

NBER WORKING PAPER SERIES

LEVEL VERSUS VARIABILITY TRADE-OFFS IN WIND AND SOLAR GENERATION INVESTMENTS:
THE CASE OF CALIFORNIA

Frank A. Wolak

Working Paper 22494
<http://www.nber.org/papers/w22494>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 2016

I would like to thank Reinhard Madlener and Carlo Andrea Bollino and two anonymous referees for helpful comments on a previous draft. Chris Bruegge provided outstanding research assistance. The views expressed herein are those of the author and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2016 by Frank A. Wolak. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Level versus Variability Trade-offs in Wind and Solar Generation Investments: The Case of California

Frank A. Wolak

NBER Working Paper No. 22494

August 2016

JEL No. Q2,Q4,Q5

ABSTRACT

Hourly plant-level wind and solar generation output and real-time price data for one year from the California ISO control area is used to estimate the vector of means and the contemporaneous covariance matrix of hourly output and revenues across all wind and solar locations in the state. Annual hourly output and annual hourly revenues mean/standard deviation efficient frontiers for wind and solar resource locations are computed from this information. For both efficient frontiers, economically meaningful differences between portfolios on the efficient frontier and the actual wind and solar generation capacity mix are found. The relative difference is significantly larger for aggregate hourly output relative to aggregate hourly revenues, consistent with expected profit-maximizing unilateral entry decisions by renewable resource owners. Most of the hourly output and hourly revenue risk-reducing benefits from the optimal choice of locational generation capacities is captured by a small number of wind resource locations, with the addition of a small number of solar resource locations only slightly increasing the set of feasible portfolio mean and standard deviation combinations. Measures of non-diversifiable wind and solar energy and revenue risk are computed using the actual market portfolio and the risk-adjusted expected hourly output or hourly revenue maximizing portfolios.

Frank A. Wolak

Department of Economics

Stanford University

Stanford, CA 94305-6072

and NBER

wolak@zia.stanford.edu

1. Introduction

This paper quantifies the total renewable energy output and revenue consequences of different combinations of wind and solar generation investments in California using hourly output and real-time market revenue data for all wind and solar generation units in the California Independent System Operator (ISO) control area from April 1, 2011 to March 31, 2012. This data is used to construct an economic model that provides an estimate of the average annual mean and covariance matrix of the hourly energy production and the hourly real-time revenues for all existing wind and solar locations in California for any amount of renewable generation capacity at each existing renewable location.

The economic model is used to compute the shares of statewide wind and solar capacity at existing wind and solar locations in California that yield the highest annual average hourly output for a given value of the annual variability of hourly output, or equivalently the lowest annual variability in hourly output for a given value of the average hourly output. These pairs of annual average hourly output and annual variability in hourly output define the mean output and standard deviation of output efficient frontier for wind and solar investments at existing wind and solar locations in California. Points on this efficient frontier are compared to the actual average hourly output and annual variability in hourly output from actual wind and solar generation units.

This same exercise is repeated for the annual average hourly revenue and annual variability in hourly revenue for wind and solar generation locations in California. Locational capacity shares at all existing wind and solar locations in California are computed that yield the highest annual average hourly revenues for a given value of the annual variability of hourly revenues, or equivalently the lowest annual variability in hourly revenues for a given value of the average hourly revenues. These annual average hourly revenues and annual variability in hourly revenues

pairs define the mean hourly revenue and standard deviation of hourly revenue efficient frontier of wind and solar investments at existing wind and solar locations in California. Points on this efficient frontier are compared to the actual hourly average revenue and annual variability in hourly revenue from actual wind and solar generation units.

Although the renewable generation locations are distributed throughout a state with a geographic area slightly smaller than France and slightly larger than Germany, these two analyses reveal a high degree of contemporaneous correlation between the hourly output of the 13 solar locations and 40 wind locations in California that produced energy throughout the 2011-2012 fiscal year. This result suggests modest opportunities to reduce the variability in the total hourly output and hourly revenue of wind and solar resource owners by optimizing where and how much solar and wind generation capacity is placed at each existing location in California. Nevertheless, the hourly output efficient frontier implies that a 48 percent increase in the annual average hourly output of solar and wind units is possible without increasing the annual standard deviation of hourly output only by changing the state-wide capacity shares of the wind and solar investments at the existing wind and solar locations in California. More modest increases in the average hourly revenue for solar and wind generation units can be obtained by optimizing the locational capacity shares of these investments. The hourly revenue efficient frontier implies that by changing the state-wide capacity shares of the wind and solar investments at the existing wind and solar locations, a 26 percent increase in the annual average hourly real-time revenue of the solar and wind generation capacity is possible without increasing the annual standard deviation in hourly revenues by changing the state-wide capacity shares of the wind and solar investments at the existing wind and solar locations. This 26 percent increase in the annual average hourly revenue

is found to be statistically significantly smaller than the 48 percentage increase in annual average hourly output.

To assess the extent to which the results described above are robust across hours of the day, separate efficient frontiers are computed for four groups of hours of the day. The largest gains in expected output or reductions in the standard deviation of output as a result of moving from the actual capacity shares of wind and solar investments to the efficient frontier appears to be during the late morning and afternoon hours of the day.

The capacity shares of wind and solar investments on the annual hourly output and annual hourly revenue efficient frontiers are similar, however they are both very different from the actual capacity shares. The capacity shares on both efficient frontiers concentrate the statewide capacity of wind and solar investments on a substantially smaller number of locations than actual wind and solar capacity investments.

Taken together, these empirical results have implications for the design of policies to stimulate the deployment of renewable generation capacity. As the share of annual energy production provided by intermittent generation resources in a region increases, the cost of managing the real-time supply and demand balance increases.¹ For the same annual average hourly output from wind and solar resources, a more variable hourly output from these generation units implies that more operating reserves are required to maintain grid reliability standards. Consequently, a substantial amount of the increased hourly renewable energy output variability associated with scaling wind and solar energy production in California could be mitigated by

¹ Many jurisdictions have significant renewable energy goals. California has a legislative mandate to achieve a 33% qualified renewable energy share by 2020 and a 50% share by 2030. Three-fifths of US states have similar mandated renewables portfolio standards.

constructing additional wind and solar capacity in resource areas that minimize the increase in the annual standard deviation of hourly output.

To demonstrate the empirical content of this statement, the elasticity of annual average hourly output with respect to the annual standard deviation of hourly output for a 1 MW increase in capacity at that location is computed for each existing wind and solar location in California starting from the current statewide locational capacity shares for wind and solar generation capacity. There is substantial heterogeneity in these elasticities across locations. For some locations the elasticity is as much as ten times larger than the typical California wind and solar location, indicating that at these locations a 1 MW increase in wind or solar capacity would have a significantly larger increase in the annual average wind and solar output for the same increase in the annual standard deviation in hourly output from these units. There are even some locations where the elasticity is negative, meaning that a 1 MW increase in wind or solar capacity at that location would increase the annual average hourly output and reduce the annual standard deviation of hourly output. Wind and solar investments at locations with the highest values of these elasticities are likely to minimize the system reliability costs and operating costs associated with achieving any statewide renewable energy goal.

The remainder of the paper proceeds as follows. The next section provides a discussion of related research. Section 3 presents a model of output and revenue risk diversification for wind and solar generation investments. Section 4 contains descriptive statistics on the California electricity market and discusses the data used to construct the efficient frontiers. Section 5 discusses the computation of the efficient frontiers and the procedure used to compute the maximum risk-adjusted expected output and the expected revenue points on these curves as well as location-specific measures of non-diversifiable wind and solar output and revenue risk. Section

6 presents the empirical results. Section 7 discusses the implications of these results and the design of potential policies to minimize the reliability and operating costs associated with meeting any renewable energy goal. Section 8 concludes.

2. Portfolio-Based Approaches to New Generation Investment Decisions

A number of papers have formulated electricity generation capacity technology mix decisions as portfolio choice problems. Bar-Lev and Katz (1976) first applied expected return and risk (standard deviation return) portfolio theory to the choice of the mix fossil fuel generation capacity—coal, natural gas, and oil—used to produce electricity. Awerbuch and Berger (2003) employed mean-variance portfolio theory to derive European Union (EU) generation portfolios that enhance energy securities objectives. They quantified the investment return risk diversification benefits of increasing the amount of renewable generation capacity in the EU. Roques, Newbery and Nuttal (2008) employ Monte Carlo simulations of natural gas, coal, and nuclear generation unit investment return distributions to solve a mean-variance portfolio choice problem for an optimal mix of these generation technologies for firms in a re-structured wholesale electricity market. Bazilian and Roques (2008) contains a number of papers that extend in a number of directions the application of expected return and risk portfolio theory to the electricity supply industry investment decision-making processes.

Westner and Madlener (2010) apply mean-variance portfolio theory to Monte Carlo simulations of distributions of the net present value (NPV) of individual combined heat and power technologies in the four largest Western European countries to study the potential for regional diversification benefits of investments in these technologies. Westner and Madlener (2011) consider investments in four combined heat and power technologies in Germany and employ Monte Carlo simulations of the joint distribution of the NPVs of each technology, and from this

compute the NPV mean-variance efficient frontier for the four technologies. Delarue, De Jongho, Belmans, and D'haeseleer (2011) apply mean-variance portfolio theory and Monte Carlo simulations to compute distributions of the cost per MWh by technology to derive the efficient frontier in cost per MWh and standard deviation-of-cost per MWh space of investments in nuclear, coal, gas, oil and wind generation units. Bhattacharya and Kojima (2012) assess the impact of adding renewable generation investments to the existing generation capacity mix on the portfolio-level return and risk profile for the Japanese electricity sector.

Several papers have also applied mean/variance portfolio theory across geographic locations. Yu (2003) considers the question of choosing efficient portfolios of energy and ancillary services sales by generation owners able to sell in spatially distributed wholesale electricity markets. Roques, Hiroux, and Sagan (2010) consider the optimal country-level deployment of wind generation units for France, Denmark, Austria, Germany and Spain, using a sample of country-level hourly wind output in these five countries. They derive a mean hourly output and standard deviation of hourly output efficient frontier for wind deployment across these five countries.

Madlener (2012) provides an accessible and comprehensive survey of this growing literature. Garnier and Madlener (2016) expands the set of risks that a potential investor faces by incorporating regulatory regime risk into a portfolio choice model for an investor in a virtual power plant composed of a mix of distributed renewable energy resources where the investor must manage both intraday energy balancing cost risk and policy regime risk.

The present paper builds on this previous work by applying the mean-variance portfolio theory to generation investment decisions in order to understand how substantial intermittent renewable energy goals can be met with minimal reliability consequences. To this end, using

actual hourly output data, I characterize the trade-offs between aggregate renewable energy output and intermittency risk for any portfolio of wind and solar resource locations in California. I then quantify the magnitude of the deviations of the actual mix of renewable generation capacity from the mix capacity on the hourly output and hourly standard deviation efficient frontier. I also explore the extent to which the location decisions of existing renewable generation capacity in California are consistent with maximizing the unit owner's expected profits. I accomplish this by comparing the actual portfolio of renewable generation investments to the portfolios on the expected revenue and standard deviation of renewable generation revenues efficient frontier. Finally, I compute location-specific measures of output and revenue risk and find that only a few actual renewable locations are necessary to achieve a frontier of the highest expected hourly capacity factor or highly hourly revenue per MW of capacity for a given value of these two location-specific risk measures.

3. Model of Renewable Output and Revenue Risk Diversification

The level versus variability tradeoffs in the system-wide hourly output and hourly total revenues for wind and solar resource locations studied in this paper can be conveyed through a simple example with N intermittent renewable resource areas. The contemporaneous correlation between the hourly output at each intermittent energy location and the impact of the level of hourly output on real-time prices at each intermittent renewable energy location determines the output and revenue volatility reductions that are possible from different intermittent renewable energy investment policies.

Suppose that during a typical hour of the day at each of these N locations 1 megawatt-hour (MWh) of intermittent renewable energy is available to be harvested with probability p . If the availability of 1 MWh energy at one location is distributed independently of the availability of 1

MWh energy at any other location, then the expected number of MWh produced in an hour across these N locations is pN MWh, in other words the probability of the event times the number of resource areas. The variance of the hourly output from the N intermittent renewable resources is equal to $Np(1-p)$. Now consider the opposite extreme in terms of correlation between the hourly outputs of the N locations. Suppose that there is perfect positive correlation between the hourly outputs at these same N locations. The expected hourly output of the facilities is still equal to pN , but the variance of the aggregate hourly output measure is equal to $p(1-p)N^2$ because of the perfect positive correlation between the hourly output across the N locations. The hourly variance is N times larger than it is for the case in which the hourly availability of energy is independently distributed across locations.

This example shows that for the same aggregate amount of renewable generation capacity in a region, the variance in the total hourly output of wind and solar generation facilities can be as much as N times larger as a result of correlation in the output across the N locations. Conversely, by accounting for the contemporaneous correlation between the hourly outputs at pairs of the N renewable energy locations, the variance in the total hourly output of renewable energy units can be reduced by installing different amounts of capacity at each of the N available locations.

To understand how this hourly renewable output variability reduction can be realized, consider the following wind and solar investment portfolio choice problem. Suppose that Q_{jh} is hourly output in megawatt-hours (MWh) at location j during hour h and K_j is the MW capacity at location j for $j=1,2,\dots,J$. Define $f_{jh} = Q_{jh}/K_j$ as the hourly capacity factor at location j , which is the ratio of the wind or solar energy actually produced during the hour at location j divided by the amount energy that could be produced during one hour at location j by fully utilizing the K_j MWs of capacity for an entire hour. Note that f_{jh} is continuously distributed on the interval $[0,1]$. Let

$K_T = \sum_{j=1}^J K_j$ equal the total amount of wind and solar generation capacity in the control area. Let $w_j^{act} = K_j/K_T$ equal the actual share of total wind and solar generation capacity at location j .

The total renewable energy produced during hour h is equal to $Q_h^T = \sum_{j=1}^J Q_{jh} = \sum_{j=1}^J w_j^{act} f_{jh} K_T$. The second equality illustrates that Q_h^T can be written as the capacity-share-weighted sum of the hourly capacity factors of the individual wind and solar resource locations multiplied by the total amount of wind and solar capacity in the control area. This result implies that the mean and standard deviation of Q_h^T depends on the J -dimensional mean and $(J \times J)$ covariance matrix of the vector of hourly location specific capacity factors $F_h = (f_{1h}, f_{2h}, \dots, f_{Jh})'$. Let the expected value of F_h , $E(F_h)$, equal the J -dimensional vector μ and the covariance matrix of F_h , $E[(F_h - \mu)(F_h - \mu)']$, equal the $(J \times J)$ symmetric, positive definite matrix, Ω .² In terms of this notation, the mean and standard deviation of Q_h^T can be expressed as:

$$E(Q_h^T) = \sum_{j=1}^J w_j^{act} \mu_j K_T \text{ and } SD(Q_h^T) = [\sum_{j=1}^J \sum_{k=1}^J w_j^{act} \omega_{jk} w_k^{act}]^{1/2} K_T. \quad (1)$$

where μ_j is the j^{th} element of μ and ω_{jk} is the $(j,k)^{th}$ element of Ω . This result shows that the expected value of Q_h^T is equal to the expected value of the capacity-share-weighted sum of the location-specific hourly capacity factors, $CF_h^T = \sum_{j=1}^J w_j^{act} f_{jh}$, multiplied by the total amount of renewable capacity, K_T . The standard deviation of Q_h^T is equal to the standard deviation of the capacity share weighted sum of the location-specific capacity factors multiplied by the total amount of renewable capacity, K_T . These decompositions imply the following two results:

$$E(Q_h^T) = E(CF_h^T) K_T \text{ and } SD(Q_h^T) = SD(CF_h^T) K_T, \quad (2)$$

which implies that for a fixed amount of renewable generation capacity, K_T , the capacity shares that yield the highest expected hourly output are those that maximize the expected value of CF_h^T ,

² Throughout the paper $E(X)$ denotes the expectation of a random variable X .

and the capacity shares that yield the lowest standard deviation of hourly output are those that minimize the standard deviation of CF_h^T .

To compute the hourly total revenues earned by wind and solar resources, let P_{jh} equal the real-time price at location j during hour h . The total revenues earned by a wind or solar resource at location j is equal to $TR_j = P_{jh} * f_{jh} * K_j$. The expected revenues earned by renewable resources at location j is equal to $E(TR_j) = [E(P_{jh}) * E(f_{jh}) + Cov(P_{jh}, f_{jh})] K_j$, where $Cov(X, Y)$ denotes the covariance between the random variables X and Y . This expression implies that there are several ways to increase expected renewable energy revenues at a location. A higher expected value of the hourly capacity factor, $E(f_{jh})$, and a higher expected price, $E(P_{jh})$, both increase the expected revenues of the renewable resources at location j . Note that a positive covariance between the hourly real-time price and hourly capacity factor at location j also increases expected revenues at that location. The variance of total revenues at location j depends on higher-order moments of the hourly capacity factor and hourly price at location j . Specifically,

$$Var(TR_j) = \{E[(P_{jh})^2]E[(f_{jh})^2] + Cov[(P_{jh})^2, (f_{jh})^2] - [E(P_{jh}) * E(f_{jh}) + Cov(P_{jh}, f_{jh})]^2\} (K_j)^2. \quad (3)$$

The total revenues earned by renewable resources during hour h , TR_h^T can be expressed as

$$TR_h^T = \sum_{j=1}^J P_{jh} w_j^{act} f_{jh} K_T. \quad (4)$$

Let $r_{jh} = P_{jh} * f_{jh}$ equal the revenue per MW of capacity paid to resources at location j during hour h . Let $R_h = (r_{1h}, r_{2h}, \dots, r_{Jh})'$ be the vector of locational revenues per MW of capacity during hour h . Let the expected value of R_h , $E(R_h)$, equals the J -dimensional vector λ and the covariance matrix of R_h , $E[(R_h - \lambda)(R_h - \lambda)']$, equals the $(J \times J)$ symmetric, positive definite matrix, Γ . In terms of this notation, the mean and standard deviation of TR_h^T can be expressed as:

$$E(TR_h^T) = \sum_{j=1}^J w_j^{act} \lambda_j K_T \text{ and } SD(TR_h^T) = [\sum_{j=1}^J \sum_{k=1}^J w_j^{act} \gamma_{jk} w_k^{act}]^{1/2} K_T, \quad (5)$$

where λ_j is the j^{th} element of λ and γ_{jk} is the $(j,k)^{\text{th}}$ element of Γ . This result shows that the expected value of TR_h^T is equal to the expected value of the capacity-share-weighted sum of the location-specific revenues per MW of capacity, $RC_h^T = \sum_{j=1}^J w_j^{act} r_{jh}$, multiplied by the total amount of renewable capacity, K_T in MWs. The standard deviation of TR_h^T is equal to the product of the standard deviation of the capacity-share-weighted sum of the location-specific revenues per MW of capacity multiplied by the total amount of renewable capacity, K_T . These decompositions imply the following two results:

$$E(TR_h^T) = E(RC_h^T)K_T \text{ and } SD(TR_h^T) = SD(RC_h^T)K_T, \quad (6)$$

which implies that for a fixed total amount of renewable capacity, K_T , the capacity shares that yield the highest expected hourly revenue are those that maximize the expected value of RC_h^T and the capacity shares that yield the lowest standard deviation of hourly revenue are those that minimize the standard deviation of RC_h^T .

4. Renewables Output and Revenues in California Electricity Market

This section describes the data used to estimate the parameters of the mean and covariance matrices of F_h , the vector of hourly location-specific capacity factors, and R_h , the vector of hourly location-specific revenues per MW of capacity. To motivate the efficient frontier modeling results presented in Section 6, this section first presents summary statistics on the California wholesale electricity market.

4.1. Descriptive Statistics for California Wholesale Electricity Market

The sample period is the California ISO fiscal year from April 1, 2011 to March 31, 2012. There were 40 wind generation locations and 13 utility-scale solar locations in the California ISO control area, with 3,040 MW of wind capacity and 499 MW of solar capacity, available to operate throughout this time period. Table 1 lists the amount of generation capacity in the California ISO

control area by technology as of April 1, 2011. The vast majority of generation capacity is natural gas-fired. There is also a significant amount of hydroelectric capacity and nuclear capacity.

[Insert Table 1 Here]

[Insert Figure 1 Here]

The left panel of Figure 1 graphs the mean within-day pattern of aggregate demand and the within-day pattern of the standard deviations of each hour-of-the-day demand level. If QD_{hd} is the aggregate demand for hour h of day d , then each point along the curve is equal to $\overline{QD}_h = \frac{1}{D} \sum_{d=1}^D QD_{hd}$, which is the mean of the demand during hour-of-the-day h over the $D = 365$ days in the sample. Each point along the hourly standard deviation curve is equal to $SD(QD)_h = [\frac{1}{D-1} \sum_{d=1}^D (QD_{hd} - \overline{QD}_h)^2]^{1/2}$, the standard deviation of Q_{hd} over all days in the sample period. The left panel of Figure 1 repeats these two graphs for the QWS_{dh} , the total output of wind and solar generation units during hour h of day d . The graph of $SD(QWS)_h$ is only slightly lower than the graph of \overline{QWS}_h , ranging between 70 and 60 percent of \overline{QWS}_h across all hours of the day. In contrast, the graph of $SD(QD)_h$ is substantially below the graph of \overline{QD}_h , ranging between 9 percent and 16 percent of \overline{QD}_h across all hours of the day. The volatility (as measured by the standard deviation) of hourly wind and solar production is a much larger fraction of the mean hourly wind and solar production than the volatility of hourly demand is relative to the mean of hourly demand.

The California ISO and all other formal wholesale markets in the United States operate a multi-settlement locational marginal pricing market which sets a potentially different price at each generation unit or load withdrawal location in the control area for every hour of the day. The price depends on the configuration transmission network, the location of electricity demand, the location and level of output of dispatchable generation units, and the level of wind and solar energy

production.³ High levels of close-to-zero marginal cost of wind and solar energy production are likely to displace the more costly fossil fuel energy production to the extent that the configuration of the transmission network and the location of demand allow it. Conversely, low levels of wind and solar energy production are likely to require significantly more production from fossil fuel generation units to the extent that the transmission network and location of demand allow it. If either of these outcomes is limited by transmission capacity or other operating constraints, this will result in different prices at different locations in the transmission network. Generation-rich locations will have lower locational marginal prices (LMPs) and generation-deficient locations will have higher LMPs, so that no more than the amount of energy equal to the capacity of each transmission link will flow between any two locations in the transmission network.

The other important feature of the California market and all United States electricity markets in general, is that they have at least one formal forward market prior to real-time in which participants can purchase and sell firm financial commitments to buy or sell electricity. For example, the California ISO operates a day-ahead market where participants submit offers to sell and bids to buy electricity in each of the 24 hours of the following day. The total amount of energy sold in the day-ahead market by a specific generation unit is typically referred to as the unit's day-ahead energy schedule.

This day-ahead energy schedule is a firm financial commitment that must be settled in the real-time market, which operates every five minutes in the California ISO control area. For example, a generation unit owner that sells 100 MWh of energy from a generation unit during the 10 am to 11 am period of the following day must either produce that amount of energy from its unit to cover this sale or purchase the 100 MWh from the real-time market at the real-time LMP

³ Jha and Wolak (2013) provide a description of the key features of the operation of a multi-settlement, locational marginal pricing wholesale electricity market.

at that location. By similar logic, a load that purchased 70 MWh in the day-ahead market for this same hour can sell some or all of this energy in the real-time market at the real-time LMP at that location by consuming less than 70 MWh during that hour. Alternatively, if the load consumes more than 70 MWh during the 10 am to 11 am hour, then it must purchase the additional energy from the real-time market at the real-time price for that location. Similar logic applies to a generation unit. If it produces more than the 100 MWh it sold in the day-ahead market, this additional energy is paid the real-time price at that location.

The above logic implies that the real-time price at a generation unit's location is the relevant price for valuing any deviations from a supplier's day-ahead schedule. Because wind and solar resources are intermittent and their actual output can differ significantly from what might be scheduled in the day-ahead market, they face exposure to the real-time price. The hourly real-time price is computed as the arithmetic mean of the twelve five-minute real-time prices during that hour. Because the real-time price is the opportunity cost of energy consumed or produced within the hour, this is the hourly price used to compute a generation unit owner's revenue stream throughout the remainder of this paper.

[Insert Figure 2 Here]

Suppose that p_{jhd} is the real-time LMP at location j during hour h of day d and q_{jhd} is the actual energy produced by generation units at location j during hour h of day d . For each hour of the sample, I compute the quantity-weighted average LMP for all $J = 630$ generation units in the California ISO control area. For each day d and hour of the day h , this price, p_{whd} , is equal to

$$p_{whd} = \frac{\sum_{j=1}^J p_{jhd} q_{jhd}}{\sum_{j=1}^J q_{jhd}}.$$

I also compute the quantity-weighted average price for all 53 wind and solar generation units in the California ISO control area for each hour in the sample. The left panel of Figure 2 reports the hourly mean and standard deviation of p_{whd} for $h=1,2,\dots,24$ across all days in

our sample for all generation units. The right panel of Figure 2 repeats these same two plots for the quantity-weighted average LMP for all wind and solar generation units. Although significantly less pronounced than was the case for the hourly demand versus the hourly output of all wind and solar units, the coefficient of variation of the quantity-weighted average price paid to wind and solar resources is uniformly higher than the coefficient of variation for the quantity-weighted average price for all generation units. This result is consistent with the greater hour-to-hour variability in the energy produced from wind and solar resources, which can also result in greater hour-to-hour variation in the LMPs paid to wind and solar units when transmission and other relevant operating constraints are binding.

To provide a measure of the extent of variability in prices across locations within each hour during our sample period, I construct the following measure of the hourly range of real-time prices paid to different generation units. Define

$$DF_{hd} = \max_{0 \leq j \leq J} P_{jhd} - \min_{0 \leq j \leq J} P_{jhd}, \quad (7)$$

which is the difference between the maximum real-time LMP over all $J = 630$ generation units during hour h of day d and the minimum real-time LMP over all $J = 630$ generation units during hour h of day d . The upper panel of Figure 3 plots the hourly means of DF_{hd} and the pointwise upper and lower 95 percent confidence intervals on realizations of DF_{hd} across days of our sample for each hour of the day.

[Insert Figure 3 Here]

I also compute a version of DF_{hd} taking the maximum over all of the real-time LMPs paid to all 53 wind and solar generation units during hour h of day d and the minimum of the real-time LMPs over these same units during the same hour and day. The lower panel of Figure 3 plots the hourly means of DF_{hd} and the pointwise upper and lower 95 percent confidence intervals on

realizations of DF_{hd} across days of our sample for each hour of the day all wind and solar generation units.

Two results emerge from Figure 3. First, the mean of the DF_{hd} for $h=1,2,\dots, 24$ across days is largest for the highest demand hours of the day. This is consistent with the logic described earlier that high demand hours are more likely to require managing transmission and other operating constraints by setting different prices at different locations in the transmission network. The second result is that although the hourly means and length of the 95 percent confidence interval of DF_{hd} is larger for the all 630 generation units in the California ISO control area, the corresponding features of the distribution of DF_{hd} for the 53 wind and solar generation units account for a significant fraction of the total spatial variation in prices within each hour of the day.

[Insert Table 2 Here]

[Insert Table 3 Here]

Tables 2 and 3 present the location-specific coefficients of variation for the hourly output and the hourly real-time price for all wind and solar locations in the California ISO control area. The final column of each table contains the sample correlation between the hourly output at each location and the hourly real-time price at that location. For each of the wind locations, the hourly output is negatively correlated with the hourly price, consistent with most of the energy production from wind units occurring during the low-priced periods of the day. For each of the solar locations, hourly output is positively correlated with the hourly price, consistent with most energy production from solar units occurring during the high-priced periods of the day.

4.2. Potential for Diversification to Reduce Output and Revenue Variability

A final descriptive analysis assesses the extent which is possible to construct a portfolio of wind and solar generation units that significantly reduces the standard deviation of the hourly

capacity factor, CF_h^T , and the standard deviation of the hourly revenue per MW of capacity, RC_h^T , relative to the values of these variables using the actual capacity shares. If the off-diagonal elements of the covariance matrices of F_h and R_h are positive and large (relative to the diagonal elements of these matrices), meaning there is significant positive correlation between elements of F_h and R_h , then the potential aggregate output or revenue variability reduction from optimizing the sizes of wind and solar generation units at existing locations in California is limited.

The spectral decomposition of a covariance matrix yields a useful summary measure of the degree of contemporaneous correlation between elements of a random vector that can be used to assess the extent to which diversification can significantly reduce the variability in aggregate output or revenues. Recall that F_h , the J -dimensional vector of hourly location specific capacity factors of the wind and solar units in the California ISO control area, is assumed to have mean vector μ and covariance matrix Ω . This covariance matrix has the spectral decomposition $\Omega = QDQ'$, where D is a diagonal matrix composed of the eigenvalues of Ω , ordered from the largest to smallest, and Q is an orthonormal matrix composed of the eigenvectors of Ω . Let $d_{11} \geq d_{22}, \dots, \geq d_{JJ}$ be these eigenvalues sorted from largest to smallest. Note that because Ω is a positive definite matrix, all of its eigenvalues are positive. The relative magnitude of the J eigenvalues quantifies the extent to which elements of F_h can be explained by common factors that cause them to be correlated. Note that if the elements of F_h were uncorrelated and all had the same variance, then all of the eigenvalues of Ω would be equal.⁴

The spectral decomposition theorem of a symmetric matrix can be used to reduce the matrix Ω into J orthogonal factors, where the sum of the variances of each factor is equal to the sum of the variances of the elements of F_h . This result is derived as follows. Let q_i equal the

⁴ The example presented at the beginning of Section 3 of N locations with the same independently distributed probability of providing 1 MWh of energy during the hour satisfies these conditions.

normalized eigenvalue (which is also the i^{th} column of Q') associated with the i^{th} eigenvector, d_{ii} . Because the eigenvectors are orthonormal, $q_i'q_i = 1$ for all i and $q_i'q_j = 0$ for $i \neq j$. Using the definition of an eigenvalue and eigenvector yields the following equation for each i ($i=1,2,\dots,J$), $\Omega q_i = d_{ii}q_i$. Pre-multiplying both sides of the equation by q_i' yields $q_i'\Omega q_i = d_{ii}$, because $q_i'q_i = 1$. Note that $q_i'\Omega q_i$ is also equal to the variance of $q_i'F_h$, and because $q_i'q_j = 0$, the covariance between $q_i'F_h$, and $q_j'F_h$ is zero for $i \neq j$. Thus, the i^{th} orthogonal factor is $q_i'F_h$ and the variance of this factor is d_{ii} .

Because the trace of Ω is equal to the sum of the eigenvalues of Ω and each random variable $z_{ih} \equiv q_i'F_h$ has variance equal to d_{ii} and is uncorrelated with all other z_{jh} , the sum of the variances of the elements of F_h , the trace of Ω , is equal to the sum of the variances of the z_{ih} for $i=1,2,\dots,J$. Consequently, one measure of the extent of correlation among elements of F_h is to normalize each eigenvalue by the sum of the eigenvalues of Ω , or the sum of the diagonal elements of Ω . Compute

$$S_k = \frac{\sum_{j=1}^k d_{jj}}{\sum_{j=1}^J d_{jj}} \quad \text{for } k=1,2,\dots,J. \quad (8)$$

The larger S_k is for a given value of k , the greater is the correlation between the elements of Z_h . S_k is the proportion of the sum of the variances of elements of F_h that is explained by the first k orthogonal factors, $z_{ih} \equiv q_i'F_h$ for $i=1,2,\dots,k$, each term of which has variance d_{ii} .

Figure 4 plots the values of S_k for $k=1,2,\dots,J$, for the solar units separately, the wind units separately, and the combination of wind and solar units. For a given k , the value of S_k is largest for all of the solar units. Next largest value of S_k is for the wind units alone. The combination of wind and solar units has the smallest value of S_k for a given value of k .

Each plot also contains the values of S_k for the hypothetical case that the hourly capacity factor at each of the N locations is independently distributed of the hourly capacity factor at all other locations and each location has the same variance of the hourly capacity factor. For each

generation technology, these values are straight lines using the same line-type as the actual values of S_k from the origin to the point (N,N) , where N is the number of locations with that technology. For example, for the case of just the solar units, this is the dotted line in Figure 4 from $(0,0)$ to $(13,13)$.

The results in Figure 4 demonstrate substantial differences between the case of independent hourly capacity factors with the same variance at all locations and the actual covariance matrix of hourly capacity factors across locations for: (1) solar locations, (2) wind locations, and (3) the combination of wind and solar locations. These results show that a single common factor is responsible for more than 80% of the hourly variation in the 13 solar generation units and more than 60% of the hourly variation in the 40 wind units. Even for the case of the 53 wind and solar units, one factor is responsible for more than 50% of the hourly variation.

[Insert Figure 4 Here]

[Insert Figure 5 Here]

Figure 5 presents the same calculation for R_h , the vector of unit-specific revenues per MW of capacity. The graph of the S_k for the hypothetical case of independent elements of R_h with the same variance is also presented for each of the three sets of generation locations. The actual S_k function rises at least as fast as it does for the vector of hourly capacity factors. The first factor accounts for more than 80% of hourly variation in the revenues earned by the 13 solar units. The first factor accounts for more than 70% of the hourly variation in the revenues earned by the 40 wind units. For the 53 wind and solar units, almost 60 percent of the hourly variation in revenues is accounted for by the first factor.

The results of applying the spectral decomposition theorem to the covariance matrices of F_h and R_h suggests that because of the high degree of contemporaneous correlation between the

elements of each of these vectors, there are likely to be modest reductions in the standard deviation in the aggregate of output of wind and solar units and more modest reductions in the standard deviation in the total revenues for wind and solar units from optimizing where these units are located within the state. Virtually all of the area in Figures 4 and 5 above the S_k lines that assume uncorrelated elements of F_h and R_h with identical variance is below the actual S_k curve for that combination of generation locations. Nevertheless, as we demonstrate in the next section, even with this potential for modest gains (relative to the uncorrelated and same variance for all locations case) from optimizing the portfolio of wind and solar units, significant increases in the average hourly output of renewable energy providers are possible without increasing the variability in hourly output by optimizing the locations and sizes of these generation units.

The magnitudes of the pairwise spatial correlations in the hourly capacity factors at 53 wind and solar locations in California are significantly larger in absolute value than those computed by Roques, Hiroux and Saguan (2010) for the country-level hourly wind capacity factors for Spain, Germany, Austria, Denmark and France. These capacity factors were computed using 2006 and 2007 hourly wind output data for each country normalized by the total amount of installed capacity in each country. With the exception of Austria and Germany, which have a pairwise correlation in their hourly wind outputs of 0.362, and France and Germany, which have a pairwise correlation in their hourly wind outputs of 0.147, none of the eight remaining pairwise correlations in hourly wind output is above 0.062 in absolute value.

The results in Figures 4 and 5 suggest that the use of country-wide aggregate hourly wind output may mask a higher contemporaneous correlation between nearby locations in the adjacent countries and which implies greater reliability benefits from diversification than is actually possible.

5. Computation of Efficient Frontier for Wind and Solar Output and Revenues

This section describes the computation of the mean-standard deviation of output and revenue efficient frontiers for wind investments, solar investments, and combined wind and solar investments. Each point along the efficient frontier computes the capacity shares for each existing wind and solar location that minimize the standard deviation of hourly output or hourly revenues subject to achieving a given level of expected hourly output or expected hourly revenues. This section also describes the computation of the maximum output risk-adjusted portfolio and the maximum revenue risk-adjusted portfolio and two measures of location-specific non-diversifiable wind and solar risk.

In terms of the notation from Section 3, each point on the hourly output efficient frontier is the solution to the following optimization problem in the J -dimensional vector w , for a given value C ,

$$\min_{w_j \geq 0} \left[\sum_{j=1}^J \sum_{k=1}^J w_j \hat{\omega}_{jk} w_k \right] \text{ subject to } C = \sum_{j=1}^J w_j \hat{\mu}_j \text{ and } 1 = \sum_{j=1}^J w_j \quad (9)$$

where $\hat{\omega}_{jk}$ is the estimate of ω_{jk} and $\hat{\mu}_j$ is the estimate of μ_j based on the sample values of F_h and J is equal to the number of solar locations, the number of wind locations, or the number of combined wind and solar locations. If H is the set of hours of data that forms our sample, then $\hat{\omega}_{jk}$ is the (j,k) element of $\hat{\Omega}$ and $\hat{\mu}_j$ is the j^{th} element of $\hat{\mu}$, where

$$\hat{\Omega} = \frac{1}{\text{card}(H)} \sum_{h \in H} (F_h - \hat{\mu})(F_h - \hat{\mu})' \text{ and } \hat{\mu} = \frac{1}{\text{card}(H)} \sum_{h \in H} F_h \quad (10)$$

and $\text{card}(H)$ is equal to the number of elements in the set H . This problem can be solved using standard quadratic programming methods. Solving the problem for each value of C from the minimum element of $\hat{\mu}$ to the maximum element of $\hat{\mu}$ yields the set of mean and standard deviation pairs along the efficient frontier.

Specifically, if $w_j^*(C)$ for $j=1,2,\dots,J$ is the solution to this optimization problem for a given value of C , then the point on the efficient frontier corresponding this value of C is equal to the mean-standard deviation pair:

$$(\sum_{j=1}^J w_j^*(C) \hat{\mu}_j, [\sum_{j=1}^J \sum_{k=1}^J w_j^*(C) \hat{w}_{jk} w_k^*(C)]^{1/2}). \quad (11)$$

The actual capacity-weighted hourly mean output and standard deviation of hourly output is plotted on the same diagram as the mean-standard deviation efficient frontier to illustrate how far this portfolio is from the efficient frontier. The actual capacity-weighted share point is equal to the portfolio mean-standard deviation pair:

$$(\sum_{j=1}^J w_j^{act} \hat{\mu}_j, [\sum_{j=1}^J \sum_{k=1}^J w_j^{act} \hat{w}_{jk} w_k^{act}]^{1/2}), \quad (12)$$

where the w_j^{act} ($j=1,2,\dots,J$) are the actual installed capacity weighted shares defined in Section 3.

I compute three efficient frontiers which involve different dimensions for the vector F_h . The first involves just the solar units, so that $J = 13$. The second involves just the wind units, so that $J = 40$. The third involves both wind and solar units, so that $J = 53$. In solving for the efficient frontier with both wind and solar units, the constraint is imposed that the overall capacity share of solar units must equal the actual capacity share of solar units. Suppose that the first 13 units are solar units, so this constraint becomes $w_{solar}^{act} = \sum_{j=1}^{13} w_j^{act}$. This implies that the efficient frontier for the combination of wind and solar units is the solution to:

$$\min_{w_j \geq 0} \{ \sum_{j=1,2,\dots,J}^J w_j \hat{w}_{jk} w_k \} \quad (13)$$

$$\text{subject to } C = \sum_{j=1}^J w_j \hat{\mu}_j, \quad 1 = \sum_{j=1}^J w_j, \quad w_{solar}^{act} = \sum_{j=1}^{13} w_j.$$

This efficient frontier minimizes the standard deviation of the portfolio-level hourly capacity factor subject to achieving a given expected hourly capacity factor and maintaining the same total capacity shares for wind and solar capacity as actually exists.

Two sets of hours, H , are used to compute the estimates of μ and Ω . The first uses all hours of the day and all days of the sample, so that $card(H)$ is equal to $8760 = 24 \text{ hours} \times 365 \text{ days}$. The second computes efficient frontiers for groups of hours of the day to quantify how far the actual portfolio is from the efficient frontier for different groups of hours of the day. For this calculation, hours of the day are separated into four groups, hours 1-6, 7-12, 13-18, and 19-24, so that in this case, $card(H) = 365 \times 4 = 1460$.

For each efficient frontier, I also compute the risk-adjusted maximum (RA-Max) output portfolio for that frontier. This is the point on the wind and solar efficient frontier that has the largest value of the ratio of the expected hourly output divided by the standard deviation of the hourly output. Let $w_j^*(\tilde{C})$ for $j=1,2,\dots,J$ equal the set of portfolio weights that solve this problem and \tilde{C} , the value C for these portfolio weights. The RA-Max expected hourly output portfolio is equal to:

$$\frac{\sum_{j=1}^J w_j^*(\tilde{C}) \hat{\mu}_j}{[\sum_{j=1}^J \sum_{k=1}^J w_j^*(\tilde{C}) \hat{\omega}_{jk} w_k^*(\tilde{C})]^{1/2}} \quad (14)$$

where \tilde{C} is the value of C that yields the largest value of this ratio.

Computing the efficient frontier for expected hourly revenues and the standard deviation of hourly revenues follows the exact same process but with the vector F_h replaced by the vector R_h . Each point on the hourly revenues efficient frontier is the solution to the following optimization problem in the J -dimensional vector w , for a given value D ,

$$\min_{w_j \geq 0 \{j=1,2,\dots,J\}} [\sum_{j=1}^J \sum_{k=1}^J w_j \hat{\gamma}_{jk} w_k] \text{ subject to } D = \sum_{j=1}^J w_j \hat{\lambda}_j \text{ and } 1 = \sum_{j=1}^J w_j \quad (15)$$

where $\hat{\gamma}_{jk}$ is the estimate of γ_{jk} and $\hat{\lambda}_j$ is the estimate of λ_j based on the sample values of R_h . For each of the sets of hours, H , defined above $\hat{\gamma}_{jk}$ is the (j,k) element of $\hat{\Gamma}$ and $\hat{\lambda}_j$ is the j^{th} element of $\hat{\lambda}$, where

$$\hat{\Gamma} = \frac{1}{\text{card}(H)} \sum_{h \in H} (R_h - \hat{\lambda}) (R_h - \hat{\lambda})' \text{ and } \hat{\lambda} = \frac{1}{\text{card}(H)} \sum_{h \in H} R_h. \quad (16)$$

Increasing the value of D from the minimum element of $\hat{\lambda}$ to the maximum element of $\hat{\lambda}$, and solving the problem for each value of D , yields the mean and standard deviation pairs along the efficient frontier.

Points along the efficient frontier take the form

$$(\sum_{j=1}^J w_j^*(D) \hat{\lambda}_j, [\sum_{j=1}^J \sum_{k=1}^J w_j^*(D) \hat{\gamma}_{jk} w_k^*(D)]^{1/2}). \quad (17)$$

We also plot the actual capacity-weighted expected hourly revenues per MW of capacity and standard deviation of hourly revenues per MW of capacity on the same diagram as the efficient mean-standard deviation frontier to demonstrate how far the actual portfolio of wind and solar resources is from this efficient frontier. The actual capacity-weighted share point is equal to the portfolio mean-standard deviation pair:

$$(\sum_{j=1}^J w_j^{act} \hat{\lambda}_j, [\sum_{j=1}^J \sum_{k=1}^J w_j^{act} \hat{\gamma}_{jk} w_k^{act}]^{1/2}), \quad (18)$$

where the w_j^{act} ($j=1,2,\dots,J$) are the actual installed capacity shares defined in Section 3. Similar to the case of the hourly output efficient frontier, for the combined wind and solar hourly revenues efficient frontier the constraint that the efficient capacity share of solar resources must equal the actual capacity shares of solar resources is imposed.

The same two sets of hours that are used to compute the efficient frontiers for hourly output are used to compute the estimates of λ and Γ . The RA-Max expected hourly revenue portfolio is equal to:

$$\frac{\sum_{j=1}^J w_j^*(\tilde{D}) \hat{\lambda}_j}{[\sum_{j=1}^J \sum_{k=1}^J w_j^*(\tilde{D}) \hat{\gamma}_{jk} w_k^*(\tilde{D})]^{1/2}}, \quad (19)$$

where \tilde{D} is the value of D that yields the largest value of this ratio.

Following the logic of the Capital Asset Pricing Model (CAPM), the actual capacity-share-weighted portfolio and the risk-adjusted maximum expected hourly output portfolio can be used to compute measures of the non-diversifiable hourly output risk associated with each wind and solar location in the California ISO control area. These measures of risk can be computed using the following linear regression:

$$f_{jh} = \alpha_j + \beta_j f_{mkt,h} + \varepsilon_{jh} \quad (20)$$

where $f_{mkt,h}$ is the weighted average of location-specific capacity factors that is computed as either the actual capacity share weighted average ($f_{mkt,h}^{act} = \sum_{j=1}^J w_j^{act} f_{jh}$) or as the RA-Max share weighted average ($f_{mkt,h}^{RA-Max} = \sum_{j=1}^J w_j^*(\tilde{C}) f_{jh}$). The coefficient β_j is the measure of the non-diversifiable risk associated with wind and solar production at location j . It measures the sensitivity of movements in f_{jh} to movements in $f_{mkt,h}$, the hourly weighted average capacity factor across all wind and solar locations.

A similar non-diversifiable wind and solar risk measure can be derived for the hourly revenue per MW of capacity from the linear regression

$$r_{jh} = \alpha_j + \beta_j r_{mkt,h} + \varepsilon_{jh} \quad (21)$$

where $r_{mkt,h}$ is the weighted average of location-specific revenues per MW that is computed as either the actual capacity share weighted average ($r_{mkt,h}^{act} = \sum_{j=1}^J w_j^{act} r_{jh}$) or as the RA-Max share weighted average ($r_{mkt,h}^{RA-Max} = \sum_{j=1}^J w_j^*(\hat{D}) r_{jh}$). The coefficient β_j is the measure of the non-diversifiable risk associated with hourly wind and solar revenues at location j . It measures the sensitivity of movements in hourly revenues per MW of capacity at location j , r_{jh} , to movements in the hourly weighted average of revenues per MW of capacity across all wind and solar locations.

6. Hourly Output and Hourly Revenue Efficient Frontiers

This section presents the results of computing the hourly output-per-unit-of-capacity and hourly-revenues-per-unit-of-capacity efficient frontiers. For each efficient frontier, the risk-adjusted maximum expected hourly output or expected hourly revenues portfolio is computed. The location-specific measures of non-diversifiable risk are computed for both the actual capacity-weighted-share output and revenue portfolios and the RA-Max expected-hourly-output or RA-Max expected hourly-revenues-portfolios.

[Insert Figure 6 Here]

[Insert Figure 7 Here]

Figure 6 plots the annual hourly output per MW of capacity efficient frontier for solar units alone, wind units alone, and the combination of wind and solar units. This efficient frontier computes the estimates of μ and Ω using the set H that contains all hours of the day and all days in the sample period. The actual capacity share-weighted portfolio is plotted with a “+” and the RA-Max expected hourly output maximizing portfolio is plotted with an “o”.

Several results emerge from this figure. First, the efficient frontier composed of only solar generation units lies to the right of the efficient portfolio of only wind units, suggesting that on an annual hourly basis, almost any efficient portfolio of only wind units dominates any efficient portfolio of only solar units. Second, the addition of solar units to a wind-only efficient portfolio only slightly increases the expected hourly output per MW of capacity for the same standard deviation of hourly output per MW of capacity. Third, a 48 percent increase in the expected hourly output per MW of capacity of the actual portfolio of wind and solar units (denoted by “+” in the figure) can be obtained without increasing the annual hourly variability in renewable output per MW by optimizing the capacity shares of individual wind and solar locations subject to the

constraint that the aggregate wind and solar capacity shares are equal to the actual total wind and solar capacity shares. The risk-adjusted expected hourly output per MW of capacity maximizing portfolio selects a 15 percent lower expected hourly output per MW of capacity but a 50% lower standard deviation of the hourly expected output per MW of capacity relative to the actual capacity-weighted-share portfolio.

Figure 7 presents efficient frontiers for solar, wind and the combined wind and solar resource locations for four 6-hour periods of the day. In this case, the estimates of μ and Ω are computed separately for four groups of hours of the day (1-6, 7-12, 13-18, and 19-24). For all of these groups of hours of the day, the actual portfolio of wind and solar units is significantly to the right of the efficient frontier. Moreover, this distance gets even greater during the hours of the day when there is significant wind and solar output. During hours 13-18, the benefits of combining wind and solar resources yields the greatest hourly output risk reductions relative to the efficient portfolios that contain only one of these generation technologies. This is confirmed by the larger horizontal distance between the wind only efficient frontier and the wind and solar efficient frontier for these hours.

During the daylight hours (7-12 and 13-18), the risk-adjusted output maximizing portfolio yields a lower expected hourly output per MW of capacity and a smaller standard deviation of hourly output per MW of capacity. During the night-time hours, the risk-adjusted output maximizing portfolio selects a higher expected hourly output per MW and a higher standard deviation of hourly output per MW of capacity. This result is consistent with the within-day pattern of wind output in the California ISO control area. Typically, the highest hourly capacity factors occur during the night-time hours.

[Insert Figure 8 Here]

[Insert Figure 9 Here]

Figure 8 computes the efficient frontier for hourly revenues per MW of capacity for all hours of the year. The set of hours used to estimate λ and Γ is composed of all hours of the day and all days of the sample period. The actual portfolio of wind and solar generation units yields higher expected revenues per MW of capacity than an efficient portfolio with the same hourly revenue risk composed of only solar units. The relative distance between the actual portfolio and portfolios on the efficient frontier for hourly revenue is smaller than is the case for hourly output. One explanation for this result is the positive correlation between hourly prices at solar facility locations and the hourly output at these locations shown in Table 3. For same amount of revenue risk per MW as the actual-capacity-weighted portfolio, moving to the efficient frontier increases expected hourly revenues per MW of capacity by 26 percent. Although these are sizeable percent changes, they are not as large in absolute value as the percent change for the hourly output per MW.

A bootstrap procedure can be used to compute the standard error of the difference between the 48 percent increase in the average hourly output to move to the output efficient frontier described earlier to the 26 percent increase in the average hourly revenue to move to the revenue efficient frontier. Although the actual portfolio of wind and solar generation units remains the same, the closest portfolio on each efficient frontier in terms of increasing the annual average value of output or revenue depends on the estimated values of the mean and variance of F_h and R_h . To compute this standard error, daily realizations of these two vectors are re-sampled to construct a bootstrap resample of the 8760 values of F_h and R_h . Estimates of the mean and variance of F_h and R_h are computed and the minimum distance to each efficient frontier is computed. The difference between the percentage change in the annual average hourly output to reach the efficient frontier

and the percentage change in the average hourly revenue to reach the efficient frontier is computed for each resample. This process is repeated B times and the standard deviation of these re-samples is a consistent estimate of the standard error of the difference. Dividing the observed difference, (48-26) by this standard error estimate yields a Z-statistic of 2.97 which is larger than the 0.01 percentile of a N(0,1) random variable, which implies rejection of the size 0.01 test of the null hypothesis that the two percentage changes are equal.

Figure 9 shows the hourly revenue per MW of capacity efficient frontiers for the four 6-hour periods of the day for solar locations, wind locations, and the combined wind and solar locations. In this case, the estimates of λ and Γ are computed using the 4 sets of hours, H_j , described above. These results for hourly revenue per MW of capacity are broadly similar to the results for the hourly output per MW of capacity. They demonstrate that for all hours of the day, the actual portfolio of wind and solar units is significantly to the right of the efficient frontier for that hour. The greatest revenue diversification benefits from combining wind and solar generation units occurs during the 13-18 hours of the day.

One question that immediately arises from the results in Figures 6 to 9 is: How do the capacity shares on the hourly-output-per-MW-of-capacity and hourly-revenue-per-MW-of-capacity efficient frontiers compare to the actual capacity shares? Figure 10 addresses this question for the hourly output per MW capacity frontier. The solid line in Figure 10 orders the wind and solar energy locations by the values of $w_j^*(\widehat{C})$, $j=1,2,\dots,J$, the risk-adjusted expected-hourly-output-maximizing portfolio shares. The “squares” are the values of the actual capacity shares in the same order. The most glaring difference between the two sets of weights is the relatively small number of non-zero values for the $w_j^*(\widehat{C})$ relative to the w_j^{act} , which are non-zero for all $j=1,2,\dots,J$. This figure also plots the weights associated with the point on the efficient

frontier that has the same annual standard deviation of hourly output as the actual capacity share-weighted portfolio but with a higher annual average capacity factor. These weights, in the same order of locations as the $w_j^*(\widehat{C})$, are denoted by “+” in the figure. The final series, denoted by “o” in the figure, plots the weights associated with the point on the efficient frontier that has the same annual average hourly output as the actual capacity share-weighted portfolio, but a lower standard deviation of hourly output. Both these portfolios on the efficient frontier have very few non-zero weights. No more than 15 of the 53 locations have non-zero weights in all of these portfolios. For both the “+” and “o” weights, the locations that are non-zero are very similar in relative magnitude to the RA-Max portfolio.

[Insert Figure 10 Here]

[Insert Figure 11 Here]

Figure 11 repeats the calculations in Figure 10 for the case of the hourly revenues per MW of capacity. The solid line in Figure 11 orders the wind and solar energy locations by the values of $w_j^*(\widehat{D})$, $j=1,2,\dots,J$, the RA-Max expected hourly revenue per MW of capacity portfolio shares. The “squares” are the values of the actual capacity shares in the same order. Again, the most glaring difference between the two sets of weights is the relatively small number of non-zero values for the $w_j^*(\widehat{D})$ relative to the w_j^{act} , which are all non-zero. This figure also plots the weights associated with the point on the efficient frontier that has the same annual standard deviation of the hourly revenue per MW as the actual capacity share-weighted portfolio but a higher annual average hourly revenue per MW. These weights, in the same order of locations as the $w_j^*(\widehat{D})$, are denoted by “+” in the figure. The final series, denoted by “o” in the figure, plots the weights associated with the point on the efficient frontier that has the same annual average hourly revenue per MW as the actual capacity share-weighted portfolio, but a lower annual

standard deviation of hourly revenue per MW. Both these portfolios have very few non-zero weights and the ones that are non-zero are very similar in relative magnitude to the RA-Max portfolio.

[Insert Figure 12 Here]

Figure 12 compares the RA-Max hourly output and RA-Max hourly revenue per MW portfolio weights. The solid line orders the weights according to the RA-Max mean hourly output weights. The RA-Max mean hourly revenue weights in the same order of locations are denoted by “o”. With five exceptions, the same wind and solar locations have non-zero weights for both sets of weights. In addition, for the non-zero weights, there appears to be a positive correlation between the value of the RA-Max hourly output weight and the RA-Max hourly revenue weight.

To determine whether the actual capacity shares and the three sets of portfolio weights shown in Figures 10 and 12 are statistically different, I perform the following hypothesis test:

$$H: w_i^M = w_i^{act} \text{ for } i=1,2,\dots,53 \text{ versus } K: w_i^M \neq w_i^{act} \text{ for at least one } i, \quad (22)$$

where w_i^M is the weight for wind or solar location i in efficient portfolio M , where M is either the (1) RA-Max portfolio, (2) the efficient portfolio with same mean but lower standard deviation (LS) than the actual portfolio, or (3) the efficient portfolio with the same standard deviation but higher mean (HM) than the actual portfolio. Because the portfolio weights for each of these efficient portfolios depends on the estimates of the mean and covariance matrix of F_h or R_h , I use a moving-blocks bootstrap to compute an estimate of the covariance matrix of the vector estimated portfolio weights for the RA-Max, LS and HM portfolios. For each of the $b=1,2,\dots,B$ bootstrap resamples of the estimated portfolio weights, I sample blocks of 7 consecutive days from the original sample with replacement to construct a new 365 values of F_h or R_h . Blocks of days of data on F_h and R_h

are resampled to account for the fact that the values of F_h or R_h are correlated across days. However, this correlation dies out as the time between observations increases.

The values of each set of efficient frontier portfolio weights are computed based on the estimates of the mean and covariance matrix of F_h or R_h using the resampled annual time series of values. The sample covariance matrix of the B resampled portfolio weight vectors is the moving blocks bootstrap covariance matrix estimate, $Var(W^M)$, that is used to construct the test statistic

$$(W^M - W^{act})'Var(W^M)^{-1}(W^M - W^{act}), \quad (23)$$

where W^{act} is the 52-dimensional vector of the 52 largest capacity shares for the actual portfolio, W^M is the 52-dimensional vector of capacity shares for these same locations for portfolio M , and $Var(W^M)$ is the moving blocks bootstrap estimate of the covariance matrix of W^M . (Note that the test statistic drops one capacity share because the sum of all capacity shares in all of the portfolios is equal to one.) This test statistic is asymptotically distributed as a χ^2 -distribution with 52 degrees of freedom under the null hypothesis. Table 4 reports the test statistics for the three hypothesis tests for the hourly output and hourly revenue efficient frontiers. For all cases the null hypothesis is overwhelmingly rejected, which implies that the three portfolios on the hourly output and hourly revenue efficient frontiers all have statistically different portfolio weights from the actual capacity shares.

[Insert Table 4 Here]

The results in Figures 10 to 12 imply that by concentrating California's total wind and solar capacity investments at fewer locations a capacity mix can be achieved that has either a higher annual average hourly output and a lower annual variability in hourly output or a higher average hourly revenue and lower annual variability in hourly revenue. Moreover, because of the high degree of positive correlation between the non-zero weights for the two RA-Max portfolios, it is

possible to move closer to both the hourly output and hourly revenue efficient frontiers with a different capacity mix from the present one.

[Insert Table 5 Here]

[Insert Table 6 Here]

The upper left panel of Figure 13 plots pairs of $\hat{\beta}_j$, the estimated market-specific renewable energy output risk at location j , using $r_{mkt,h}^{act} = \sum_{j=1}^J w_j^{act} r_{jh}$ as the market portfolio and the value of $\hat{\mu}_j$, the mean hourly capacity factor associated with location j . The upper right panel repeats this same plot using $r_{mkt,h}^{RA-Max} = \sum_{j=1}^J w_j^*(\hat{D}) r_{jh}$ as the market portfolio. Both of these figures also plot two estimates of the analogue of the Security Market Line from the Capital Asset Pricing Model (CAPM) relating the market-specific output risk at a location j to the expected capacity factor at location j . The solid line in the figure is the ordinary least squares fit from regressing $\hat{\mu}_j$ on $\hat{\beta}_j$ for all wind and solar locations and the dotted line is the fit using only observations with a non-zero capacity shares in the RA-Max portfolio on the hourly output efficient frontier. The bottom two panels of Figure 13 repeats these calculations for the market-specific measures of renewable energy revenue risk at each location, where the $\hat{\beta}_j$ in the bottom left panel are estimated using the actual capacity shares to compute the market portfolio revenue per MW of capacity and the $\hat{\beta}_j$ in the bottom right panel are estimated using the RA-Max portfolio weights on the hourly revenue efficient frontier to compute the market portfolio.

[Insert Figure 13 Here]

Several conclusions emerge from this figure. First, across all four panels, the fitted lines relating expected output and expected revenues to their respective market-specific risk measures using only wind and solar resource locations with non-zero shares in the RA-Max portfolios yield uniformly higher expected output and expected revenue for the same value of market-specific risk

than the fitted lines that use all wind and solar resource locations. This result clarifies why many of the actual wind and solar locations did not receive a positive weight on the output or revenue efficient frontiers. A higher expected hourly capacity factor or a higher expected revenue per MW of capacity can be obtained for the market-specific risk measure value at locations with zero weights in the RA-Max portfolio by taking a weighted average of locations with non-zero weights in the RA-Max portfolio.

Second, using the RA-Max output or revenue portfolio as the market portfolio implies a much greater range of values of market-specific risk measures across locations. The range of estimated values of the market-specific measure is roughly double the range of values for the case in which actual portfolio shares are used to compute the market portfolio. This occurs because many actual wind and solar resources locations have zero weights in the efficient RA-Max output and revenue portfolios.

Third, the fit of pairs of expected output and market-specific risk and pairs of expected revenue and market-specific risk measures to the analogue to the Security Market Line are far superior for market-specific risk measures computed using the RA-Max portfolio as the market portfolio. Consistent with our earlier result that the actual portfolio is relatively closer to the hourly revenue efficient frontier than to the hourly output efficient frontier, the fit of the expected revenue and market-specific risk line, using only locations with non-zero weights in the RA-Max revenue portfolio, is far better than the fit of the expected output and market-specific risk line which uses only locations with non-zero weights in the RA-Max output portfolio.

6.1. Qualifications and Caveats with Counterfactual Results

The empirical results presented in this section demonstrate the possibility of increasing the annual average hourly output of wind and solar resources in the California ISO control area without

increasing the annual standard deviation of this hourly output by changing the locational capacity shares of the state's wind and solar generation capacity. It is important to emphasize what is being assumed in constructing the counterfactual capacity investments implied by these efficient frontiers. Each point along the efficient frontier assumes that the mean and covariance matrix of F_h would be same regardless of how much generation capacity at each location is constructed, subject to the constraint that the total amount of wind and solar capacity built in California is equal to the aggregate amount that has actually been built.

Clearly, one explanation for higher mean capacity factors and lower standard deviation of capacity factors at some locations in the California ISO control area is the vintage of the technology installed at that location. However, the sample correlation between the age of the generation units at each wind location and the average annual capacity factor at that location does allow rejection of the 0.05 test of the null hypothesis of zero correlation. This could be explained by the fact that although wind capacity was installed at locations with the best wind resources first, the efficiency of the turbines installed over time increased fast enough to counteract this effect.

An alternative approach to this analysis would be to use wind speed data at each location and assume standardized wind energy to electricity conversion factors to compute an estimated hourly capacity factor at each location. Green and Vasilakos (2010) take this approach in their study of the impact of large amounts of wind generation on the behavior of hourly prices in the Great Britain electricity market in 2020. The strength of their approach is that a standardized technology for converting wind energy to electricity can be applied across all wind resource locations. However, this approach does not account for how the generation technology actually performs at each wind resource location. Using the actual hourly wind energy production at each wind resource location, addresses this concern. A topic for future research is to compare the hourly

output of a wind generation unit at each location estimated using hourly wind speed measurements to the actual hourly output of the wind units at that location and the hourly solar output at each location estimating using hourly solar irradiance data to the actual hourly output of the solar units at that location.

7. Implications for Wholesale Market Design with Significant Wind and Solar Goals

The empirical results presented in the previous section demonstrate the reliability and operating cost benefits of accounting for the impact on the annual average hourly output and annual variability in the hourly output of renewable energy in making wind and solar capacity investment decisions. Locating wind and solar facilities to maximize the increase in the annual average hourly output for a given increase in the annual standard deviation in hourly output can yield significant reliability benefits and operating cost savings.

The results from Section 6 argue in favor of taking these reliability and operating reserves cost savings into account in making wind and solar generation investment decisions. By doing so, investments in wind or solar generation capacity could provide the largest possible increase in average hourly renewable energy output while limiting the aggregate reliability consequences of this investment. Conversely, investments made in this manner would minimize increases in the annual variability of hourly wind and solar output for a given increase in the average hourly energy output from these generation units.

The difference between the relative distance to the hourly revenue frontier versus the hourly output frontier for the actual portfolio of wind and solar investments provides a rough measure of the extent to which failing to account for these reliability and operating cost savings distorts wind and solar location decisions. If one is willing to assume that the cost per megawatt of constructing a wind or solar generation unit is roughly the same for all existing locations in

California, then an expected profit-maximizing investor in wind or solar generation capacity in California will choose the location that yields the highest expected revenues, because the cost of constructing and operating the unit is roughly the same at all locations. As the analyses in the previous section have demonstrated, this choice will likely increase the annual standard deviation in the hourly output of all wind and solar units in the California ISO control area because the supplier does not take into consideration the impact of its investment decision on overall grid reliability and only considers the expected profits this individual investment will earn. If prices at that location are sufficiently high and/or positively correlated with renewable energy output at that location, the expected profit-maximizing location decision may not even increase the annual average hourly output of wind and solar energy.

This logic is consistent with the results presented earlier that the closest portfolio on hourly revenue efficient frontier with the same annual variability in hourly revenues as the actual portfolio of wind and solar generation investments has a 26 percent higher annual average hourly revenue, whereas the closest portfolio on the hourly output frontier with the same annual variability in hourly output as the actual portfolio has a 48 percent higher annual average hourly output. As shown in Section 6, the annual average hourly output percent increase is statistically significantly larger than the annual average hourly revenue increase. The difference in these estimates points towards significant reliability improvements and operating reserves cost savings from accounting for these reliability benefits in wind and solar capacity location decisions.

To provide quantitative evidence in favor of this claim, the left panel of Figure 14 reports a histogram of values of following variable for each location. Starting from the actual portfolio of wind and solar generation capacity in California, compute the percentage increase in the annual average hourly output of all wind and solar units associated with adding one more 1 MW of wind

(if it is a wind location) or solar capacity (if it is a solar location) at wind or solar location k . Call this variable $\% \Delta E(Q) / \Delta MW(k)$. For each location compute the percentage increase in the annual standard deviation of hourly output of all wind and solar units associated with adding one more 1 MW of wind or solar capacity at each wind or solar location k . Call this variable $\% \Delta \sigma(Q) / \Delta MW(k)$. Compute the ratio of $[\% \Delta E(Q) / \Delta MW(k)] / [\% \Delta \sigma(Q) / \Delta MW(k)]$ and call this variable $\% \Delta E(Q) / \% \Delta \sigma(Q)(k)$. This variable measures the percentage increase in the annual average hourly renewable output for a one unit increase in the annual standard deviation of hourly renewable output that results from a 1 MW investment at location k . Following this same procedure for hourly revenue produces the variable $\% \Delta E(R) / \% \Delta \sigma(R)(k)$, the percentage increase in the annual average hourly wind and solar revenue for a one unit increase in the annual standard deviation of hourly wind and solar revenue associated with a 1 MW investment at location k . The histogram of values of $\% \Delta E(R) / \% \Delta \sigma(R)(k)$ are plotted in the right panel of Figure 14.

[Insert Figure 14 Here]

The histogram of values of $\% \Delta E(Q) / \% \Delta \sigma(Q)(k)$ demonstrate that substantial reliability and cost savings that are possible from a policy that sites wind and solar facilities based on the value of $\% \Delta E(Q) / \% \Delta \sigma(Q)(k)$ at that location. There are renewable locations with values of $\% \Delta E(Q) / \% \Delta \sigma(Q)(k)$ that are ten times larger than the values at the majority of locations. There are a significant number of locations where the value of $\% \Delta E(Q) / \% \Delta \sigma(Q)(k)$ is 5 times larger than the majority of values, and even a small number of locations where an additional MW of renewable capacity would both increase the annual average hourly output and reduce the annual standard deviation of hourly output, which yields a negative value of $\% \Delta E(Q) / \% \Delta \sigma(Q)(k)$. The heterogeneity in values of $\% \Delta E(R) / \% \Delta \sigma(R)(k)$ is substantially less than the heterogeneity in values of $\% \Delta E(Q) / \% \Delta \sigma(Q)(k)$. The values of $\% \Delta E(Q) / \% \Delta \sigma(Q)(k)$ range from slightly more than -5 to

less than 12, whereas the values of $\% \Delta E(R) / \% \Delta \sigma(R)(k)$ range from slightly more than 0.5 to 2.5, which is consistent with the logic that profit-maximizing entry decisions focus on maximizing expected revenues.

[Insert Figure 15 Here]

[Insert Figure 16 Here]

Figure 15 repeats these calculations for just wind generation capacity. Figure 16 repeats them for just the solar generation capacity. The major result to emerge from these figures is that accounting for the reliability and cost savings associated with wind investments and combinations of wind and solar investments is likely to yield greater operating cost savings than for solar investments alone. The heterogeneity in the values of $\% \Delta E(Q) / \% \Delta \sigma(Q)(k)$ for the wind-only portfolio is substantially greater than for solar-only portfolio. This result is consistent with the results presented in Figure 4 and 5 showing much greater contemporaneous correlation in both hourly output and hourly revenues across solar locations than wind locations.

7.1. Methods for Incorporating Reliability and Costs Benefits

These results emphasize that a major challenge in wholesale electricity markets with significant wind and solar energy goals is to provide the appropriate incentives for wind and solar project developers to locate and size these investments to maximize system-wide wind and solar output subject to a given increase in the standard deviation in system-wide hourly wind and solar output. By providing incentives for wind and solar project developers to internalize the impact their location decisions impose on the cost of managing real-time system balance, a lower system-wide cost of achieving a given wind and solar energy production goal can be achieved.

There are a number of potential approaches to achieving this goal. A first step is for control areas to compile and make publicly available information similar to what is reported in the

previous section. Values for $\% \Delta E(Q) / \% \Delta \sigma(Q)(k)$ at each wind or solar resource location would provide valuable information to the control area operator, new wind and solar capacity investors, and transmission planners on the reliability implications of different new capacity investment decisions, because they would know which locations result in the greatest risk-adjusted increase in the annual average hourly output of renewable energy.

The aggregate wind and solar energy risk associated with the location and scale of investments in these technologies could be made a factor in the transmission planning process to support wind and solar energy capacity expansion in any control area. Specifically, keeping all other factors the same, expanding transmission capacity into wind and solar resource regions with the largest values of $\% \Delta E(Q) / \% \Delta \sigma(Q)(k)$ would yield significant savings in real-time system operating costs.

The next step could be to eliminate any implicit support mechanisms for renewable investments that unnecessarily increase the incentive of wind and solar developers to build new capacity at locations that increase the aggregate volatility of wind and solar energy production. Long-term contracts or feed-in-tariffs (FITs) for wind and solar resources that pay a fixed per MWh price for all of the energy produced each hour of the year are prime example of the sort of financing scheme that provides strong incentives to enhance, rather than reduce, hourly energy output volatility from wind and solar resources. Transitioning wind and solar forward contracting to the standard fixed-price and fixed-quantity contract used to finance dispatchable generation units would reduce this incentive to enhance wind and solar output volatility.

In the multi-settlement markets that exist throughout the United States, making wind and solar generation unit owners liable for the hourly imbalance charges associated with deviations between the amount of energy they sell in the day-ahead market and the amount of energy the units

actually produce in real-time would increase the likelihood that new wind and solar investments are made with some consideration for the aggregate wind and solar output volatility associated with these investments.

A final mechanism could be to charge differential grid interconnection fees inversely related to the value of $\frac{\Delta E(Q)}{\Delta \sigma(Q)}(k)$ to provide the price signals necessary to cause investors in new renewable generation capacity to take into account the increased system operating costs associated with their location decisions. However, given the substantial complexity associated with the determining the appropriate interconnection fees, this kind of policy must be implemented with great care.

8. Conclusions

Because a growing number of jurisdictions have set ambitious renewable energy goals, understanding how the aggregate output risk of intermittent wind and solar resources scale with the location and magnitude of these capacity investments is increasingly important to maintaining a reliable supply of electricity. The mean and standard deviation efficient frontiers for the hourly output and hourly revenues derived above are important tools for determining the magnitude of the level versus variability trade-offs in total wind and solar output.

Using hourly output data from the California ISO control area for a one-year time period, this paper computed the hourly output and hourly revenue efficient frontier for all wind and solar resource locations in the California ISO control area. Points along the efficient frontier are compared to the actual capacity-weighted portfolio of wind and solar resources in California. For both mean/standard deviation efficient frontiers, economically meaningful differences between portfolios on the efficient frontier and the actual wind and solar capacity mix are found. The relative difference was particularly large for the hourly wind and solar output frontier, with a 48

percent increase in the expected hourly output of wind and solar energy possible without increasing the standard deviation of the total renewable hourly output beyond its actual level. For the hourly revenues efficient frontier, a 26 percent increase in mean hourly total renewable revenues was found to be possible without increasing the standard deviation of hourly revenues beyond its current level. These results were shown to be consistent with expected profit-maximizing behavior by solar and wind capacity investors under the current renewables interconnection policies.

Most of the hourly output and hourly revenue risk-reducing benefits of optimal wind and solar generation capacity location decisions in California can be captured by employing a mix of wind resources, with the addition of solar resources only slightly increasing the set of feasible portfolio mean and standard deviation combinations. The diversification benefit from including solar generation units in the portfolio of renewable resources is particularly modest for the hourly revenue efficient frontier.

The risk-adjusted maximum expected hourly output portfolio on the efficient hourly output frontier and the risk-adjusted maximum expected hourly revenue maximizing portfolio on the efficient hourly revenue frontier were computed. In both cases, weights for these portfolios focused on a small number of wind and solar locations, with the vast majority of existing locations having a zero portfolio weight. The market-specific risk measures for the risk-adjusted maximum output portfolio and the risk-adjusted maximum revenue portfolio found approximately linear relationships between the actual expected output at a location and the market-specific risk output measure at that location and actual expected revenue at a location and the market-specific revenue risk measure at that location, particularly for the locations with non-zero portfolio weights in risk-adjusted maximum output and revenue portfolios.

References

- Awerbuch, Shimon, and Berger, Martin (2003) “Applying portfolio theory to EU electricity planning and policy-making,” IEA/EET Working paper No. EET/2003/03, IEA/OECD, Paris.
- Bar-Lev, Dan and Katz, Steven (1976) “A portfolio approach to fossil fuel procurement in the electricity utility industry,” *Journal of Finance*, 31(3), 933-947.
- Bazilian, Morgan and Roques, Fabien (editors) (2008) *Analytical methods for energy diversity and security—portfolio optimization in the energy sector: a tribute to the work of Dr. Shimon Awerbuch*, Elsevier, Oxford/Amsterdam.
- Jha, Akshaya, and Wolak, Frank A. (2013) “Testing for Market Efficiency with Transactions Costs: An Application to Financial Trading in Wholesale Electricity Markets,” available from <http://www.stanford.edu/~wolak>.
- Bhattacharya, Anindya and Kojima, Satoshi (2012) “Power sector investment risk and renewable energy: A Japanese case study using portfolio risk optimization method,” *Energy Policy*, 40, 69-80.
- Delarue, Erik; Jonghe, Cedric; Belmans, Ronnie; and D’haeseleer, William (2011) “Applying portfolio theory to the electricity sector: Energy versus power,” *Energy Economics*, 33, 12-23.
- Garnier, Ernesto and Madlener, Reinhard (2016) “The Influence of Regime Risks on Investments on Innovative Energy Technology,” *The Energy Journal*, this issue. (please insert correct page numbers).
- Green, Richard and Vasilakos, Nicholas (2010) “Market behavior with large amounts of intermittent generation,” *Energy Policy*, 38, 3211-3220.
- Madlener, Reinhard (2012) “Portfolio Optimization of Power Generation,” in Zheng, Qipeng P.; Rebennack, Steffen; Pardalos, Panos M.; Pereira, Mario V.F.; and Iliadis, Nikos A. (editors) *Handbook of CO₂ in Power Systems*, Springer, Berlin/New York, 275-296.
- Roques, Fabian; Newbery, David M.; and Nuttal, William J. (2008) “Fuel mix diversification incentives in liberalized electricity markets: A Mean-Variance Portfolio theory approach,” *Energy Economics*, 30, 1831-1849.
- Roques, Fabian; Hiroux, Celine; and Sagan, Marcelo (2010) “Optimal wind power deployment in Europe—A portfolio approach,” *Energy Policy*, 38, 3245-3256.
- White, Halbert (1980) “Using Least Squares to Approximate Unknown Regression Functions,” *International Economic Review*, vol. 21, no. 1, 149-170.
- Westner, Günther, and Madlener, Reinhard (2010) “The benefit of regional diversification of cogeneration investments in Europe: A mean-variance portfolio analysis,” *Energy Policy*, 38, 7911-7920.

Westner, Günther, Madlener, Reinhard (2011) Development of cogeneration in Germany: A mean-variance portfolio analysis of individual technology's prospects in view of the new regulatory framework," *Energy*, 36, 5301-5313.

Yu, Zuwei (2003) "A spatial mean-variance MIP model for energy market risk analysis, *Energy Economics*, 25(3), 255-268.

Table 1: Total Generation Capacity by Technology

Fuel Type	Total Capacity (MW)
Wind	3,040
Solar	499
Small Hydro	1,144
Large Hydro	7,776
Biomass	502
Geothermal	1,404
Nuclear	4,550
Gas	32,428
Other	4,252
Total	55,595

Table 2: Hourly Output and Hourly Price Coefficient of Variations by Wind Location

Plant ID	Output CV	Price CV	Corr(Output,Price)
1	1.285	1.764	-0.083
2	1.244	1.764	-0.081
3	0.950	1.766	-0.100
4	1.110	1.765	-0.094
5	1.024	1.765	-0.092
6	1.530	1.382	-0.090
7	1.071	1.198	-0.105
8	1.084	1.198	-0.116
9	0.979	1.198	-0.106
10	1.019	1.198	-0.114
11	1.114	1.326	-0.049
12	0.930	1.325	-0.047
13	0.847	1.464	-0.099
14	1.008	1.191	-0.130
15	1.246	1.191	-0.114
16	0.987	1.326	-0.037
17	1.080	1.326	-0.069
18	0.986	1.242	-0.016
19	1.299	1.188	-0.066
20	1.508	1.191	-0.108
21	1.390	1.191	-0.094
22	0.899	1.805	-0.107
23	0.933	1.323	-0.068
24	0.976	1.323	-0.064
25	0.969	1.323	-0.058
26	1.023	1.326	-0.078
27	1.615	1.190	-0.120
28	1.010	1.326	-0.043
29	1.354	1.191	-0.103
30	1.379	1.192	-0.097
31	1.008	1.206	-0.120
32	1.122	1.198	-0.114
33	1.299	1.198	-0.098
34	1.404	1.191	-0.095
35	2.022	1.200	-0.094
36	1.084	1.690	-0.076
37	1.070	1.687	-0.080
38	0.939	1.326	-0.066
39	1.164	1.319	-0.114
40	1.360	1.191	-0.106

Table 3: Hourly Output and Hourly Price Coefficient of Variations by Solar Location

Plant ID	Output CV	Price CV	Corr(Output,Price)
1	1.325	1.333	0.083
2	1.404	1.317	0.126
3	1.353	1.367	0.072
4	1.881	1.301	0.095
5	1.288	1.303	0.116
6	1.309	1.305	0.119
7	6.655	1.752	0.080
8	1.512	1.193	0.098
9	1.417	1.228	0.088
10	1.676	1.306	0.091
11	2.136	1.300	0.126
12	1.431	1.190	0.090
13	1.377	1.316	0.120

Table 4: Hypothesis Tests that Efficient Frontier Portfolio Weights Equal Actual Capacity Shares

Portfolio	χ^2 Statistic
Hourly Output Efficient Portfolio Weights	
RA-Max	356
Lower SD	407
Higher Mean	961
Hourly Revenue Efficient Portfolio Weights	
RA-Max	389
Lower SD	261
Higher Mean	639
99% Critical Value from χ^2_{52}	78.62

Figure 1: Within-Day Hourly Means and Standard Deviations of System Demand and Wind and Solar Output

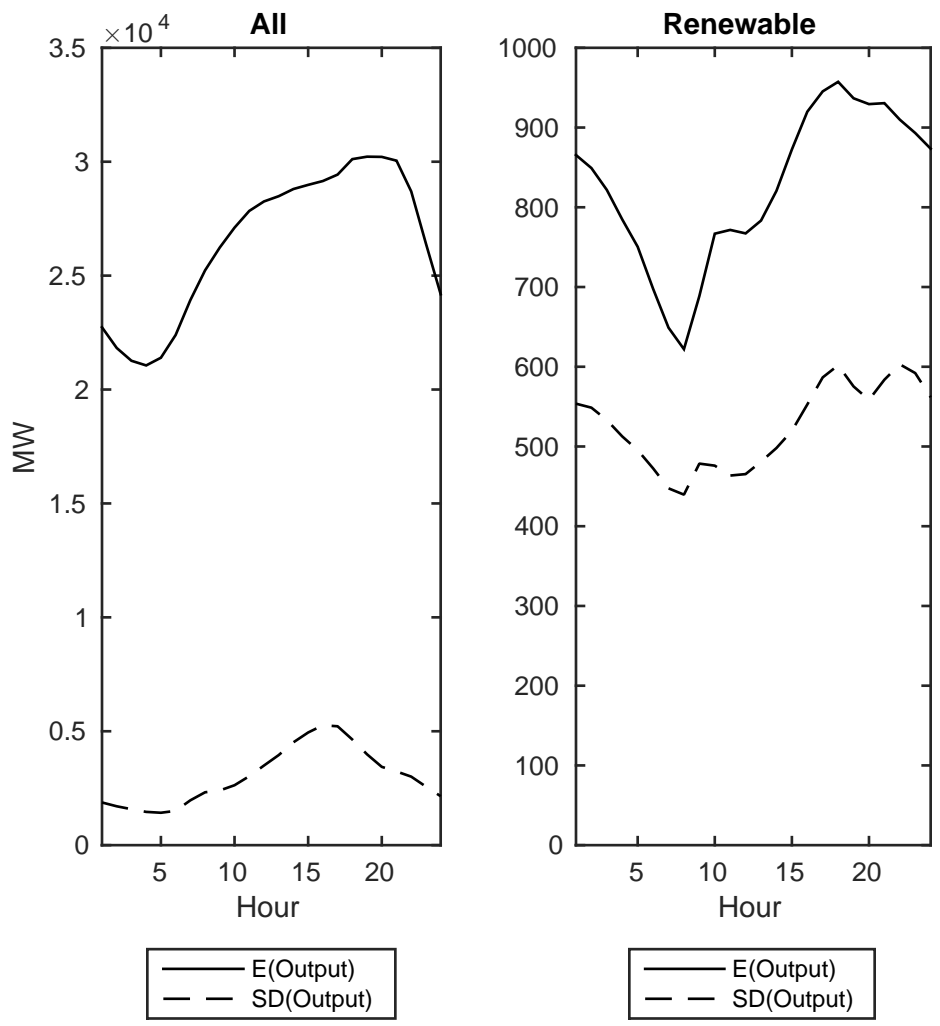


Figure 2: Within-Day Hourly Means and Standard Deviations of Weighted Average Real-Time Price and Weighted Average Wind and Solar Real-Time Price

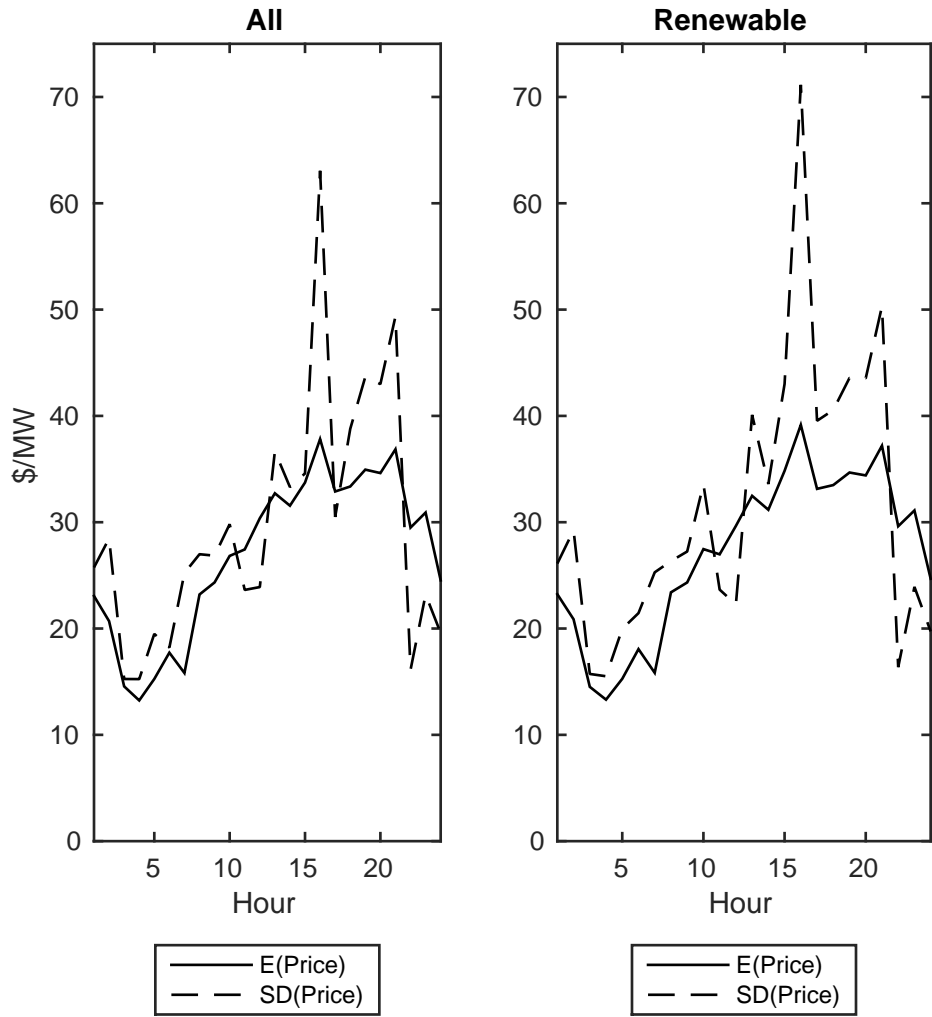


Figure 3: Hourly Means and 95 Percent Confidence Intervals for Spatial Price Range

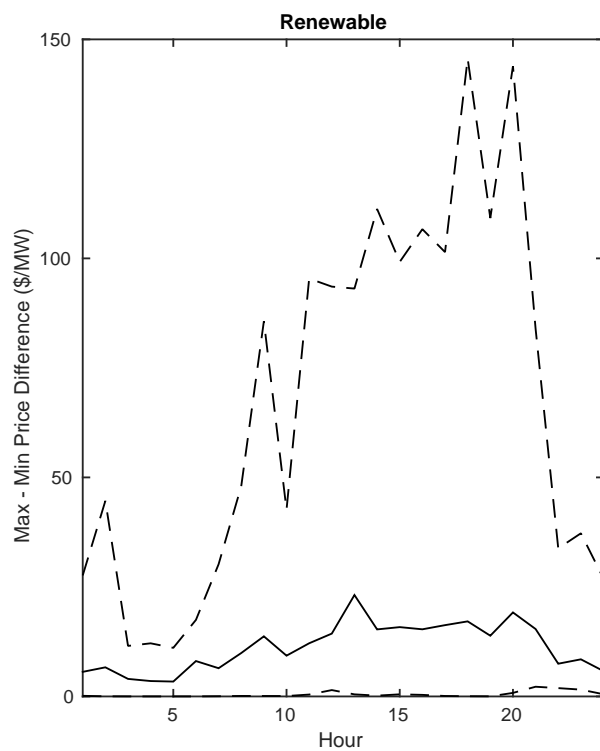
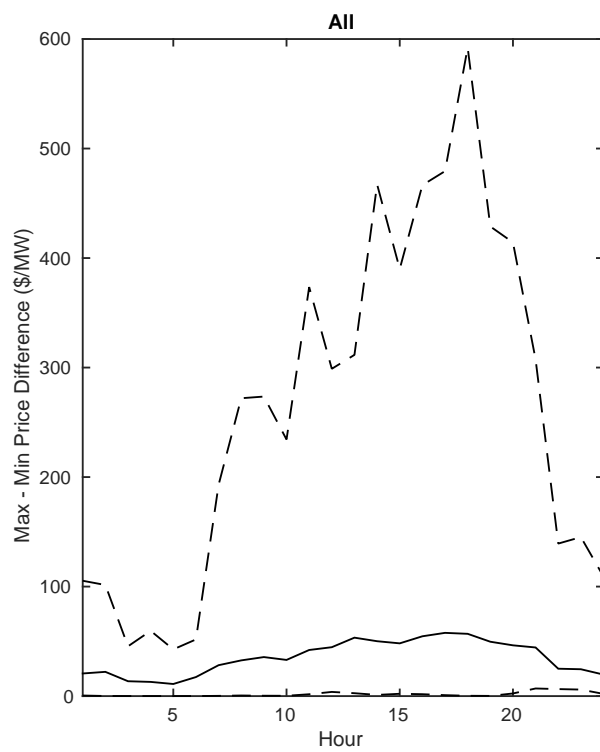


Figure 4: Sum of Normalized Eigenvalues of Covariance Matrix of Hourly Location-Specific Capacity Factors- Ω

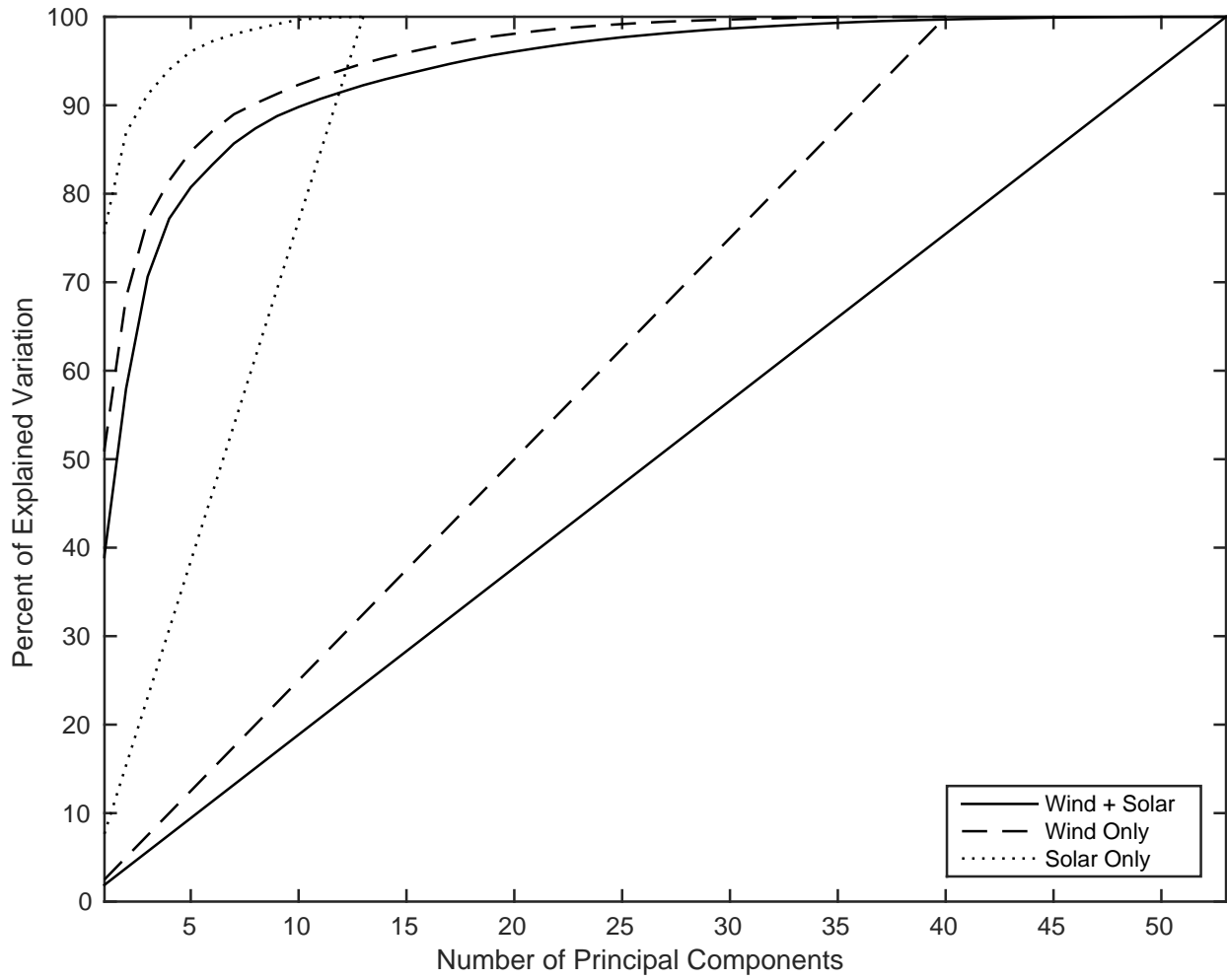


Figure 5: Sum of Normalized Eigenvalues of Covariance Matrix of Hourly Location-Specific Revenues per MW- Γ

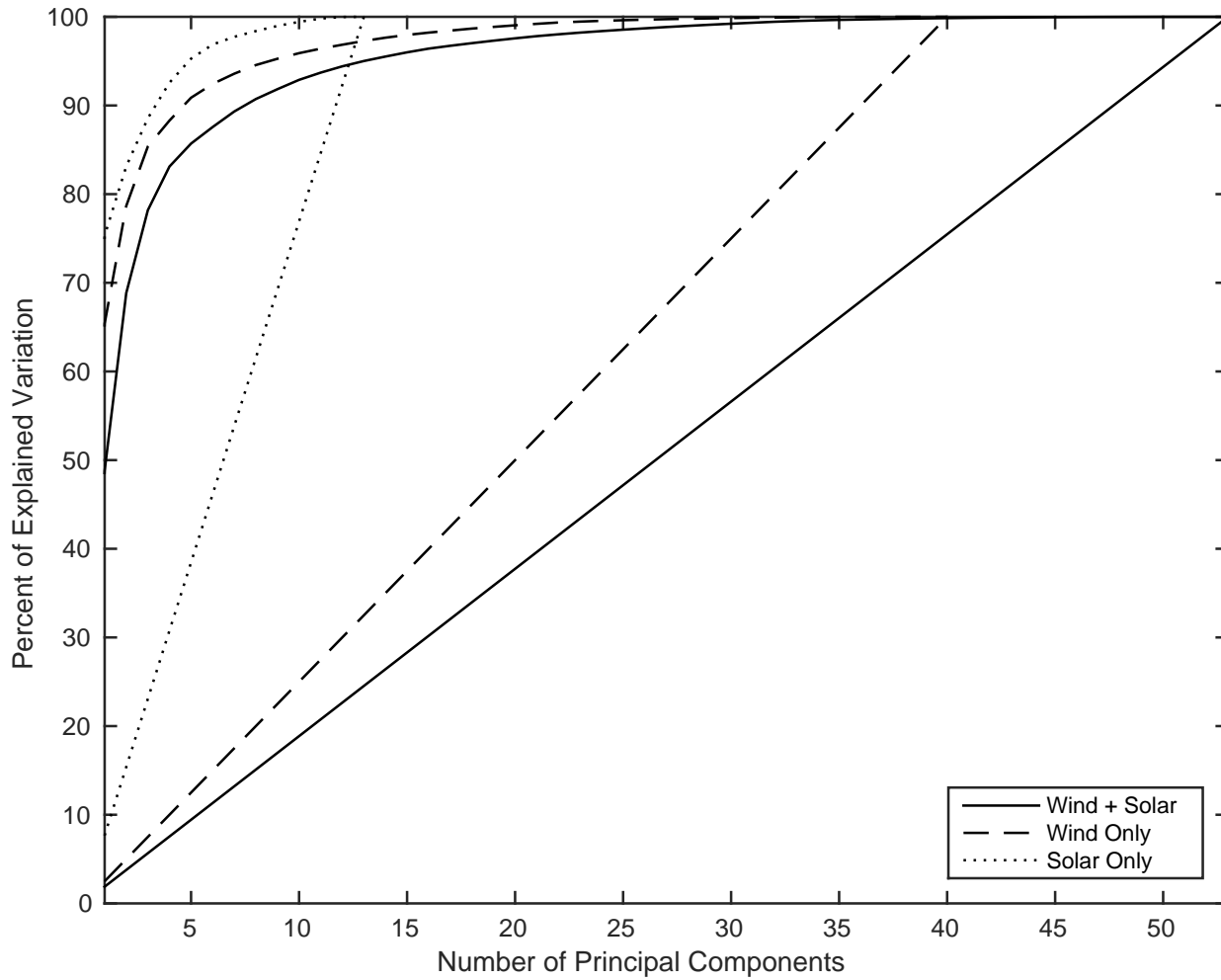


Figure 6: Expected Hourly Output per MW and Standard Deviation of Hourly Output per MW Efficient Frontier

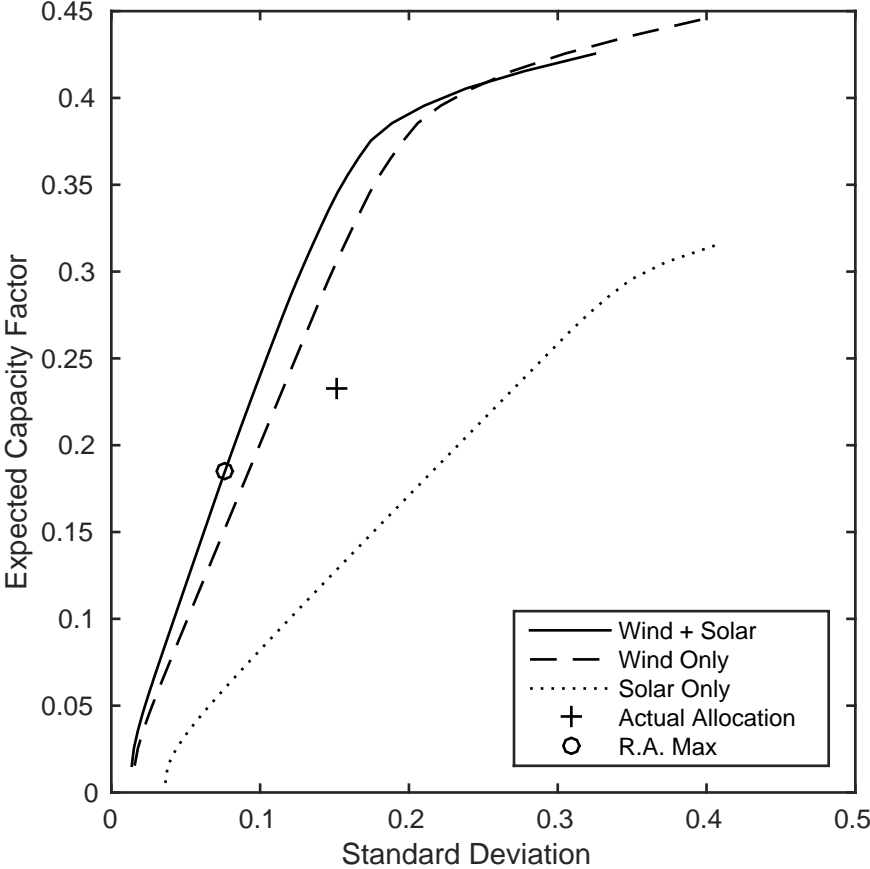


Figure 7: Expected Hourly Output per MW and Standard Deviation of Hourly Output per MW Efficient Frontier by Groups of Hours of the Day

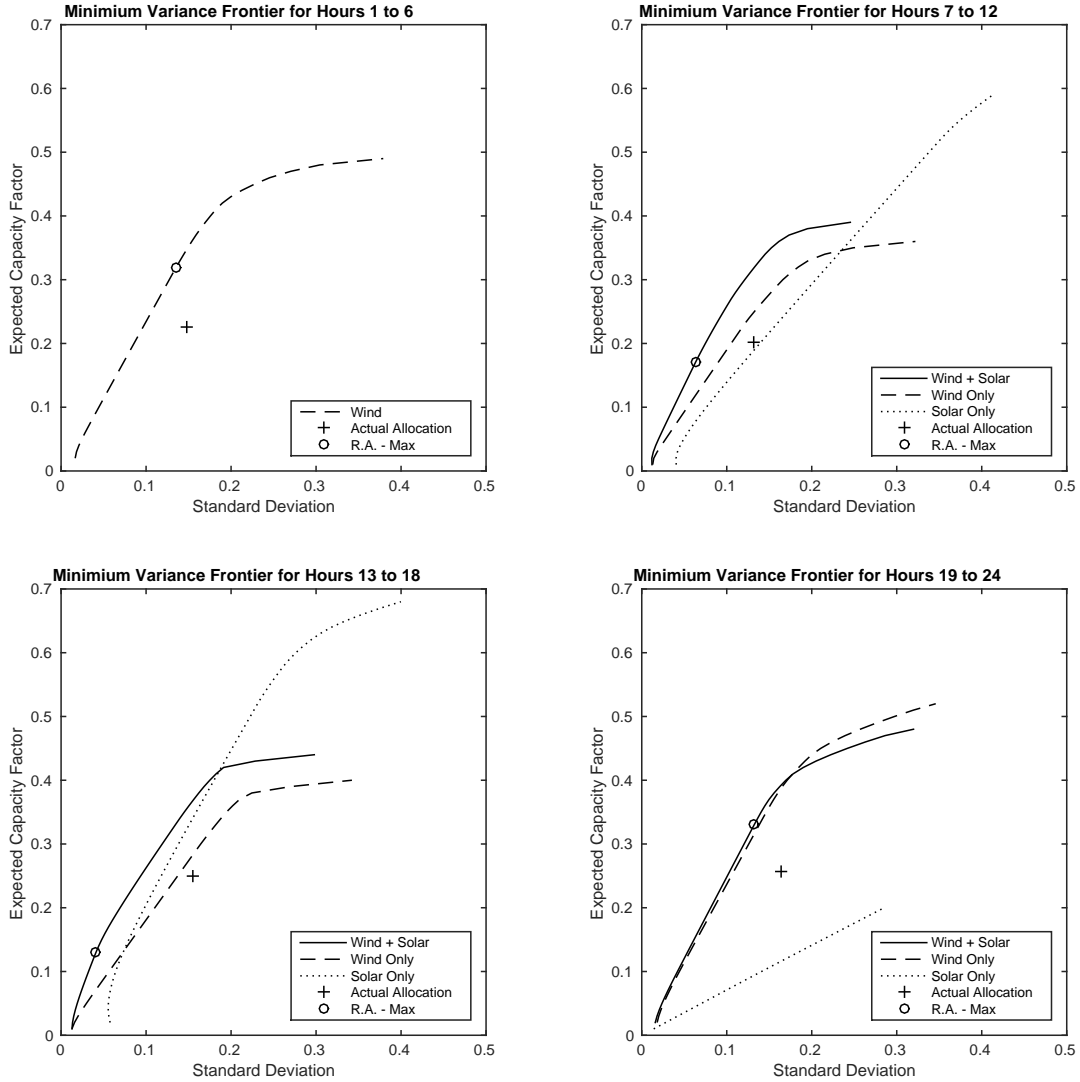


Figure 8: Expected Hourly Revenues per MW and Standard Deviation of Hourly Revenues per MW Efficient Frontier

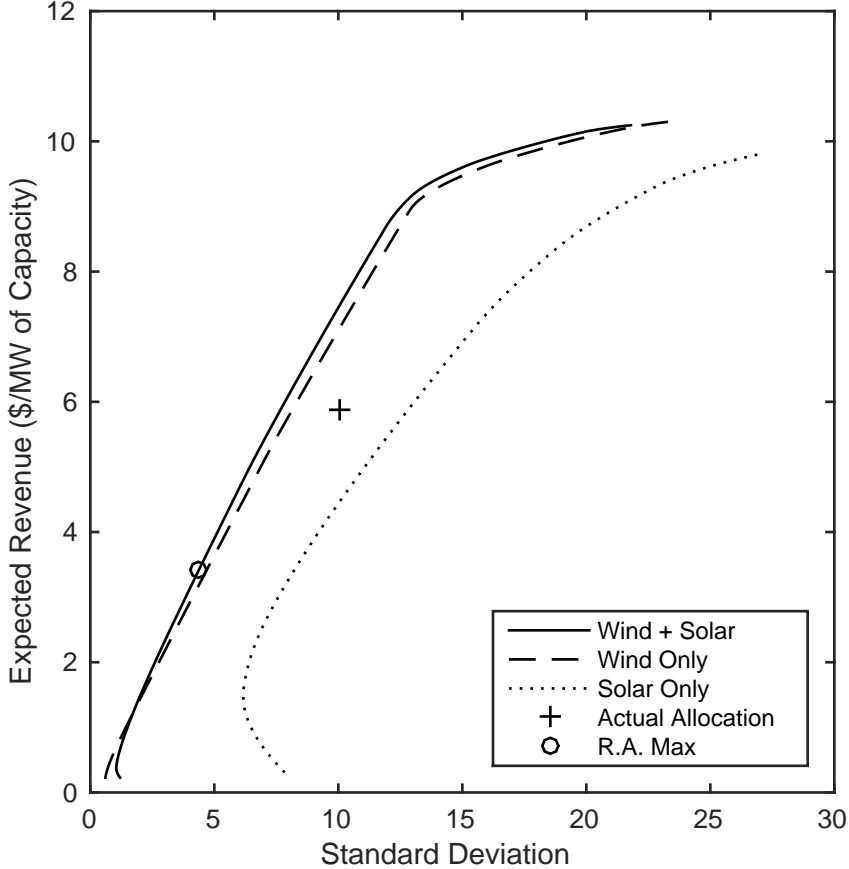


Figure 9: Expected Hourly Revenues per MW and Standard Deviation of Hourly Revenues per MW Efficient Frontier by Groups of Hours of the Day

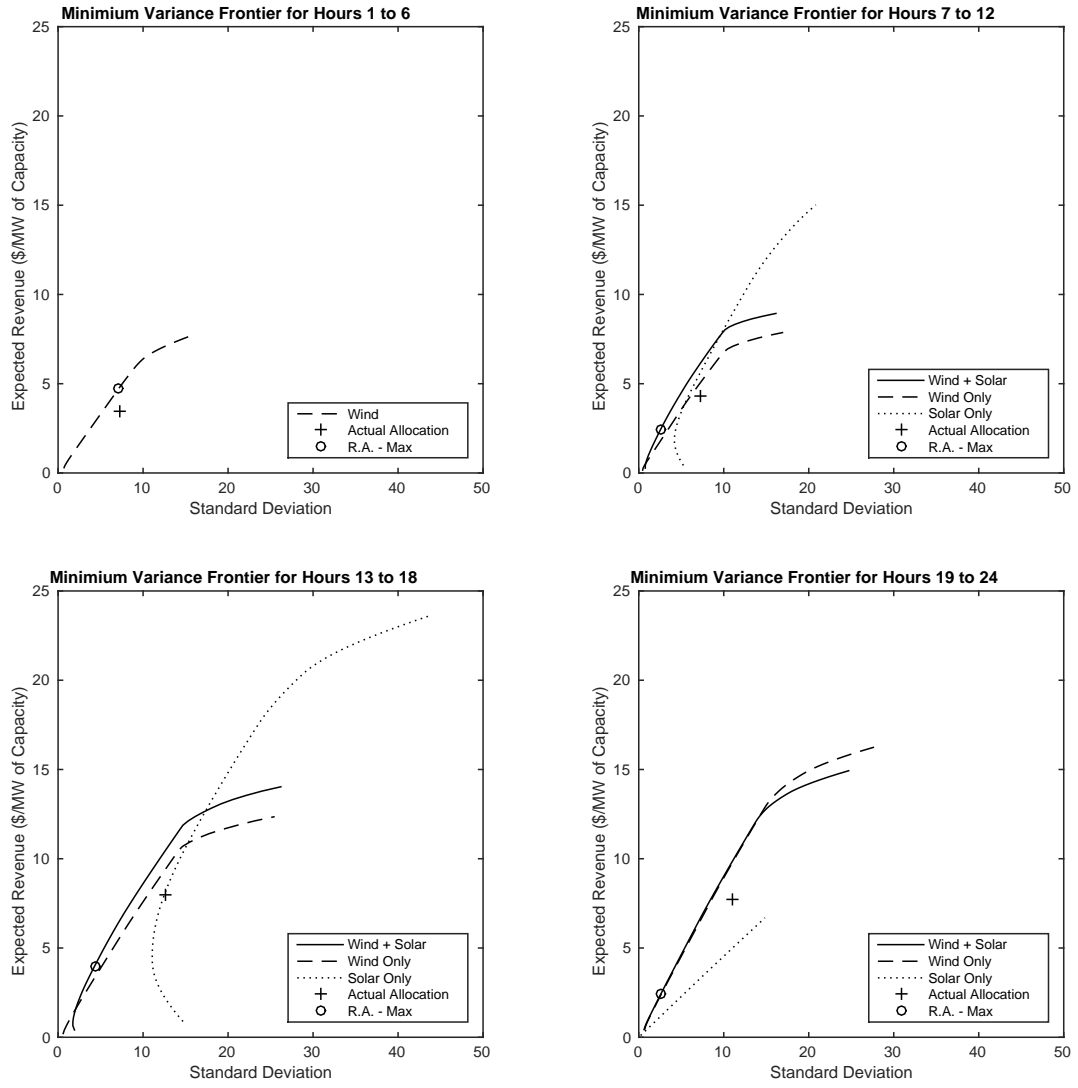


Figure 10: Wind and Solar Hourly Output Portfolio Weights Ordered from Largest to Smallest RA-Max Portfolio Weights

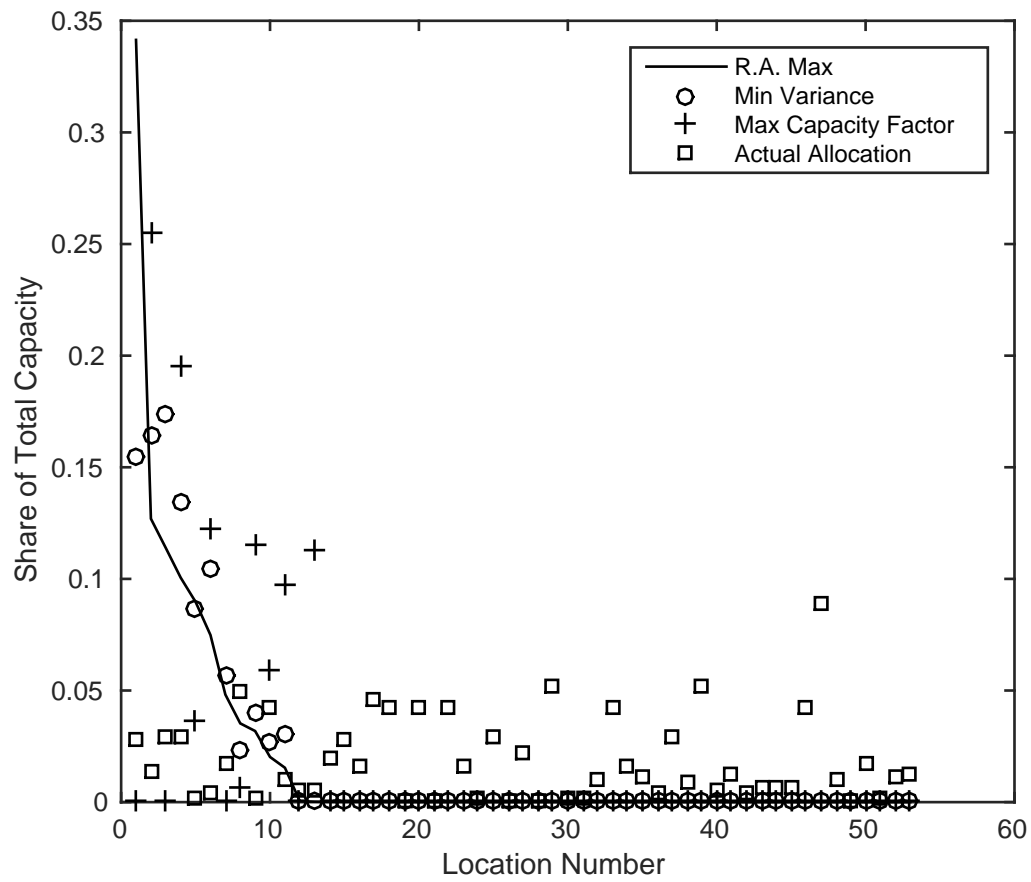


Figure 11: Wind and Solar Hourly Revenue Portfolio Weights Ordered from Largest to Smallest RA-Max Portfolio Weights

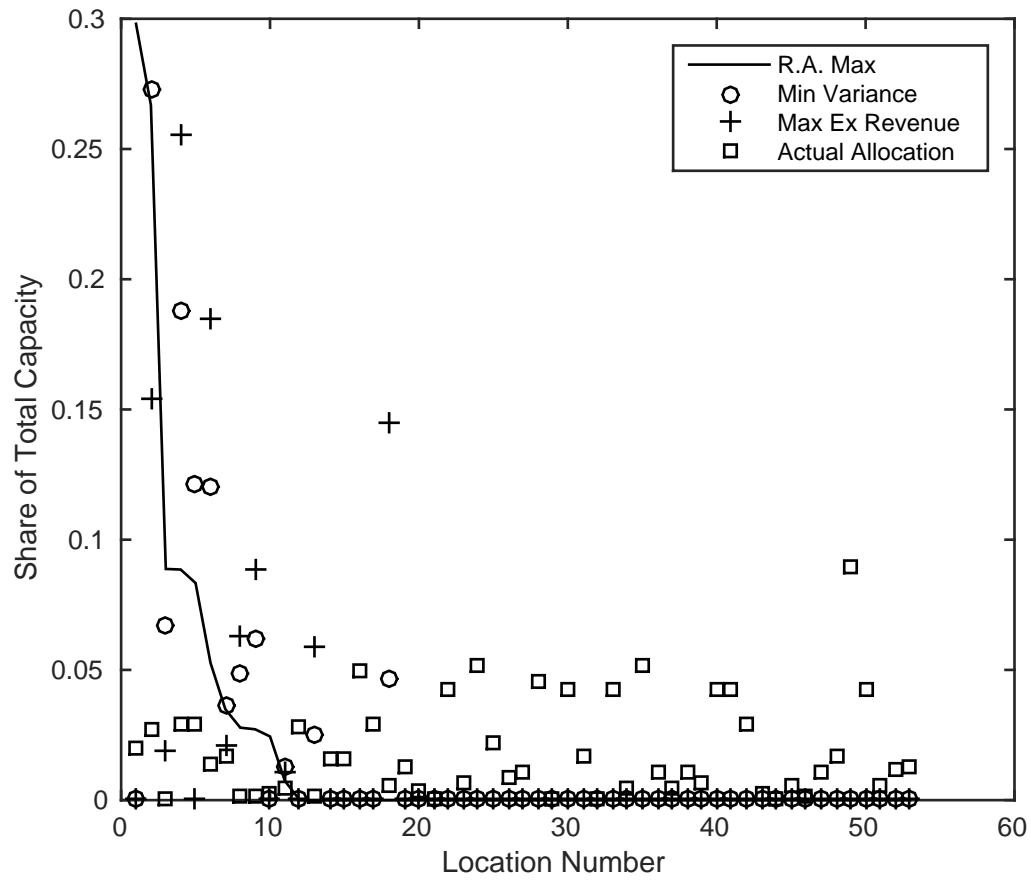


Figure 12: Comparison RA-Max Output versus RA-Max Revenue Portfolio Weights Ordered by RA-Max Output Weights

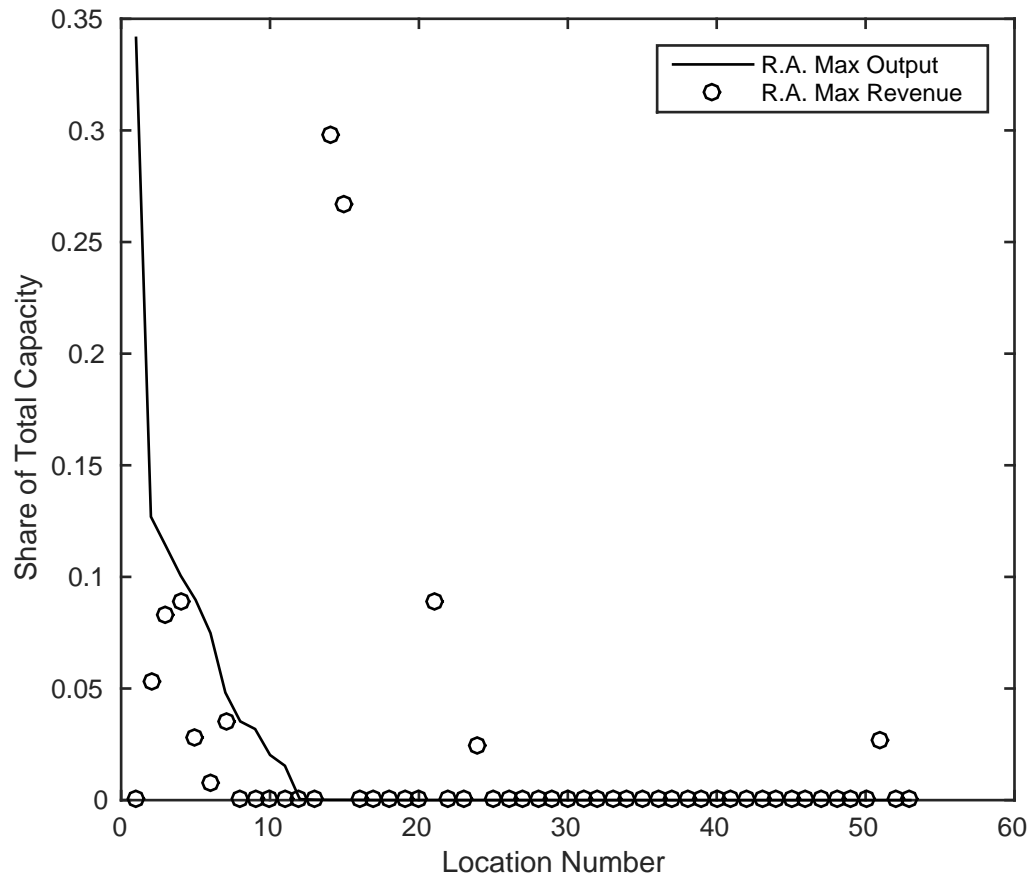


Figure 13: Market-Specific Output Risk and Expected Capacity Factor and Market-Specific Revenue Risk and Expected Revenue

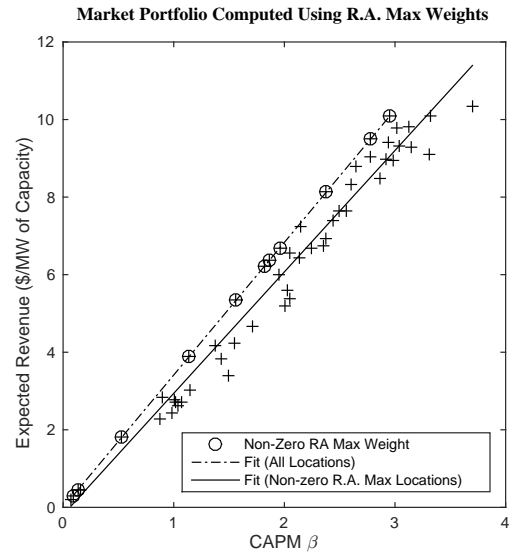
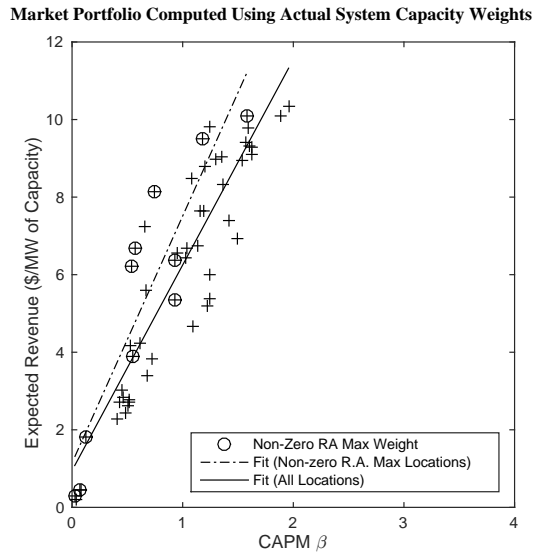
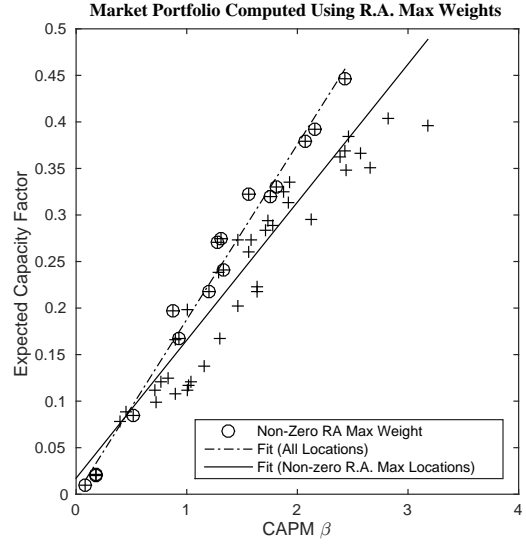
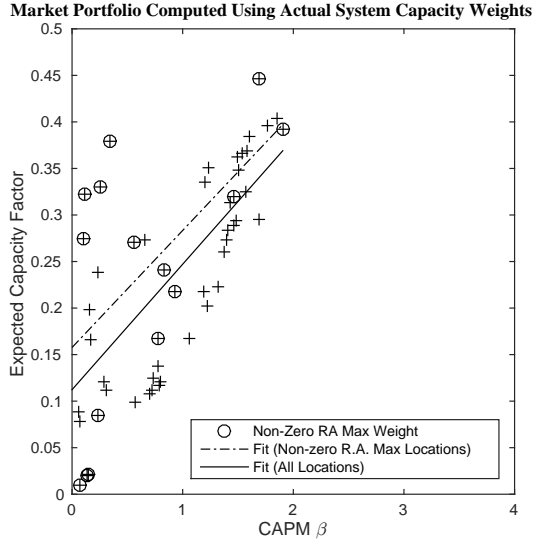


Figure 14: $\% \Delta E(X) / \% \Delta(\sigma(X))$ for X=Output, Revenue - All Locations

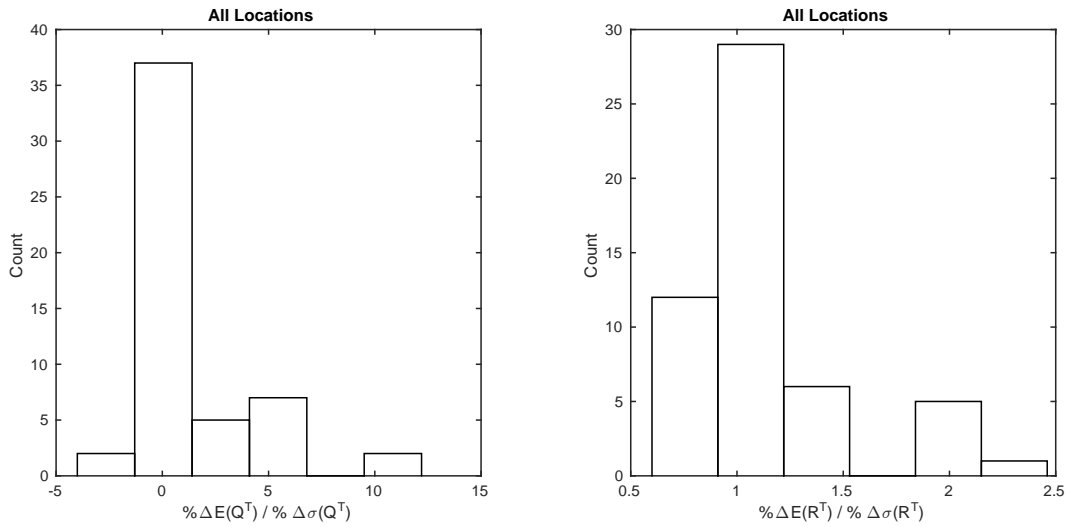


Figure 15: $\% \Delta E(X) / \% \Delta(\sigma(X))$ for X=Output, Revenue - Wind Locations

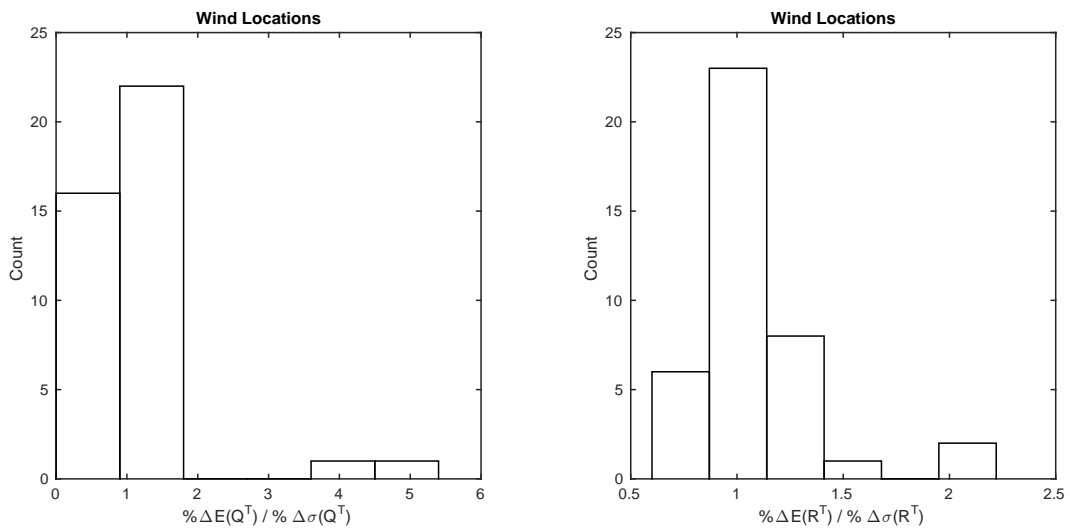


Figure 16: $\% \Delta E(X) / \% \Delta(\sigma(X))$ for X=Output, Revenue - Solar Locations

