i.e., flexible and inflexible, can be viewed as an onshore and an (reliable) offshore supplier, respectively; the intermittent renewable source can be viewed as an (unreliable) onshore source. Different from their findings, our results suggest that not all sources are substitutes. Interestingly, any two source combination (e.g., flexible and renewable, or flexible and inflexible) of our model gives the same result as Sting and Huchzermeier (2012), indicating that our model extends their model into a three source setting.

2.2 Supply Chain Flexibility

Our definition of operational flexibility is similar to volume flexibility, where the production quantity can be altered, perhaps at a cost, depending on the realized demand. In this domain, Van Mieghem and Dada (1999) consider postponing the production decision in a single source setting. Tomlin (2006) finds the importance of volume flexibility for a firm that can either source from an unreliable supplier or a reliable and flexible supplier. A similar notion of volume flexibility is quick response, where a firm can place additional orders after observing some initial information on demand (c.f., Fisher and Raman 1996, Fisher et al. 2001, and Bensoussan et al. 2011). Goyal and Netessine (2011) also analyze volume flexibility by comparing the profit of a firm that invests in a volume flexible or a dedicated source. We consider three sources where each source is either flexible, inflexible, or unreliable to study the interaction between these sources. As noted by Goyal and Netessine (2011),
the notion of flexibility that we consider is different from the flexibility in production facilities because energy sources are either fully flexible or fully inflexible.

Another flexibility type is the product mix or process flexibility, studied first by Fine and Freund (1990). This type of flexibility refers to the ability to manufacture different products at the same facility. In this domain, Jordan and Graves (1995) introduce the concept of long chain and show that limited flexibility, if configured in the right way, can perform almost as good as full flexibility (i.e., all products can be produced in all facilities). Van Mieghem (1998) studies optimal capacity investments in two inflexible (dedicated) and one flexible source to meet stochastic demand for two products. Bish and Wang (2004) extend the model of Van Mieghem (1998) by considering price flexibility and show that similar insights hold. Bernstein et al. (2007) study the impact of decentralization in an assemble-to-order system where two dedicated and one common component is used to produce two end-products. They show that decentralization, i.e., component capacity levels being determined by independent suppliers, might reduce the value of operational hedging. In a similar setting, Bernstein et al. (2011) investigate how aggregating demand information affects the profit level and the allocation of the common component between the two end-products. Tomlin and Wang (2008) investigate flexibility in product-mix and pricing in a two-product, two customer class setting. In the process mix flexibility literature, the two inflexible sources are complements with each other and the flexible source is a substitute for both of them (c.f., Van Mieghem 1998). On the other hand, in our context, we find that the renewable and flexible sources are complements and the inflexible source is a substitute for both.

2.3 Sustainable Operations and Energy Economics

Our paper is directly related to the growing literature on sustainable operations (see Drake and Spinler 2013 for a review) and particularly sustainability of energy systems. Many papers including, Aflaki and Netessine (2015), Hu et al. (2015), and Kök et al. (2015) model capacity investments in renewable and conventional sources. Their results indicate that conventional and renewable sources are substitutes. We refine this conclusion by modeling operational flexibility of conventional sources and show that a renewable and a conventional source are substitutes (complements, respectively) if the conventional source is inflexible (flexible, respectively). In an earlier study, Gardner and Rogers (1999) investigate the capacity investment problem under different lead times of construction for
power plants. They do not consider renewable energy investments (hence supply uncertainty) and operational flexibility, which jointly constitute the main focus of our paper. Wu and Kapuscinski (2013) also consider operational flexibility to determine ways to cope with intermittency, but capacity investment is not endogenous in their model.

There is an extensive literature in economics dealing with energy investments. See Crew et al. (1995) for a review. More recently, through a simulation study, Chao (2011) observes that wind energy is substituted by combined cycle, natural gas turbines (i.e., inflexible source) and complemented by regular gas turbines (i.e., flexible source). Lee et al. (2012) also point to potential reasons for complementarity between natural gas and wind investments. Baranes et al. (2015) focus on a similar question as ours but they only consider investments in a renewable source under deterministic demand without modeling operational flexibility.

3 Model

To facilitate the formulation of our model, we first describe how a monopolist utility firm makes its capacity investment decision in practice. A typical process starts from forecasting the electricity demand and the intermittency of renewable energy supply in a targeted demographic region. Using these demand and supply forecasts as an input, the utility firm makes a strategic decision on its investment level in inflexible, renewable, and flexible energy sources. The invested capacity level of a source becomes a constraint for the energy output from that source. In the daily operations, the utility firm’s objective is to match the random demand with electricity supply for each operating period, which is often set to be five minutes. The utility firm uses the five-minute ahead forecasts of demand and supply as inputs and decides how much electricity to generate from the renewable and flexible sources in each operating period. The inflexible source, on the other hand, is dispatched at a constant level throughout the day. This is because a utility firm cannot frequently change the output of an inflexible source due to technical reasons (c.f., Shively and Ferrare 2008, p. 39, Denholm et al. 2010, and DOE 2011).

The utility firm uses the short-term forecasts as inputs in the dispatch decision because they are quite accurate. In Figure 1, we plot the Mean Absolute Percentage Error (MAPE) for demand.

\footnote{In most cases, renewable energy is not curtailed. That is, the entire capacity of the renewable source is dispatched as its generation cost is negligible.}
and intermittency forecasts in 2014 for the Southwest Power Pool (SPP), the network of utility firms in the southwest region of the U.S. Each circle represents one of the 288 operating periods (i.e., five-minute intervals) during a day. On the vertical axis, we plot the MAPE of day-ahead forecasts, and on the horizontal axis, we plot the MAPE of five-minute ahead forecasts. All circles remain well above the 45 degree line, indicating that the forecasts made five-minutes ahead are much more accurate compared to the forecasts made a day ahead. Thus, a typical utility firm has relatively reliable forecasts of demand and supply before determining the dispatch quantities.

![Figure 1: Forecast Errors in Southwest Power Pool (SPP), 2014](image)

The costs involved in the above process are the investment costs and the generation costs of electricity. Specifically, the generation cost of the renewable source is negligible and the generation cost of inflexible sources, such as nuclear or coal-fired power plants, is usually smaller than that of flexible sources, such as natural gas. In some rare occasions, blackouts occur if the demand cannot be fulfilled by the dispatched supply. Blackouts are costly because a utility firm usually has to purchase electricity from external sources to avoid fines imposed by governmental regulations. The objective of the utility firm is to minimize the total cost consisting of the investment costs, generation costs, and penalty costs due to potential blackouts. We refer to the latter two costs as the generation-related costs.

We formulate our problem as a stochastic program with recourse based on the above practice. We consider a representative day with $N$ operating periods. That is, on the operational level, we consider multi-period dispatch decisions. On the strategic level, we only focus on one-time capacity investments. The problem consists of two stages: the first stage is related to the initial capacity investments.
investment decision and the second stage is related to the dispatch decision to match demand with supply. Let the variable generation cost (in dollars per unit capacity for a period) of the inflexible and flexible sources be $c_I$ and $c_F$, respectively. We normalize the variable generation cost of the renewable source to zero ($c_R = 0$). The distribution of the five-minute ahead demand forecast is given as a bounded, nonnegative random variable $\epsilon_n$. Available capacity of renewable energy is denoted as $\Theta_n k_R$, where $\Theta_n$ is a random variable with a support of $[0,1]$, representing the distribution of intermittency forecast. The sequence of events is illustrated in Figure 2.

We formulate the problem backwards. Let $q_I$, $q_R$, and $q_F$ represent the dispatch levels of the inflexible, renewable, and flexible sources, respectively. Similarly, $k_i$ denotes the investment level in source $i \in \{I, R, F\}$. Any unmet demand results in a penalty cost, with rate $r$, proportional to the amount of electricity demand that cannot be satisfied by the dispatched electricity from the three sources. This linear penalty cost is consistent with the literature (c.f., Crew et al. 1995). The second stage problem of the utility firm is to minimize the sum of the generation-related costs for each period $n$ after observing demand and intermittency forecasts $\xi$ and $\theta$:

$$C(q_I, k_R, k_F, \xi, \theta) = c_I q_I + \begin{cases} \min_{q_R, q_F \geq 0} & c_F q_F + r (\xi - q_I - q_R - q_F)^+ \\ \text{subject to} & q_R \leq \theta k_R \\ & q_F \leq k_F \end{cases},$$

where $(x)^+ = \max\{x, 0\}$.

In the above formulation, the decision variables are the dispatch levels of the renewable and flexible sources, whereas the dispatch level of the inflexible source $q_I$ is given as a state variable. This is to reflect the fact that the inflexible source is dispatched at a constant level and cannot be
adjusted in each period. Thus, we shall view \( q_I \) as a long-term decision, which will be optimized in the subsequent stage 1 problem. This formulation implicitly assumes that the inflexible source will be dispatched earlier than the other two sources, which is consistent with the current practice.\(^2\) Notice that \( \xi \) and \( \theta \) in the above formulation are the point forecasts made five-minutes ahead of the dispatch decision. As explained above, these forecasts are quite reliable. Hence, as in Wu and Kapuscinski (2013), we ignore any remaining uncertainty by considering \( \xi \) and \( \theta \) as realizations of demand and supply uncertainty, respectively.

We now turn to the stage 1 problem. The utility firm determines its nonnegative capacity investment levels along with the dispatch decision of the inflexible source so as to minimize its expected total cost:

\[
\min_{k \in \mathbb{R}^+} \Pi(k) = \alpha_I k_I + \alpha_R k_R + \alpha_F k_F + \min_{0 \leq q_I \leq k_I} E \left[ \sum_{n=1}^{N} C(q_I, k_R, k_F, \epsilon_n, \Theta_n) \right],
\]

(2)

where \( E[\cdot] \) denotes the expectation operator, \( k = (k_I, k_R, k_F) \), and \( C(q_I, k_R, k_F, \epsilon_n, \Theta_n) \) is the solution of the second stage problem given in (1). The expectation is with respect to non-stationary random variables \( \epsilon_n \) and \( \Theta_n \), which represents the demand and supply uncertainty, respectively, in the planning stage for the utility firm. Here, in addition to the capacity investment levels, the utility firm determines the dispatch level of the inflexible source.

Notice that we consider a monopolist utility firm that does not have access to an electricity spot market in our base model. In this setting, the firm is responsible for matching supply and demand by using its own generation sources. This is not an uncommon setting because approximately half of the U.S. utility firms operate in such settings (FERC 2015b). In Section 8, we extend our model by incorporating an electricity spot market.

In the remainder of the paper, we use terms “increasing,” “decreasing,” and “convex” in the weak sense. We denote the gradient operator as \( \nabla \). For a random variable \( X \), we let \( f_X(\cdot) \) be the probability density function. Finally, “\( X|\cdot \)” denotes the conditional probability. All proofs are given in Appendix B.

\(^2\)In fact, it is also optimal to first dispatch the renewable source in this case. That is, it is also optimal to set \( q_R = \theta k_R \) in all periods as \( c_R = 0 \) and there is no overproduction penalty. Nevertheless, we explicitly consider \( q_R \) to ensure consistency with the second stage problem of the spot market setting given in (14)–(17). In the presence of a spot market, it might not be optimal to dispatch all renewable capacity as we explain in Section 8.
4 Optimal Capacity Investments

In this section, we characterize the optimal capacity investments of a utility firm. We first simplify the problem given in (2) by showing that at optimality, the dispatch level of the inflexible source is always equal to its capacity investment level, i.e., \( q_I = k_I \). The intuition is that the firm should always dispatch all of its inflexible capacity at every period because the firm can otherwise achieve a strictly lower cost by decreasing \( k_I \).

**Lemma 1.** Consider the investment problem given in (2). It is optimal to set \( q_I = k_I \).

Lemma 1 is consistent with the practice as the utilization of nuclear power plants in the U.S. is close to 90% (EIA 2015b), indicating that these plants operate continuously during the day. By using Lemma 1, we substitute \( k_I \) for \( q_I \) in the second stage dispatch problem given in (1) and obtain:

\[
C(k, \xi, \theta) = \min_{q_R, q_F \geq 0} \quad c_F q_F + r (\xi - k_I - q_R - q_F)^+ \quad \text{(3)}
\]

subject to

\[
q_R \leq \theta k_R \quad \text{(4)}
\]

\[
q_F \leq k_F. \quad \text{(5)}
\]

Similarly, under Lemma 1, the capacity investment problem in the first stage becomes:

\[
\min_{k \in \mathbb{R}_+^3} \Pi(k) = E \left[ \sum_{n=1}^N C(k, \epsilon_n, \Theta_n) \right] + (\alpha_I + c_I N) k_I + \alpha_F k_F + \alpha_R k_R, \quad \text{(6)}
\]

where we charge the generation cost of the inflexible source to its entire capacity for each of the \( N \) periods. In the remainder of the paper, we focus on these simplified formulations of the first and second stage problems.

We next characterize the optimal capacity investments by backward induction, i.e., by first solving the second stage problem given in (3)–(5). Let \( q_i^*(k, \xi, \theta) \) be the optimal dispatch level of energy source \( i \in \{ R, F \} \) given an investment vector \( k \), demand forecast \( \xi \), and intermittency forecast \( \theta \). The optimal dispatch policy for renewable and flexible sources is shown in the following lemma.

**Lemma 2.** Consider the dispatch problem given in (3)–(5). The optimal dispatch policy is to set \( q_R^*(k, \xi, \theta) = \min (\theta k_R, \xi - k_I)^+ \) and \( q_F^*(k, \xi, \theta) = \min (k_F, \xi - k_I - \theta k_R)^+ \).
Lemma 2 shows that the utility firm first dispatches its renewable source up to its available capacity \( \theta k_R \) if demand forecast \( \xi \) exceeds the inflexible source capacity \( k_I \) in a period. Then, the flexible source is dispatched for the remaining demand. This is due to the fact that the renewable source incurs a negligible generation cost compared to the flexible source. Lemmas 1 and 2 conclude the optimal dispatch policy: in every period, all of the inflexible capacity is dispatched, followed by the renewable source, and then by the flexible source.

We next use this optimal dispatch policy to characterize the optimal capacity portfolio. Our analysis involves constructing the dual of the dispatch problem in (3)–(5) such that \( \lambda_i^* (k, \xi, \theta) \) denotes the optimal dual variable associated with the capacity constraint related to source \( i \in \{ I, R, F \} \). We present this dual problem in the proof of Proposition 1, where each dual variable represents the shadow price of the associated capacity constraint.

Table 1: Shadow Prices of Capacity Constraints for Demand and Intermittency Space Partitions

<table>
<thead>
<tr>
<th>Partition for ((\xi, \theta) \in \mathbb{R}_+ \times [0, 1])</th>
<th>( \lambda_I^* (k, \xi, \theta) )</th>
<th>( \lambda_R^* (k, \xi, \theta) )</th>
<th>( \lambda_F^* (k, \xi, \theta) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \Omega_1 (k) = { (\xi, \theta)</td>
<td>\xi \leq k_I + \theta k_R } )</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>( \Omega_2 (k) = { (\xi, \theta)</td>
<td>k_I + \theta k_R \leq \xi \leq k_I + \theta k_R + k_F } )</td>
<td>( c_F )</td>
<td>( \theta c_F )</td>
</tr>
<tr>
<td>( \Omega_3 (k) = { (\xi, \theta)</td>
<td>k_I + \theta k_R + k_F \leq \xi } )</td>
<td>( r )</td>
<td>( \theta r )</td>
</tr>
</tbody>
</table>

Lemma 2 is obtained by solving the dispatch problem based on the realizations of demand and supply uncertainty. There are three regions of the uncertainty space in each of which the optimal dispatch decision as well as the dual variables have the same structure. We present these regions in Table 1. For example, in \( \Omega_1 \), \( \xi \) and \( \theta \) are such that \( \xi \leq k_I + \theta k_R \), i.e., the demand is less than the sum of the inflexible and available renewable capacity. Then, it is optimal to set \( q_R = (\xi - k_I)^+ \) and \( q_F = 0 \) as also indicated by Lemma 2. Furthermore, in this case, no capacity constraint is binding so that all dual variables are zero. The uncertainty regions are identical across all \( N \) periods but the probability that a pair of \( \xi \) and \( \theta \) falls into a specific region in each period depends on the (non-identical) distributions of \( \epsilon_n \) and \( \Theta_n \). In addition, since \( \bar{\Pi}(k) \) is convex, the Karush-Kuhn-Tucker (KKT) conditions are necessary and sufficient for the investment problem given in (6). Moreover, we can show that \( \nabla_k E[C(k, \xi, \theta)] = -E[\lambda(k, \xi, \theta)] \). That is, the derivative and expected value can be interchanged, where the expected value of the dual variables can be easily computed by using Table 1. With these observations, in Proposition 1, we present the KKT conditions of the investment problem given in (6), where \( v \) is the vector of Lagrange multipliers of the nonnegativity
Proposition 1. Consider the problem given in (3)–(6). An investment vector $\mathbf{k}^* \in \mathbb{R}^3_+$ is optimal if and only if there exists a $\mathbf{v} \in \mathbb{R}^3_+$ such that

$$\sum_{n=1}^{N} \begin{bmatrix} c_F & \Theta_n c_F & 0 \\ r & \Theta_n r & 0 \\ c_F & r - c_F & 0 \end{bmatrix} \begin{bmatrix} \Omega_2 (\mathbf{k}^*) \\ \Omega_3 (\mathbf{k}^*) \end{bmatrix} P^n (\Omega_2 (\mathbf{k}^*)) + E \begin{bmatrix} c_F & \Theta_n c_F & 0 \\ r & \Theta_n r & 0 \\ c_F & r - c_F & 0 \end{bmatrix} \begin{bmatrix} \Omega_2 (\mathbf{k}^*) \\ \Omega_3 (\mathbf{k}^*) \end{bmatrix} P^n (\Omega_3 (\mathbf{k}^*)) = \begin{bmatrix} \alpha_I + Nc_I - v_I \\ \alpha_R - v_R \\ \alpha_F - v_F \end{bmatrix}.$$

(7)

$$\forall i \in \{I, R, F\} : k_i v_i = 0.$$  

(8)

Equation (7) is obtained by taking the partial derivative of the Lagrangian function with respect to $k_I, k_R,$ and $k_F,$ respectively. Based on the KKT conditions, there are a total of eight cases that we should consider in order to find the optimal investment levels. These eight cases form four investment strategies: (i) no investments (i.e., $\mathbf{k}^* = 0$), (ii) single sourcing (three cases, e.g., $k^*_I > 0$ and $k^*_R = k^*_F = 0$), (iii) dual sourcing (three cases, e.g., $k^*_I, k^*_R > 0$ and $k^*_F = 0$), (iv) triple sourcing (i.e., $\mathbf{k}^* > 0$). No investments strategy is optimal if $rN < \alpha_i + c_i N$ for $i \in \{I, F\},$ and $r \sum_{n=1}^{N} E[\Theta_n] < \alpha_R,$ i.e., when the investment costs are higher than the penalty cost. Unfortunately, we are not able to analytically characterize the range of the cost parameters that ensures the optimality of the rest of the investment strategies due to the nonstationarity in demand and supply uncertainty. Nevertheless, based on the estimates of the cost parameters and the electricity demand data of Texas, we observe that the triple sourcing strategy is optimal. This is consistent with the practice that utility firms simultaneously invest in inflexible, renewable, and flexible sources (FERC 2015a). Motivated by these facts, in the subsequent discussion, we shall focus on the triple sourcing strategy as this is the most interesting and relevant case. We also investigate the other strategies in Section 8.

Proposition 1 provides a method to find the optimal investment levels for the triple sourcing investment strategy. The idea is to solve three newsvendor problems simultaneously with $\mathbf{v}^* = 0$ in (7), each corresponding to one energy source. Specifically, for the inflexible source, the underage cost includes the expectation of the two events associated with the demand exceeding the capacity of this source. In the first case, the capacity of the flexible source is sufficient to meet the remaining constraints and $P^n(\Omega)$ is the probability that $\xi$ and $\theta$ is in $\Omega$ for $\epsilon_n$ and $\Theta_n$. 

14
demand. In the second, the total demand may exceed the entire capacity and a penalty cost \( r \) is incurred in addition to the generation cost of the flexible source. Hence, the underage cost for the inflexible source is the probability weighted sum of these two costs. The overage cost for the inflexible source, on the other hand, is the investment and the generation cost. Note that we include the generation cost of the inflexible source in the overage cost because the entire capacity of this source is dispatched at every period even if its capacity exceeds the demand.

For the renewable source, the underage cost is similar to the inflexible source. However, supply uncertainty \( \Theta_n \) is also considered while computing the expectation. The overage cost only includes the investment cost but not the variable generation cost for two reasons. First, we assume that the variable cost is zero for the renewable source. Second, even in the absence of this assumption, the utility firm would not dispatch the renewable source when its capacity exceeds the demand, so the variable generation cost should not be included in the overage cost.

For the flexible source, the underage cost only involves the event of demand exceeding the total capacity. In this case, the penalty cost is incurred and the underage cost is given as \( (r - \alpha_F - c_F) \). Note that we deduct the investment and generation cost from the penalty cost, i.e., as in the classical Newsvendor model, we consider the net underage cost. The overage cost for the flexible source is only the capacity cost, \( \alpha_F \).

In summary, the optimality condition suggests that there is a pair of underage cost and overage cost that determines the optimal investment level for each energy source. The utility firm balances underage and overage costs of inflexible, renewable, and flexible sources for demand and supply realizations as we illustrate in Figure 3. The thick line in the figure represents the maximum demand that the firm is able to serve. By adjusting its investments, the utility firm determines the probability of each region so that the underage cost is balanced with the overage cost for each energy source.

Next, we consider the relationship between the investments and the reliability of the electricity grid. In the energy economics literature, reliability is defined based on the so-called loss-of-load probability (LOLP), i.e., the probability that the demand exceeds the supply of electricity (Telson 1975). This definition is similar to the concept of service level in the supply chain management.
Notes. In Figure 3, for illustration purposes, we assume that $k_F > k_R$.

literature. Let $\rho^*$ denote the LOLP corresponding to the optimal investment levels:

$$\rho^* = \sum_{n=1}^{N} P^n(\Omega_3(k^*)),$$

where $\Omega_3(k^*)$ is the demand and intermittency space region in which demand exceeds available supply.

**Corollary 1.** $\rho^* = \alpha_F / (r - c_F)$.

In the triple source strategy (i.e., $k^* > 0$), this corollary immediately follows from the third dimension of the optimality condition (i.e., with respect to $k_F$) given in (7) in Proposition 1. It suggests that the reliability of the electricity grid is only affected by the penalty cost rate $r$ and the cost parameters of the flexible source. That is, the newsvendor critical fractile of the flexible source determines the service level. Intuitively, the flexible source is the last option for the utility firm to satisfy the demand and the firm finds the optimal investment level in this source by comparing the penalty cost of not satisfying demand and the investment cost. This result is an extension of the similar observations made in the energy economics (c.f., Chao 1983) and the dual sourcing literature (c.f., Sting and Huchzermeier 2012) to our setting.

Corollary 1 suggests an important policy insight. Because subsidies for the renewable or the inflexible source do not affect $r$, $c_F$, or $\alpha_F$, these subsidies do not change the reliability level of the grid. This result provides a different perspective than the claims that renewable energy subsidies undermine the reliability and nuclear subsidies enhance the reliability (c.f., Gronewold
2011, Garman and Thernstrom 2013, Karnitschnig 2014, Fisher 2015, and Smith 2015). This is because, our model optimizes investments in all energy sources simultaneously and can identify the impact of subsidies on the entire capacity portfolio rather than considering the impact on a single source.

5 Interaction Between Energy Sources

In this section, we investigate how providing a subsidy for one energy source affects the investment in other sources. Two consumption goods are substitutes if a decrease in the price of one good leads to a lower level of consumption in the other (Singh and Vives 1984). From the utility firm’s perspective, energy sources are consumption goods and their price is the investment cost. Hence, we define two energy sources as substitutes if a decrease in one’s investment cost leads to a decrease in the other’s investment level. That is, sources $i$ and $j$ are substitutes if a decrease in $\alpha_i$ leads to a decrease in $k_j^*$ (i.e., $dk_j^*/d\alpha_i > 0$) and vice-versa (i.e., $dk_i^*/d\alpha_j > 0$). Analogously, we define two sources as complements if a decrease in one’s investment cost leads to an increase in the other’s investment level. We refer to the decrease in the investment cost as an investment subsidy. In practice, such a decrease is not necessarily limited to the subsidies provided by the government but can also represent a technological breakthrough that reduces the cost of investment. For example, a new technology has reduced investment cost for coal-fired power plants (Duke Energy 2015b), which can be considered as a decrease in $\alpha_I$. We first present a preliminary result before identifying the interaction between energy sources (i.e., how a subsidy for one source affects investment in others).

**Proposition 2.** For $i, j \in \{I, R, F\}$, (i) $\frac{dk_i^*}{d\alpha_i} \leq 0$, (ii) $\frac{dk_i^*}{d\alpha_j} = \frac{dk_j^*}{d\alpha_i}$.

Proposition 2 part (i) shows that providing a subsidy for an energy source leads to a higher investment level in that source. Intuitively, the subsidy leads to a lower investment cost, in response, the utility firm increases its investment. Part (ii) shows that the cross effect of a subsidy is symmetric: the change in the investment level for source $i$ in response to a change in the investment cost of source $j$ is equivalent to that for source $j$ in response to a change in the investment cost of source $i$. Next, we present our main result under the following assumption, which we impose in the remainder of the paper.
Assumption 1. (i) In each period $n$, $\epsilon_n$ is independent of $\Theta_n$. (ii) Demand distribution $\epsilon_n$ is strictly log-concave. (iii) Intermittency distribution $\Theta_n$ follows a Bernoulli distribution:

$$
\Theta_n = \begin{cases} 
1 & \text{with probability } q_n \\
0 & \text{with probability } 1-q_n 
\end{cases} \quad (9)
$$

The first part of this assumption is consistent with the forecast errors presented in Figure 1. This is because the correlation coefficient of demand and intermittency forecast errors are $-0.0219$, indicating no linear dependence. The second part of the assumption requires $\log f_{\epsilon_n}(\cdot)$ to be a strictly concave function. Many well-known probability distributions including Normal, Logistic, Extreme Value, and Gamma (with the shape parameter greater than one), satisfy this condition (Bagnoli and Bergstrom 2005). Finally, because we consider $N$ periods with each period corresponding to a five-minute interval, Assumption 1 part (iii) is not too restrictive. This assumption is also made in the literature for intermittency of renewables (c.f., Aflaki and Netessine 2015, Baranes et al. 2015, and Køk et al. 2015) as well as supply disruptions (c.f., Tomlin and Wang 2005, Tomlin 2006, and Yang et al. 2012). Moreover, this assumption can be relaxed, as we explain in Section 8, if the demand is assumed to be stationary. Below we present our main result.

Proposition 3. (i) The inflexible and renewable sources are substitutes. (ii) The inflexible and flexible sources are substitutes. (iii) Suppose $q_n = q$ for all $n$ and $g(\cdot) = \sum_{n=1}^{N} f_{\epsilon_n}(\cdot)$ is log-concave, then the renewable and flexible sources are complements.

Proposition 3 (i) and (ii) indicate that a subsidy for the inflexible source leads to a lower investment level in the renewable and flexible sources. However, Proposition 3 (iii) shows that a subsidy for the flexible source leads to a higher investment in the renewable source under two sufficient conditions. First, the intermittency distribution needs to be stationary. Second, the sum of the density functions for demand over $N$ periods is required to be log-concave. Notice that the sum of log-concave density functions is not necessarily log-concave. Although these two conditions are required for tractability, the case study presented in Section 7 reveals that our insights hold in the general case when real electricity demand and generation data is used. Moreover, in the absence of these two conditions, in Proposition 4, we show that the results of Proposition 3 hold as long as the demand follows a log-concave distribution and is independent and identical across
Proposition 4. Suppose $\epsilon_n$ is log-concave and independent and identically distributed. (i) The inflexible and renewable sources are substitutes. (ii) The inflexible and flexible sources are substitutes. (iii) The renewable and flexible sources are complements.

The intuition behind Proposition 3 and 4 can be explained by considering a utility firm that forms a portfolio with its capacity investments. The inflexible source and the renewable source share similar characteristics in this portfolio, as both of these sources have high investment and low generation costs. Furthermore, these sources are not responsive to the demand because the renewable source is intermittent and the utility firm cannot increase the output of the inflexible source on-demand. Hence, these two sources are substitutes. On the other hand, the renewable source and the flexible source possess opposite features: the flexible source has high generation and low investment cost, and its output can be ramped up or down quickly according to the demand. Thus, the flexible source complements the renewable source. Finally, the inflexible and the flexible sources are substitutes because the flexible source already complements the renewable source.

6 Effects of Other Policies

In this section, we first consider the effects of a decrease in the generation cost of inflexible and flexible sources. Such a decrease can be either due to a governmental subsidy policy or decreased input prices in commodity markets. For example, natural gas prices in the U.S. have fallen considerably in recent years due to the increase in the supply of shale gas (Puko 2015). This corresponds to a decrease in $c_F$ in our model.

Proposition 5. (i) A decrease in the generation cost of the inflexible source leads to a higher investment level in the inflexible source but a lower investment level in the renewable and flexible sources. (ii) A decrease in the generation cost of the flexible source leads to a lower investment level in the inflexible source but a higher investment level in the flexible source.

Proposition 5 part (i) indicates that the effects of a generation subsidy and an investment subsidy for the inflexible energy source are equivalent. That is, we find that, both subsidies result in the same amount of change in investment levels of the three sources. Furthermore, part (ii)
shows that the effect of a generation subsidy and an investment subsidy for the flexible source are directionally the same on the inflexible source. Unfortunately, we are not able to analytically characterize the impact of a generation subsidy for the flexible source on the renewable energy investment. Nevertheless, we numerically observe that, similar to the investment subsidy for the flexible source, a generation subsidy for the flexible source also leads to a higher investment level in the renewable source. The intuition behind these findings remains to be the substitution and complementarity effects between these energy sources.

We next investigate the effects of the carbon tax policy, which, in various forms, has been adopted by many countries to reduce carbon emissions. Consider a carbon tax given as \( t \), and let the emission intensity of a source be denoted as \( e_i \) for \( i \in \{I, F\} \). The emission intensity of the renewable source is zero, i.e., \( e_R = 0 \) because a utility firm does not burn any fossil fuel to generate electricity from a renewable source. Under the carbon tax, the investment problem of the utility firm becomes:

\[
\begin{aligned}
\min_{k \in \mathbb{R}^3_+} \Pi(k) &= (\alpha_I + (c_I + te_I)N)k_I + \alpha_F k_F + \alpha_R k_R + E \left[ \sum_{n=1}^N C(k, e_n, \Theta_n) \right], \\
\end{aligned}
\]

(10)

where

\[
C(k, \xi, \theta) = \min_{0 \leq q_R \leq \theta k_R, 0 \leq q_F \leq k_F} (c_F + te_F)q_F + r (\xi - k_I - q_R - q_F)^+. 
\]

(11)

**Proposition 6.** Assume that \( e_I = 0 \), that is, the inflexible source is carbon-free such as nuclear energy. Then, a carbon tax leads to a higher investment level in the inflexible source and a lower investment level in the flexible source.

Proposition 6 shows that in response to a carbon tax, a utility firm increases its investment in the inflexible source and decreases its investment in the flexible source, provided that the inflexible source is carbon-free (e.g., nuclear energy). This is because carbon tax increases the cost of generating electricity from the flexible source, whereas it does not affect the inflexible source. In this case, although we cannot analytically characterize the effect of the tax on the renewable investment, one can conjecture that the tax would lead to a lower investment in the renewable source due to the complementarity effect with the flexible and the substitution effect with the inflexible source. We present a numerical study in Figure 4a that confirms this intuition by plotting optimal investment levels in response to the carbon tax when \( e_I = 0 \). In addition, we present another numerical study
in Figure 4b assuming that the inflexible source is carbon-intensive such as coal power. In this case, both the inflexible and flexible source are taxed so the overall impact on the renewable source is not clear and depends on the emission intensity of the sources. As seen in Figure 4b, the carbon tax leads to a lower investment in the inflexible source, whereas investments in the flexible and renewable sources are higher.

Figure 4: Effect of Carbon Tax on Optimal Investment Levels for Different Inflexible Sources

Notes. In Figure 4, we set $N = 1$, $e_F = 1.21$ and for coal $e_I = 2.07$ (EIA 2015a). We use similar cost parameters as in the case study. Note that, compared to panel (a), we report a limited range of $t$ levels in panel (b). This is because for high carbon tax levels, coal power becomes uneconomic and the utility firm does not invest in it.

We close this section by considering the effect of the penalty rate $r$, which corresponds to the cost when the utility firm is not able to satisfy the demand. One way to reduce such a cost is to introduce a real-time electricity spot market, such as the Energy Imbalance Market of Southwest Power Pool. Such markets enable utility firms to buy and sell electricity, effectively reducing the costs associated with demand and supply imbalance. We present the effect of $r$ below. Recall that $g(\cdot) = \sum_{n=1}^{N} f_{\epsilon_n}(\cdot)$.

**Proposition 7.** Suppose $q_n = q$ for all $n$ and $g(\cdot)$ is log-concave. Then, a decrease in the penalty rate $r$ leads to (i) a decrease in the investment levels of the inflexible and the flexible sources, and (ii) an increase in the investment level of the renewable source.

Proposition 7 shows that penalty rate $r$ affects investment in the inflexible and flexible sources similarly: investment in both conventional sources decrease with a decrease in $r$. The effect
opposite for the renewable source: the utility firm increases its renewable investment if $r$ decreases. This is because as $r$ decreases, demand and supply mismatch becomes less costly, enabling the utility firm to make more (risky) investment in the intermittent renewable source rather than the conventional sources. This result indicates that an electricity spot market can help a utility firm in dealing with intermittency. Motivated by this observation, we consider an extension of our model with an electricity spot market in Section 8.

7 Case Study: Texas Data

In this section, we validate our main insight by using real electricity generation and demand data from the state of Texas in 2010. In our analytical model, presented in (3)–(6), we assume that, between consecutive periods, the output of a flexible source can be changed at any rate and the output of an inflexible source cannot be changed at all. However, in practice, operational flexibility depends on plant-level generation characteristics. For example, there are limits on how fast the output of a flexible source can be ramped up. In this case study, by considering these generation characteristics, we validate our conclusion on the complementarity and substitution effects between energy sources.

Table 2: Sample Plant Characteristics for Texas, 2010

<table>
<thead>
<tr>
<th>Plant Name</th>
<th>Fuel Type</th>
<th>Startup Cost ($)</th>
<th>Minimum Output (MW)</th>
<th>Ramp Up Limit (%/min)</th>
<th>Minimum Down Time (hr)</th>
</tr>
</thead>
<tbody>
<tr>
<td>South Texas Project</td>
<td>Nuclear</td>
<td>15,000</td>
<td>812</td>
<td>1</td>
<td>168</td>
</tr>
<tr>
<td>Morgan Creek</td>
<td>Natural Gas</td>
<td>1,203</td>
<td>122</td>
<td>10</td>
<td>0.5</td>
</tr>
</tbody>
</table>

Table 2 illustrates generation characteristics that determine operational flexibility for a representative nuclear and natural gas power plant in Texas (Cohen 2012). Here, startup cost, minimum output, and minimum downtime are all greater for the nuclear power plant compared to the natural gas plant. Furthermore, a utility firm can increase the output of the natural gas plant by 10% of its capacity every minute, whereas nuclear can only be ramped at a rate of 1% of its capacity. In practice, a utility firm takes these salient features into account and determines the least costly way of satisfying the electricity demand with its available set of generators. In doing so, the firm uses a so-called unit commitment and dispatch model (UCDM), a mixed integer program that minimizes
the generation cost subject to electricity system constraints such as capacity limits, ramp up/down constraints, and minimum up/down times. We use Cohen (2012)’s dispatch model that mimics the operations in Texas electricity system to determine how providing a subsidy for one source affects the investment in others.

Next, we describe the data required to run the UCDM. As an input, the UCDM uses the demand data and generation characteristics of available power plants. We use the observed 15-minute demand data from the state of Texas in 2010. For generation mix (available set of power plants), we use the same data sources as in Kök et al. (2015). That is, we utilize the rich dataset given in Cohen (2012) that reports various generation characteristics including those related to the operational flexibility of the 144 conventional power plants in Texas. Furthermore, as the renewable source, we consider wind energy and use the 15-minute output data which is also provided by Cohen (2012).

We now turn to our analysis. We first determine the optimal capacity investment in inflexible, renewable, and flexible sources in Texas electricity system for current estimates of investment costs. Then, to investigate how optimal investment levels change, we decrease the investment cost of each source sequentially, which corresponds to providing a subsidy for each source. Specifically, the utility firm minimizes its generation and investment cost by determining its investment level in the three energy sources:

$$\min_{k_I, k_R, k_F} \Pi(k_I, k_R, k_F) = G(k_I, k_R, k_F) + \alpha_I k_I + \alpha_R k_R + \alpha_F k_F,$$

where $G(k_I, k_R, k_F)$ corresponds to the output of the UCDM given inflexible, renewable, and flexible source investment of $k_I, k_R, \text{and } k_F$, respectively. The remaining terms are the investment costs. In essence, we use the UCDM instead of the second stage problem of the optimization model presented in (3)–(6).

To determine the optimal investment levels, we next evaluate $G(k_I, k_R, k_F)$ at various $k_I, k_R, \text{and } k_F$ levels. Each evaluation takes 1.2 CPU hours on average; hence, we only consider a limited set of investment levels. In particular, we take the current level of investment in energy sources as a basis and evaluate $G(k_I, k_R, k_F)$ at current investment levels as well as when additional investments are made. We consider nuclear energy as the inflexible source and open cycle gas turbines as the flexible source. We allow additional investments of $\{0, 1000, 3000, 5000\} \text{ MW}$ for both con-
ventional sources. As the renewable source, we consider wind energy with additional investments of \{0, 5000, 10000, 15000, 20000\} MW. The reason that we consider a maximum investment level of 20,000 MW for wind energy is to ensure that the actual generation from the renewable and conventional sources are similar. This is because wind source is intermittent with a capacity factor of approximately 0.3, meaning that the maximum effective capacity is \(6,000 = (20,000 \times 0.3)\) MW for wind energy.

In summary, we enumerate \(G(k_I, k_R, k_F)\) for 80 cases (=4 levels of nuclear investments by 4 levels of natural gas investments by 5 levels of renewable investments). Among these cases, under current cost estimates, we identify the optimal investment levels by selecting the one with the lowest cost. Then, we sequentially provide a 30\% subsidy for each of the sources and compare the new investment levels against the original investments. We report our main findings in Figure 5 and further details are provided in Appendix A.

![Figure 5: Effect of Subsidies](image)

Figure 5 plots the change in the optimal investment levels when a subsidy is provided for a source compared to the original investments. We note that, for wind energy, we report the effective investment level which accounts for the intermittency of wind energy by multiplying its optimal investment level with its capacity factor as described above. In the leftmost panel, we observe that
providing a subsidy for the nuclear energy source leads to an increase in the capacity of that source and a decrease in the capacity of renewable and flexible sources. In the middle panel, we consider a subsidy for wind energy. In this case, inflexible investment decreases, whereas the investment in renewable and flexible sources increases. Finally, in the rightmost panel, we observe that a subsidy for the natural gas results in a lower investment in the inflexible source whereas investments in other sources increase. These results validate Proposition 3 as the same complementarity and substitution effects are found between energy sources.

To sum, in this case study, we use a practical dispatch model to refine our definition of operational flexibility. We observe that our main insight still holds in this setting. That is, from the perspective of a utility firm, renewable and flexible sources are complements, whereas renewable and inflexible sources are substitutes in a realistic setting that is not subject to the limitations of our analytical model.

8 Extensions

8.1 Spot Market

In our main model given in (3) to (6), we consider a vertically-integrated utility firm that does not participate in a spot market to buy or sell electricity. In practice, more than half of the U.S. utility firms use spot markets, such as the real-time Energy Imbalance Market in SPP (EIA 2011). In this section, we consider the effect of a spot market on the capacity investments of a utility firm.

In an electricity market, a utility firm can procure electricity either from its own generation sources (self-schedule) or from other suppliers through bilateral contracts and spot markets (FERC 2015b, p. 62). The most common way for a utility firm to procure electricity is self-schedule. For example, in the largest electricity market of the U.S. (PJM Interconnect), utility firms have generated more than 60% of their electricity from their own sources in 2014 (Monitoring Analytics 2015, p. 97). The remaining electricity can be purchased from a spot market in which the price varies stochastically during the day. Also, this market, such as the one in PJM, has a relatively thin volume so that the price might be affected by the amount of electricity traded. Considering
these factors, we assume that the utility firm faces the following price in the spot market:

\[ p^S_n (\Gamma, q_S) = \Gamma + \frac{b_n}{2} q_S, \tag{13} \]

where \( \Gamma \) is a random variable representing price uncertainty, \( q_S \) is the amount of electricity bought by the utility firm, and \( b_n > 0 \) is the price responsiveness parameter in period \( n \). We note that \( q_S \) is negative if the utility firm sells electricity in the market, which causes the market price to decrease. On the other hand, if the utility firm buys electricity from the market, \( q_S \) is positive, which causes the market price to increase. We note that our results hold for any positive \( b_n \), that is, our results are robust to the magnitude of the impact of the utility firm on the market price. Similar models are considered for price formation in spot markets in the literature (c.f., Martínez-de-Albéniz and Simchi-Levi 2005).

In the presence of the spot market, we modify the second stage of the utility firm’s problem as:

\[
C_n (k, \xi, \theta, \gamma) = \min_{q_R, q_F \geq 0, q_S \in \mathbb{R}} c_F q_F + p^n_S (\gamma, q_S) q_S
\tag{14}
\]

subject to

\[
q_R \leq \theta k_R \tag{15}
\]

\[
q_F \leq k_F \tag{16}
\]

\[
q_S = \xi - k_I - q_R - q_F. \tag{17}
\]

Following Lemma 1 that it is optimal for the utility firm to dispatch the entire inflexible capacity at every period, the utility firm minimizes its generation and market transaction cost based on the dispatch levels of the renewable and flexible sources as well as the quantity traded in the spot market \( (q_S) \). In this stage, the utility firm observes the forecast of \( \Gamma \) as \( \gamma \). Furthermore, \( q_S \) is defined in (17) as the difference between demand level and dispatched electricity from the utility firm’s own investments. Recall that \( q_S \) is negative if the firm sells electricity in the market. In this case, the second term in (14) i.e., \( p^n_S (\gamma, q_S) q_S \), is also negative, indicating a decrease in the cost for the utility firm. On the other hand, if \( q_S \) is positive, the utility firm buys electricity from the market, and the second term in (14) is positive, indicating an increase in the cost for the utility firm. With this per period cost, the first stage problem can be casted as:

\[
\min_{k \in \mathbb{R}_+^3} \bar{\Pi} (k) = E \left[ \sum_{n=1}^N C_n (k, \epsilon_n, \Theta_n, \Gamma_n) \right] + (\alpha_I + c_I N) k_I + \alpha_F k_F + \alpha_R k_R. \tag{18}
\]
We next derive the optimal dispatch policy. In this case, in addition to supply and demand uncertainties, we also consider a spot price uncertainty, which complicates our analysis considerably.

**Lemma 3.** Consider the dispatch problem given in (14)–(17). (i) The optimal dispatch policy is to set
\[ q_R^*(k, \xi, \theta, \gamma) = \min \left( \theta k_R, \xi - k_I + \frac{\gamma}{b_n} \right)^+ \]
and
\[ q_F^*(k, \xi, \theta, \gamma) = \min \left( k_F, \xi - k_I - \theta k_R + \frac{\gamma - c_F}{b_n} \right)^+ \].
(ii) The first stage problem given in (18) is convex in \( k \).

Lemma 3 is the extension of Lemma 2 to the spot market setting. In this case, in addition to the demand and intermittency forecasts, the optimal dispatch levels also depend on the market price forecast. If the forecast of the market price is too low (\( \gamma \) is small), neither the renewable nor the flexible source is dispatched. That is, unlike the main model, the utility firm might find it optimal not to use all the renewable energy capacity in all periods. As the market price gets higher, the renewable source, followed by the flexible source is dispatched. We also note that the dispatch level of the inflexible source is equal to its capacity as in the main model.3

Next, we present our main result that identifies the interaction between energy sources in the spot market setting. We continue to consider the interior solution case (i.e., \( k^* > 0 \)) and impose Assumption 1 with a slight generalization in the first part. That is, we now consider that the three uncertainty sources (\( \epsilon_n, \Theta_n, \Gamma_n \)) are independent of each other. In addition, to identify the relationship between the renewable and flexible sources we require the following assumption.

**Assumption 2.** (i) The utility firm dispatches all of its available renewable energy in each period, i.e., \( q_R^* = \Theta k_R \). (ii) Demand distribution \( \epsilon_n \) is bounded above by \( D_n \). (iii) Market price uncertainty \( \Gamma_n \) follows a uniform distribution between \( L_n \) and \( U_n \) such that \( L_n \leq -b_n D_n \).

Part (i) of Assumption 2 ignores the possibility that the utility firm does not use (i.e., curtails) its renewable source. This is a good approximation of the practice because curtailment as a fraction of wind capacity is less than 4% in the U.S. in 2014 (Bird et al. 2014). Second part of the assumption bounds the demand distribution from above. This is also not very restrictive because such a distribution can be closely approximated by an unbounded random variable (e.g. Normal) as long as \( D_n \) is large enough compared to the variance (Petruzzi and Dada 1999). The last part of the

---

3Based on this optimal dispatch policy, optimal investment levels can be characterized similar to the multi-dimensional newsvendor solution in the main model. In addition, cross effects of subsidies are equivalent, i.e.,
\[
\frac{d k^*_i}{d \alpha_j} = \frac{d k^*_j}{d \alpha_i}, \forall i, j \in \{I, R, F\}. \]
Proofs of these results are available from the authors.
assumption suggests that the market price follows a nonstationary uniform distribution and it can be negative. Note that negative prices are observed frequently in practice (c.f. Zhou et al. 2015).

**Proposition 8.** (i) The inflexible and renewable sources are substitutes. (ii) The inflexible and flexible sources are substitutes. (iii) Suppose Assumption 2 holds and $\Theta_n$ follows a stationary Bernoulli distribution, then the renewable and flexible sources are complements.

Proposition 8 shows that our main insight also holds when a spot market is considered. That is, the relationship between a renewable and a conventional source is determined by operational flexibility. If the conventional source is inflexible it substitutes the renewable source, otherwise, it complements the renewable source.

### 8.2 Dual Sourcing

Throughout the paper, we assume that the triple sourcing strategy is optimal, i.e., $k^* > 0$. In some cost parameters, a dual sourcing strategy may be optimal (e.g., $k^*_I, k^*_R > 0$ and $k^*_F = 0$). In any dual sourcing case, the two sources included in the optimal portfolio are substitutes. The details of the proof is available from the authors. We note that this conclusion is the same as that of the dual sourcing literature (c.f., Sting and Huchzermeier 2012).

### 8.3 General Intermittency Distribution

In Assumption 1, we impose a two-point intermittency distribution which helps us obtain analytical insights when the demand is nonstationary. If the demand is assumed stationary, i.e., $\epsilon_n = \epsilon$ for all $n$ and $\epsilon$ has a log-concave density, our main results hold for a general intermittency distribution. In particular, Proposition 3 parts (i) - (iii) and Proposition 7 part (i) hold in this case. We show these generalizations in the proofs of the related propositions.

### 8.4 Energy Efficiency and Demand Response

Our model considers capacity investment in energy sources, which is related to managing the supply side of energy systems. Some of the incentives that reduce the demand, such as Energy Efficiency (EE) and Demand Response (DR), can be incorporated into the current model. Specifically, EE refers to the incentives that a utility firm provides to its customers so that the customers reduce
their total electricity demand by using more efficient devices. For example, Duke Energy, through its Appliance Recycling Program, offers a rebate to those who want to replace their old refrigerators with more efficient ones (Duke Energy 2015a). DR, on the other hand, aims to reduce the demand only during peak demand periods. For instance, MidWest Energy compensates farmers that curtail usage of water pumps upon a service call during high demand hours (Midwest Energy 2015).

From the perspective of a utility firm, EE is equivalent to the inflexible source and DR is equivalent to the flexible source. Specifically, the rebate paid to the customers under EE corresponds to the investment cost of the inflexible source and the curtailment payments made under the DR correspond to the generation cost of the flexible source. Furthermore, EE is used to reduce the baseload demand similar to nuclear energy, whereas DR is used to reduce demand at high demand periods similar to the natural gas. In terms of cost structures, the rebates given under EE are very costly similar to the high investment cost of the inflexible source. In contrast, DR involves a low initial cost but high curtailment payments similar to the high generation cost of the flexible source. Finally, similar to the flexible source, DR contracts are capable of curtailing demand within seconds because they involve automated response (DOE 2011). Thus, our model indicates that EE and renewable energy investment are substitutes, whereas DR and renewable investment are complements. As the share of renewables increases, the need for and the importance of DR will increase as well.

9 Conclusion

In this paper, we consider capacity investments of a utility firm in renewable and conventional sources with different levels of operational flexibility. We characterize the optimal investment levels and determine the role of operational flexibility in identifying the interaction (i.e., complement versus substitute) between energy sources. Specifically, a renewable and conventional source are substitutes (complements, respectively) if the conventional source is inflexible (flexible, respectively). We validate this result by using real electricity generation and demand data from Texas.

This paper has significant policy implications and it can provide guidelines for designing policies to promote renewables. First, we show that, from the perspective of a utility firm, the intermittency problem can be best alleviated by flexible energy sources, such as natural gas-fired power plants.
Thus, to promote renewables, policymakers should keep natural gas prices low, for example, by investing in pipeline infrastructure and issuing more permits for drilling. Second, policymakers should refrain from providing a subsidy for an inflexible source (e.g., nuclear or coal power) because this subsidy leads to lower investment in renewables. Third, a carbon tax is only effective in increasing renewable investment if the inflexible source is carbon-intensive such as coal power. Thus, given the high share of nuclear energy as the inflexible source in the U.S., the tax might not lead to increased renewable investment. Finally, policymakers should introduce electricity spot markets that enable utility firms to procure electricity at low cost when their own capacity is not sufficient to meet the demand due to intermittency. Such markets can reduce the adverse effects of intermittency from the utility firm’s perspective, thereby promoting investment in renewables.

Acknowledgments

The authors thank Aaron Beach, Sean Ong, and Stuart Cohen of National Renewable Energy Laboratory (NREL) for their contributions to the case study.

References

Aflaki, Sam, Serguei Netessine. 2015. Strategic investment in renewable energy sources: the effect of supply intermittency. Working paper, INSEAD.


EIA. 2013a. Further sensitivity analysis of hypothetical policies to limit energy-related carbon dioxide emissions. Energy Information Administration (EIA).


Kelly-Detwiler, Peter. 2014. Existing nuclear power plants may be renewable energy’s best friend. Forbes.


Monitoring Analytics. 2015. State of the market report for PJM.


Telson, Michael L. 1975. The economics of alternative levels of reliability for electric power generation systems. *Bell J. Econom.* 6(2) 679–694.


Walls, Margaret. 2015. Why now is the right time for a carbon tax. The Wall Street Journal.


Online Appendix

A Details of the Case Study

To estimate the investment cost parameters, we use Transparent Cost Database (TCDB, available at http://en.openei.org/apps/TCDB/), which tracks various publications that estimate cost figures for power plants. The estimate for the overnight capital cost of wind energy varies significantly throughout years. The median estimate is $1.57M per MW whereas the most recent one is $1.73M per MW. Based on this data, we set $\alpha_R = $1.65M per MW. For nuclear source, there is not enough data for recent years in the TCDB. Thus, we rely on a report of the Energy Information Administration (EIA 2013b) and observe that the overnight capital cost is $5.53M per MW. Furthermore, we assume that the economic life of a nuclear power plant is twice that of a wind plant. Thus, we set $\beta_I = $5.53/2 ≈ $3M per MW. Finally, for the open cycle gas turbine, we set $\alpha_F = $50,000 per MW. This is because, in the TCDB, cost estimates are as low as $200,000 per MW and the open cycle gas turbines also have a much longer economic life than wind plants.

Table 3: Optimal Investment Levels in the Case Study

<table>
<thead>
<tr>
<th>Optimal Investment Level (MW)</th>
<th>Original Costs</th>
<th>Nuclear Subsidy</th>
<th>Wind Subsidy</th>
<th>Natural Gas Subsidy</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k_I^*$</td>
<td>3,000</td>
<td>5,000</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$k_R^*$</td>
<td>5,000</td>
<td>0</td>
<td>20,000</td>
<td>10,000</td>
</tr>
<tr>
<td>$k_F^*$</td>
<td>1,000</td>
<td>0</td>
<td>3,000</td>
<td>5,000</td>
</tr>
</tbody>
</table>

Table 3 reports the optimal investment levels under original cost estimates and subsidies.

B Proofs

We first present the proofs of the Lemma 1 and 2.

Proof of Lemma 1. Let $q_I^* < k_I^*$ be the optimal solutions in (2). Let $\bar{k}_I = k_I^* - \epsilon$ for some $\epsilon > 0$ such that $q_I^* < \bar{k}_I$. Note that $\bar{k}_I$ is a feasible solution with a strictly lower cost, contradicting the optimality of $k_I^*$.

Proof of Lemma 2. Considering (3), we observe that this function decreases in $q_R$ at rate $r$, and it decreases in $q_F$ at rate $r - c_F$. Hence, it is optimal to dispatch the renewable source first, followed by
the flexible source, i.e., \( q^*_R(k, \xi, \theta) = \min(\theta k_R, \xi - k_I)^+ \) and \( q^*_F(k, \xi, \theta) = \min(k_F, \xi - k_I - \theta k_R)^+ \).

Before proceeding to the proofs of the propositions, we introduce the following lemma to be used later on.

**Lemma 4.** Consider a log-concave function \( f(\cdot) \). Let \( x, y, \) and \( z \) be positive scalars. Then, the following holds

\[
\frac{f(x+y)}{f(x)} \geq \frac{f(x+y+z)}{f(x+z)}. \tag{B.1}
\]

**Proof of Lemma 4.** This inequality holds if and only if \( \log f(x+y) + \log f(x+z) \geq \log f(x+y+z) + \log f(x) \). Let \( \lambda = \frac{y}{y+z} \), by the definition of log-concavity:

\[
\log f(\lambda (x+y+z) + (1-\lambda) x = x+y) \geq \lambda \log f(x+y+z) + (1-\lambda) \log f(x)
\]

\[
\log f((1-\lambda)(x+y+z) + \lambda x = x+z) \geq (1-\lambda) \log f(x+y+z) + \lambda \log f(x).
\]

Adding these inequalities side by side, we observe that \( \log f(x+y) + \log f(x+z) \geq \log f(x+y+z) + \log f(x) \). Hence, the inequality given in (B.1) holds. \( \square \)

**Proof of Proposition 1.** This proof follows similar arguments as in Van Mieghem (1998) and Sting and Huchzermeier (2012). We first note that the dispatch problem given in (3)-(5) can be equivalently expressed as:

\[
C(k, \xi, \theta) = \min_{q_R, q_F, s \geq 0} c_F q_F + rs
\]

subject to

\[
\frac{q_R}{\theta} \leq k_R \leftarrow \lambda_R
\]

\[
q_F \leq k_F \leftarrow \lambda_F
\]

\[
-s - q_R - q_F \leq k_I - \xi \leftarrow \lambda_I,
\]

where \( \lambda_i \)'s are the decision variables of the following dual problem:

\[
\max_{\lambda \in \mathbb{R}^+_I} (\xi - k_I) \lambda_I - k_R \lambda_R - k_F \lambda_F
\]

subject to

\[
\lambda_I \leq \frac{\lambda_R}{\theta}
\]

\[
\lambda_I \leq \lambda_F + c_F
\]

\[
\lambda_I \leq r.
\]

Since \( C(\cdot, \xi, \theta) \) is the minimal solution of a linear program, it is convex. Thus, \( \Pi(\cdot) \) is also convex.
because convexity is preserved under expectation and summation. Therefore, KKT conditions (7) and (8) are sufficient and necessary to identify the minimizer of $\bar{\Pi}(\cdot)$.

We next show that $E[C(k, \xi, \theta)] = -E[\lambda(k, \xi, \theta)]$. First, we note that the primal problem is always finite as the demand and intermittency distributions are bounded. Hence, the dual problem and the primal has the same objective value when they are both optimal. Let $\lambda^*(k, \xi, \theta)$ be the optimal dual solution for given $k, \xi, \theta$ and fix some $k^0 \in \mathbb{R}_+^3$. Then, for any $k \in \mathbb{R}_+^3$: $C(k, \xi, \theta) \geq \xi \lambda^*_k(k^0, \xi, \theta) - k^0 \lambda^*(k^0, \xi, \theta)$. Combining this with $C(k^0, \xi, \theta) = \xi \lambda^*_k(k^0, \xi, \theta) - k^0 \lambda^*(k^0, \xi, \theta)$, we obtain that

$$-C(k, \xi, \theta) \leq -C(k^0, \xi, \theta) + (k - k^0) \lambda^*(k^0, \xi, \theta).$$

Taking expectations of both sides, we observe that $E[\lambda(k, \epsilon, \Theta)]$ is a subgradient of $E[-C(k, \epsilon, \Theta)]$ evaluated at $k^0$. Because $C(k, \epsilon, \Theta)$ is convex, it is differentiable almost everywhere (a.e.) except for a set whose Lebesgue measure is zero as demand and intermittency are continuous. Thus, $\nabla_k C(k, \epsilon, \Theta)$ is single-valued a.e. This implies that the subgradient is unique for any $k \in \mathbb{R}_+^3$ and the KKT conditions given in (7) and (8) jointly define the optimal solution.  

**Proof of Proposition 2.** Note that in the triple sourcing case (i.e., $k > 0$), by expressing the demand and intermittency partitions explicitly, the FOCs with respect to (wrt) $k_I, k_R$, and $k_F$ can be sequentially written as:

$$F(k) = -\sum_{n=1}^N \int_{\theta=0}^1 \left[ r - (r - c_F) F_{\epsilon_n|\Theta_n}(k_I + \theta k_R + k_F|\theta) - c_F F_{\epsilon_n|\Theta_n}(k_I + \theta k_R|\theta) \right] f_{\Theta_n}(\theta) d\theta + (\alpha_I + Nc_I)$$

$$G(k) = -\sum_{n=1}^N \int_{\theta=0}^1 \theta \left[ r - (r - c_F) F_{\epsilon_n|\Theta_n}(k_I + \theta k_R + k_F|\theta) - c_F F_{\epsilon_n|\Theta_n}(k_I + \theta k_R|\theta) \right] f_{\Theta_n}(\theta) d\theta + \alpha_R$$

$$H(k) = -\sum_{n=1}^N \int_{\theta=0}^1 (r - c_F) \left(1 - F_{\epsilon_n|\Theta_n}(k_I + \theta k_R + k_F|\Theta_n)\right) f_{\Theta_n}(\theta) d\theta + \alpha_F.$$ 

Here, $F_{\epsilon_n|\Theta_n}(\cdot|\cdot)$ denotes the (conditional) cumulative distribution function (cdf). In addition, we define two random variables for later use as $X_n = (r - c_F) f_{\epsilon_n|\Theta_n}(k_I + \Theta_n k_R + k_F|\Theta_n)$ and $Y_n = c_F f_{\epsilon_n|\Theta_n}(k_I + \Theta_n k_R|\Theta_n)$. (i) We first show that $\frac{dX_n^*}{d\alpha_I} \leq 0$. Using implicit differentiation and
Cramer’s rule, it can be shown that
\[
\frac{dk_I^*}{d\alpha_I} = \frac{\begin{vmatrix}
\frac{\partial F}{\partial \alpha_I} & \frac{\partial F}{\partial k_R} & \frac{\partial F}{\partial k_F} \\
\frac{\partial G}{\partial \alpha_I} & \frac{\partial G}{\partial k_R} & \frac{\partial G}{\partial k_F} \\
\frac{\partial H}{\partial \alpha_I} & \frac{\partial H}{\partial k_R} & \frac{\partial H}{\partial k_F}
\end{vmatrix}}{H},
\]
where we drop the arguments of the FOCs for notational simplicity and \( H \) is the determinant of the Hessian matrix which is positive. Thus, to show that \( \frac{dk_I^*}{d\alpha_I} \leq 0 \), it suffices to show that the numerator is negative. The numerator is given as
\[
- E \left[ \sum_{n=1}^{N} \Theta_n^2 (X_n + Y_n) \right] E \left[ \sum_{n=1}^{N} X_n \right] + E \left[ \sum_{n=1}^{N} \Theta_n X_n \right] E \left[ \sum_{n=1}^{N} \Theta_n X_n \right],
\]
where \( X_n \) and \( Y_n \) are defined earlier. By Cauchy-Schwarz inequality, this is negative. Next, we consider \( \frac{dk_2^*}{d\alpha_I} \) and following similar steps as above, one can show that this derivative is equal to \( - E \left[ \sum_{n=1}^{N} X_n \right] E \left[ \sum_{n=1}^{N} \Theta_n Y_n \right] / H \), and hence negative. Finally, \( \frac{dk_F^*}{d\alpha_F} = - E \left[ \sum_{n=1}^{N} (X_n + Y_n) \right] E \left[ \sum_{n=1}^{N} \Theta_n^2 (X_n + Y_n) \right] / H + E \left[ \sum_{n=1}^{N} \Theta_n (X_n + Y_n) \right] E \left[ \sum_{n=1}^{N} \Theta_n (X_n + Y_n) \right] / H \), which is also negative. (ii) Using a similar technique, it can be shown that \( \frac{dk_I^*}{d\alpha_I} = \frac{dk_I^*}{d\alpha_I} \) for \( i, j \in \{ I, R, F \} \).

**Proof of Proposition 3.** We evaluate these derivatives by using implicit differentiation. (i) One can show that \( \frac{dk_I^*}{d\alpha_I} = \frac{dk_R^*}{d\alpha_I} = E \left[ \sum_{n=1}^{N} X_n \right] E \left[ \sum_{n=1}^{N} \Theta_n Y_n \right] / H \geq 0 \). Hence, the inflexible source and the renewable source are substitutes. (ii) It can be shown that \( \frac{dk_2^*}{d\alpha_I} = \frac{dk_F^*}{d\alpha_F} \geq Z c_F (r - c_F) / H \), where
\[
Z = E \left[ \sum_{n=1}^{N} X_n \right] E \left[ \sum_{n=1}^{N} \Theta_n^2 Y_n \right] - E \left[ \sum_{n=1}^{N} \Theta_n X_n \right] E \left[ \sum_{n=1}^{N} \Theta_n Y_n \right] / \left[ c_F (r - c_F) \right].
\]
Thus, to show that \( \frac{dk_F^*}{d\alpha_I} \geq 0 \), it suffices to show that \( Z \geq 0 \). Using the definitions of \( X_n \) and \( Y_n \):
\[
Z = \sum_{n=1}^{N} \int_{\theta=0}^{1} f_{\epsilon_n} (k_I + \theta k_R + k_F) f_{\Theta_n} (\theta) d\theta \times \sum_{n=1}^{N} \int_{\theta=0}^{1} \theta^2 f_{\epsilon_n} (k_I + \theta k_R) f_{\Theta_n} (\theta) d\theta
\]
\[
- \sum_{n=1}^{N} \int_{\theta=0}^{1} \theta f_{\epsilon_n} (k_I + \theta k_R + k_F) f_{\Theta_n} (\theta) d\theta \times \sum_{n=1}^{N} \int_{\theta=0}^{1} \theta f_{\epsilon_n} (k_I + \theta k_R) f_{\Theta_n} (\theta) d\theta
\]
\[
= \sum_{n=1}^{N} \sum_{m=1}^{N} \int_{\theta=0}^{1} \int_{\zeta=0}^{1} \theta (\theta - \zeta) f_{\epsilon_n} (k_I + \theta k_R) f_{\epsilon_m} (k_I + \zeta k_R + k_F) f_{\Theta_n} (\theta) f_{\Theta_m} (\zeta) d\zeta d\theta.
\]
This integral can also be expressed as
\[
\int_{\theta=0}^{1} \int_{\zeta=0}^{1} \theta (\theta - \zeta) f_{\epsilon_n} (k_I + \theta k_R) f_{\epsilon_m} (k_I + \zeta k_R + k_F) f_{\Theta_n} (\theta) f_{\Theta_m} (\zeta) d\zeta d\theta + \int_{\theta=0}^{1} \int_{\zeta=0}^{1} \theta (\theta - \zeta) f_{\epsilon_n} (k_I + \theta k_R) f_{\epsilon_m} (k_I + \zeta k_R + k_F) f_{\Theta_n} (\theta) f_{\Theta_m} (\zeta) d\zeta d\theta.
\]
The second integral is equivalent to
\[
- \int_{\zeta=0}^{1} \int_{\theta=0}^{1} (\zeta - \theta) f_{\epsilon_n} (k_I + \theta k_R) f_{\epsilon_m} (k_I + \zeta k_R + k_F) f_{\Theta_n} (\theta) f_{\Theta_m} (\zeta) d\theta d\zeta.
\]
In this expres-
sion, renaming $\theta$ as $\zeta$, $m$ as $n$ and vice-versa, we observe that

$$Z = \sum_{n=1}^{N} \sum_{m=1}^{N} \int_{\theta=0}^{\theta} \left[ \int_{\zeta=0}^{\theta} (\theta - \zeta) (f_{\epsilon_n}(kI + \theta kR) f_{\epsilon_m}(kI + \zeta kR + kF) - \zeta f_{\epsilon_m}(kI + \zeta kR) f_{\epsilon_n}(kI + \theta kR + kF)) \right] \frac{f_{\epsilon_n}(\theta)}{f_{\epsilon_m}(\zeta)} d\zeta d\theta.$$

Thus, $Z \geq 0$ if

$$\frac{\theta f_{\epsilon_n}(kI + \theta kR)}{\zeta f_{\epsilon_m}(kI + \zeta kR)} \geq \frac{f_{\epsilon_n}(kI + \theta kR + kF)}{f_{\epsilon_m}(kI + \zeta kR + kF)},$$

for $\theta > \zeta$ and $1 \leq n, m \leq N$. This condition is satisfied if $\Theta_n$ follows a two-point distribution. Furthermore, if $\epsilon_n = \epsilon$ for all $n$, i.e., $\epsilon_n$ is stationary as in Proposition 4, above inequality holds for any intermittency distribution as long as $\epsilon$ is log-concave by Lemma 4. (iii) Finally, $\frac{dk_F}{d\alpha} = \frac{dk_R}{d\alpha} = -TCF (r - c_F)/H$, where

$$T = \left( \frac{E \left[ \sum_{n=1}^{N} X_n \right] E \left[ \sum_{n=1}^{N} \Theta_n Y_n \right] - E \left[ \sum_{n=1}^{N} \Theta_n X_n \right] E \left[ \sum_{n=1}^{N} Y_n \right] }{c_F (r - c_F)} \right), \quad (B.2)$$

Thus, it is sufficient to show that $T \geq 0$ to prove that the flexible source and the renewable source are complements. Using a similar technique as above, one can show that

$$T = \sum_{n=1}^{N} \sum_{m=1}^{N} \int_{\theta=0}^{\theta} \left[ \int_{\zeta=0}^{\theta} (\theta - \zeta) (f_{\epsilon_n}(kI + \theta kR) f_{\epsilon_m}(kI + \zeta kR + kF) - \zeta f_{\epsilon_m}(kI + \zeta kR) f_{\epsilon_n}(kI + \theta kR + kF)) \right] \frac{f_{\epsilon_n}(\theta)}{f_{\epsilon_m}(\zeta)} d\zeta d\theta.$$

Thus, $T \geq 0$ if

$$\frac{f_{\epsilon_n}(kI + \theta kR)}{f_{\epsilon_m}(kI + \zeta kR)} \geq \frac{f_{\epsilon_n}(kI + \theta kR + kF)}{f_{\epsilon_m}(kI + \zeta kR + kF)},$$

for $\theta > \zeta$ and $1 \leq n, m \leq N$. This condition is satisfied if $\epsilon_n = \epsilon$ for all $n$, and the density function of $\epsilon$ is log-concave as in Proposition 4. Next, consider that $\Theta_n$ follows a two point distribution as in Assumption 1 and $q_n = q$ for all $n$. Then,

$$T = \sum_{n=1}^{N} \sum_{m=1}^{N} q (1 - q) (f_{\epsilon_n}(kI + kR) f_{\epsilon_m}(kI + kF) - f_{\epsilon_m}(kI) f_{\epsilon_n}(kI + R + kF)),$$

which is positive if

$$\sum_{n=1}^{N} f_{\epsilon_n}(kI + kF) \geq g(kI + kF) \frac{g(kI + kF)}{g(kI + kR)} = \sum_{n=1}^{N} f_{\epsilon_n}(kI + kR + kF),$$

i.e., as long as $g(\cdot)$ is log-concave by Lemma 4.

**Proof of Proposition 4.** See the Proof of Proposition 3.

**Proof of Proposition 5.** (i) Since $c_I$ is charged to the entire capacity $k_I$, the effect of $c_I$
on investment levels is the same as that of $\alpha_l$. (ii) Consider Assumption 1. It can be shown that
\[
\frac{dk_i}{dc_F} \times H = A_1A_2 \left[N - \sum_n q_n F_{3,n} - \sum_n (1 - q_n) F_{4,n}\right] + A_1B_1 \left[\sum_n (1 - q_n) F_{2,n} - \sum_n (1 - q_n) F_{4,n}\right] + A_2B_1 \left[N - \sum_n (1 - q_n) F_{4,n} - \sum_n q_n F_{1,n}\right],
\]
where $H$ is the determinant of the Hessian matrix, $A_1 = \sum_n q_n (r - c_F) f_{\epsilon_n} (k_I + k_R + k_T), A_2 = \sum_n (1 - q_n) (r - c_F) f_{\epsilon_n} (k_I + k_F), B_1 = \sum_n q_n c_F f_{\epsilon_n} (k_I + k_R), B_2 = \sum_n (1 - q_n) c_F f_{\epsilon_n} (k_I), F_{1,n} = F_{\epsilon_n} (k_I + k_R + k_F), F_{2,n} = F_{\epsilon_n} (k_I + k_F), F_{3,n} = F_{\epsilon_n} (k_I + k_R + k_F)$ and $F_{4,n} = F_{\epsilon_n} (k_I)$. Hence, $\frac{dk_i}{dc_F} \geq 0$, indicating that a generation subsidy for the flexible source leads to a lower investment in the flexible source. Furthermore, $\frac{dk_i}{dc_F} \times H = -A_1A_2 \left[N - \sum_n q_n F_{3,n} - \sum_n (1 - q_n) F_{2,n}\right] - A_1B_1 \left[N - \sum_n q_n F_{1,n} - \sum_n (1 - q_n) F_{4,n}\right] - B_1B_2 \left[N - \sum_n q_n F_{1,n} - \sum_n (1 - q_n) F_{2,n}\right]$, which is negative. Thus, a generation subsidy for the flexible source leads to a higher investment in the flexible source. \hfill \Box

**Proof of Proposition 6.** Defining $\bar{c}_i = c_i + t c_i$ for $i \in \{I, R, F\}$, one can show that this problem is jointly convex in its arguments. Furthermore, suppose $c_I = 0$. Then, $\frac{\partial F(k_I, k_R, k_F)}{\partial t} = c_F \frac{\partial F(k_I, k_R, k_F)}{dc_F}$, where $F(k_I, k_R, k_F)$ is the FOC wrt $k_I$. Similar relationships also hold for the FOCs wrt $k_R$ and $k_F$. Thus, through implicit differentiation it can be shown that $\frac{dk_i}{dc_F}$ has the same sign as $\frac{dk_i}{dt}$. \hfill \Box

**Proof of Proposition 7.** (i) We first consider $\frac{dk_i}{dc_F}$ for a general intermittency distribution. By using implicit differentiation, it can be shown that $\frac{dk_i}{dc_F} = E \left[\sum_{n=1}^N \Theta_n X_n \right] E \left[\sum_{n=1}^N Z_n \right] - E \left[\sum_{n=1}^N X_n \right] E \left[\sum_{n=1}^N \Theta_n Z_n \right] / (r - c_F)$ and $Z_n = \bar{F}_{\epsilon_n} (k_I + \Theta_n k_R + k_F)$. Here, $\bar{F}(\cdot)$ denotes the complementary cdf. By using a similar strategy as in the proof of Proposition 3, one can show that
\[
R = \sum_{n=1}^N \sum_{m=1}^N \int_0^1 \int_0^\theta (\theta - \zeta) \left[f_{\epsilon_n}(k_I + \theta k_R + k_F) \bar{F}_{\epsilon_n}(k_I + \zeta k_R + k_F) -
\right]
\]
\[
f_{\epsilon_n}(k_I + \zeta k_R + k_F) \bar{F}_{\epsilon_n}(k_I + \theta k_R + k_F) \right] d\Theta_m(\theta) d\Theta_m(\theta) d\zeta d\theta.
\]
Notice that $R \geq 0$ if
\[
\frac{f_{\epsilon_n}(k_I + \theta k_R + k_F)}{\bar{F}_{\epsilon_n}(k_I + \theta k_R + k_F)} \geq \frac{f_{\epsilon_n}(k_I + \zeta k_R + k_F)}{\bar{F}_{\epsilon_n}(k_I + \zeta k_R + k_F)}
\]
for $\theta > \zeta$ and $1 < n, m < N$. If demand distribution is stationary, i.e., $\epsilon_n = \epsilon$ for all $n$, this condition is satisfied as long as $\epsilon$ is log-concave, which implies that the failure rate is increasing (see Bagnoli and Bergstrom 2005, Corollary 2). Furthermore, consider that the intermittency follows a two-point distribution as given in (9) and $q_n = q$ for all $n$, then $R = \sum_{n=1}^N \sum_{m=1}^N q (1 - q) \left[f_{\epsilon_n}(k_I + k_R + k_F) \bar{F}_{\epsilon_n}(k_I + k_F) - f_{\epsilon_n}(k_I + k_F) \bar{F}_{\epsilon_n}(k_I + k_R + k_F) \right]$. In this
case, $R \geq 0$ if
\[
\sum_{n=1}^{N} f_{e_n}(k_I + k_R + k_F) = g(k_I + k_R + k_F) \geq \frac{g(k_I + k_F)}{G(k_I + k_F)} = \frac{\sum_{n=1}^{N} f_{e_n}(k_I + k_F)}{\sum_{n=1}^{N} \tilde{F}_e(x, k_I + k_F)},
\]
where we define $G(x) = \sum_{n=1}^{N} \tilde{F}_e(x)$. We next show that this inequality holds if $g(x)$ is log-concave. As long as $g(x)$ is log-concave, $G(x)$ is also log-concave. This can be seen from Bagnoli and Bergstrom 2005, Lemma 4 because $dG(x)/dx = -g(x)$. Due to log-concavity of $G(x)$ and $g(x)$, the ratio $g(x)/G(x)$ (hazard rate) is increasing as shown in Bagnoli and Bergstrom 2005, Corollary 2. Hence, the inequality above holds if $g(x)$ is a log-concave function. This proves that $R \geq 0$, implying that $\frac{dk^*_F}{dr} \geq 0$. We next consider the flexible source and note that
\[
\frac{dk^*_F}{dr} \geq \left(E \left[ \sum_{n=1}^{N} Z_n \right] U + E \left[ \sum_{n=1}^{N} \Theta_n Z_n \right] T \right) c_F \left( r - c_F \right) / H,
\]
\[
U = \left( E \left[ \sum_{n=1}^{N} (1 - q_m) (r - c_F) f_{e_n}(k_I + k_R + k_F) f_{e_m}(k_I) \right] \right) / r - c_F, \text{ and } T \text{ is defined in (A.3) and shown to be positive if } \Theta_n \\text{ follows a stationary two-point distribution and } g(\cdot) \\text{ is log-concave. Hence, it sufficies to show that } U \geq 0, \text{ to prove that } \frac{dk^*_F}{dr} \geq 0.
\]
(iii) We finally consider $\frac{dk^*_F}{dr} = -E \left[ \sum_{n=1}^{N} Y_n \right] R(r - c_F) / H$. We already show the conditions that ensure $R \geq 0$. Hence, whenever $\frac{dk^*_F}{dr}$ is positive, $\frac{dk^*_F}{dr}$ is negative.

**Proof of Lemma 3.** (i) By plugging $q_S$ into the objective function in (14), we define $\bar{C}_n(k, \xi, \theta, \gamma) = c_F q_F + (\gamma + b_F^2 (\xi - k_I - q_R - q_F)) (\xi - k_I - q_R - q_F)$, where its derivative wrt $q_R$ and $q_F$ is given as $-\gamma + b_n (k_I + q_R + q_F - \xi)$ and $c_F - \gamma + b_n (k_I + q_R + q_F - \xi)$, respectively. To determine optimal dispatch policy, we consider five cases. Case 1: $\gamma \leq b_n (k_I - \xi)$, $\frac{\partial \bar{C}_n}{\partial q_R} \geq 0$ and $\frac{\partial \bar{C}_n}{\partial q_F} \geq 0$, hence $q_R = q_F^* = 0$. Case 2: $b_n (k_I - \xi) \leq \gamma \leq b_n (k_I + \theta k_R - \xi)$, $q_R^* = \xi - k_I + \frac{\gamma}{b_n}$ so that $\frac{\partial \bar{C}_n}{\partial q_R} = 0$, and $q_F^* = 0$ as $\frac{\partial \bar{C}_n}{\partial q_F} \geq 0$. Case 3: $b_n (k_I + \theta k_R - \xi) \leq \gamma \leq c_F + b_n (k_I + \theta k_R - \xi)$, $q_R^* = \theta k_R$ so that $\frac{\partial \bar{C}_n}{\partial q_R}$ can be as close to 0 as possible subject to the constraint (15) and $q_F^* = 0$ as $\frac{\partial \bar{C}_n}{\partial q_F}$ is still positive. Case 4: $c_F + b_n (k_I + \theta k_R - \xi) \leq \gamma \leq c_F + b_n (k_I + \theta k_R + k_F - \xi)$, similar to the previous case $q_R^* = \theta k_R$ but in this case $q_F^* = \xi - k_I - k_R + \frac{\gamma}{b_n}$. Case 5: $c_F + b_n (k_I + \theta k_R + k_F - \xi) \leq \gamma$, $q_R^*$ remains to be $\theta k_R$ and $q_F^* = k_F$ so that $\frac{\partial \bar{C}_n}{\partial q_F}$ can be as close to 0 as possible subject to the constraint (16). This optimal dispatch policy can be defined as $q_R^* (k, \xi, \theta, \gamma) = \min \left( \theta k_R, \xi - k_I + \frac{\gamma}{b_n} \right)$ and $q_F^* (k, \xi, \theta, \gamma) = \min \left( k_F, \xi - k_I - \theta k_R + \frac{\gamma}{b_n} \right)$. (ii) With this optimal dispatch policy, the Hessian of the first stage problem can be shown to be positive so that the problem is jointly convex in its arguments.
Proof of Proposition 8. We first define two random variables:

\[ X'_n = b_n E \left[ F_{\Gamma_n|\epsilon_n,\Theta_n} (c_F + b_n (k_I + \Theta_n k_R + k_F - \epsilon_n)) | \epsilon_n, \Theta_n \right] \]

\[ Y'_n = b_n E \left[ F_{\Gamma_n|\epsilon_n,\Theta_n} (c_F + b_n (k_I + \Theta_n k_R - \epsilon_n)) | \epsilon_n, \Theta_n \right] - F_{\Gamma_n|\epsilon_n,\Theta_n} (b_n (k_I + \Theta_n k_R - \epsilon_n)) | \epsilon_n, \Theta_n \right]. \]

(i) As in the proof of Proposition 2, we use implicit differentiation to evaluate these derivatives.

\[ \frac{dk_I}{d\alpha_R} = \frac{dk_I}{d\alpha_I} = E \left[ \sum_{n=1}^{N} X'_n \right] E \left[ \sum_{n=1}^{N} \Theta_n Y'_n \right] / H', \] where \( H' \) is the determinant of the Hessian (which is positive). Thus, \( \frac{dk_I}{d\alpha_R} \geq 0 \), indicating that the inflexible and renewable sources are substitutes. (ii)

It can be shown that \( Z = \left[ \sum_{n=1}^{N} b_n (1 - q_n) \int_{0}^{\infty} \left[ \int_{\gamma=c_F+b_n(k_I+k_R-\xi)}^{\infty} f_{\Gamma_n} (\gamma) d\gamma \right] f_{\epsilon_n} (\xi) d\xi \right] \times \left[ \sum_{n=1}^{N} b_n q_n \int_{0}^{\infty} \left[ \int_{\gamma=c_F+b_n(k_I+k_R-\xi)}^{\infty} f_{\Gamma_n} (\gamma) d\gamma \right] f_{\epsilon_n} (\xi) d\xi \right] \) under the Bernoulli intermittency assumption. Hence, \( Z \) is positive so that the inflexible and renewable sources are substitutes as well. (iii)

Finally, considering Assumption 2, \( \frac{dk_I}{d\alpha_F} = \frac{dk_I}{d\alpha_R} = -T/H, \) where \( T = E \left[ \sum_{n=1}^{N} X'_n \right] E \left[ \sum_{n=1}^{N} \Theta_n Y'_n \right] - E \left[ \sum_{n=1}^{N} \Theta_n X'_n \right] E \left[ \sum_{n=1}^{N} Y'_n \right]. \) Under the Bernoulli intermittency distribution with \( q_n = q \) for all \( n, T = q (1 - q) \left[ \sum_{n=1}^{N} b_n c_F (1 - q_n) \right] \left[ \sum_{n=1}^{N} b_n \int_{0}^{\infty} \left[ \int_{\gamma=c_F+b_n(k_I+k_R-\xi)}^{\infty} f_{\Gamma_n} (\gamma) f_{\epsilon_n} (\xi) d\gamma \right] d\xi \right]. \) Since \( T \) is positive, \( \frac{dk_I}{d\alpha_F} \leq 0, \) hence the renewable and flexible sources are complements. \( \square \)

A8