The Other Renewable: Hydropower Upgrades and Renewable Portfolio Standards

Stein-Erik Fleten[*], Johannes Mauritzen[†], Carl J. Ullrich[‡]

April 14, 2017

Abstract

A total of 29 U.S. states and the District of Columbia have in place mandatory Renewable Portfolio Standards (RPS) which require that a minimum amount of energy come from renewable resources. We investigate the role of hydropower vis-a-vis other renewables under RPS. Using a Bayesian multilevel model, we find that hydropower plants subject to RPS are more likely to plan upgrades. These planned upgrades appear to be a substitute for solar and wind rather than complementary reserve generation.

Keywords: Renewable portfolio standards, Hydropower, Bayesian methods, Multilevel models, Planned investment.
JEL classification: C11, D92, G18, G31, L94

1 Introduction

In this paper, we investigate the effect of renewable portfolio standards (RPS) on planned investments in hydropower upgrades. We also ask whether hydropower is

[*]Norwegian University of Science and Technology, NO–7491, Trondheim Norway, Contact: stein-erik.fleten@iot.ntnu.no.

[†]Norwegian University of Science and Technology and BI Norwegian Business School, Trondheim, Norway, Contact: johannes.mauritzen@bi.no.

[‡]James Madison University, Department of Finance and Business Law, Harrisonburg, VA, 22807, Contact: ullriccj@jmu.edu.

a substitute or complement to other renewables. In the United States, RPS are the primary state-level policy instruments for encouraging investment in renewable energy generation. RPS are targets set by individual states for the amount of electrical energy which must come from renewable sources by some specified date. Currently 29 states and the District of Columbia have in place mandatory RPS. A further eight states have voluntary RPS.

In many states, RPS targets are aimed at promoting increased solar and wind power.[1] Large penetration of solar and wind power presents challenges to system reliability because these resources are intermittent, i.e., system operators are unable to directly control their production. Because electricity is non-storable, system operators must maintain reserve generation capacity – often in the form of gas-fired combustion turbines – to meet demand when intermittent technologies are unavailable. Hydropower is both renewable and, in some cases, controllable.

Hydropower plants vary by technology type. Hydropower capacity that has a river as a direct water source – run-of-river hydropower – lacks a storage reservoir and in general has limited ability to adjust production. Similar to solar and wind technologies, run-of-river hydropower capacity is not controllable by system operators.

Other hydropower plants store water in a reservoir. A hydropower plant with a reservoir can quickly and at low cost adjust production levels and therefore provide reserve generation. Some reservoir facilities also have pumped storage capabilities,

[1] For example, RPS in Delaware, Washington D.C., Maryland, New Jersey, and Ohio specify that a portion of RPS energy must come from solar power. Illinois requires that a portion of RPS generation come from wind power.

i.e., the ability to pump water uphill from a low-lying reservoir to a higher-lying reservoir. A significant advantage of pumped storage facilities is that water can be pumped into the reservoir during off peak periods, when demand is low and generating capacity is plentiful. The water can be released from the reservoir during peak periods when demand is high and capacity is in short supply. Hence hydropower has the potential to help states meet their RPS goals while overcoming the intermittency problem.

Many states place restrictions on the RPS eligibility of new hydropower facilities. The construction of new dams can pose a threat to water quality and fish passage. According to U.S. Department of Energy [18] upgrading existing plants is a primary source of new hydropower capacity in the United States.

We study the interaction of RPS, planned upgrades to hydropower facilities, and existing renewable solar and wind capacity. We find that while hydropower plants affected by RPS have higher probabilities of upgrades, the upgrade probability is inversely related to the amount of existing intermittent capacity. In other words, upgrades to hydropower facilities tend to be substitutes for intermittent generation, rather than complementary sources of reserve generation.

Our data cover planned upgrades to hydropower plants in the continental United States from 2012 through 2014. We control for substantial variation in geography, technology type, and policy environments using a Bayesian multilevel model with Markov chain Monte Carlo (MCMC) techniques. This method allows us to model naturally the hierarchical structure of the data and obtain reliable inference. To our knowledge, this is the first study to make use of a Bayesian multilevel model estimated by MCMC techniques in the field of energy finance.

Estimating the relationship between RPS policies and investment is subject to endogeneity issues that may bias coefficient estimates. Environmental and other regulations vary from state to state and can impact the timing and realization of investment decisions. These regulations could plausibly be correlated with RPS laws. Unobserved state-specific variables could be correlated with both our dependent investment indicator as well as our policy variable of RPS-status.

Our model and estimation address these concerns. The use of planned[2] – rather than realized – investment addresses causality concerns by clearly establishing the timing. We use existing solar and wind power capacity as our independent variable, but planned upgrades as our dependent variable. Planned investment avoids the influence of market, policy, regulatory, and/or construction-related issues which arise subsequent to the investment decision.

The multilevel structure of our Bayesian model allows us to control for unobserved variables while still estimating the effect of RPS status. We allow for idiosyncratic variation between 234 Transmission Areas (TAs) in whose service areas the hydropower plants are located, by estimating random effect coefficients for each entity.[3] The partial pooling property of the multilevel model and the corresponding estimation of meta-parameters have the dual role of avoiding undue influence of outliers as well as providing inference from TA status, which would otherwise lead to

[2]A planned upgrade is costly in that it involves permitting, environmental studies, and engineering work.

[3]Our Transmission Area (TA) corresponds to "Transmission or Distribution System Owner" in EIA860. See for example row 41 of the Field Directory tab in the LayoutY2013 spreadsheet available from EIA.

perfect collinearity with the TA coefficients in a traditional fixed-effects model.

2 Literature Review

Given the overlapping character of federal, state, and local energy policy in the U.S., policy comparison and interaction has been an active area of research. The focus has been on the effects of energy policies on prices and generation costs (Palmer and Burtraw [16], Wiser and Bolinger [19], and Fischer [4]). Others investigate the interaction of RPS with other energy or climate policies (Bird, Chapman, Logan, Sumner, and Short [2], Tsao, Campbell, and Chen [17], Amundsen and Mortensen [1], Böhringer and Rosendahl [3], Linares, Santos, and Ventosa [11]).

While the literature is weighted heavily towards large-scale simulation studies, a few econometric studies of the effectiveness of RPS appear. Menz and Vachon [15] use a cross-sectional analysis and find a positive correlation between RPS policies and renewable energy generation. However, the combination of a cross-sectional design and the use of a realized investment variable leads to difficulties in establishing causality. Yin and Powers [20] take into account variation between years and states by employing a panel data set and an econometric model with year and state fixed effects. They find that RPS laws have had a strong positive effect on renewable energy development.

Both Menz and Vachon [15] and Yin and Powers [20] use annual data on the realized percentage of renewables in state. They do not take into account the effect that RPS laws have on development in neighboring states. In many cases, renewable energy imported from a neighboring state are eligible under RPS. We include renewable generation from neighboring states.

Linnerud, Andersson, and Fleten [12] analyze investment decisions for small hydropower plants in Norway under a green certificates program. Green and Vasilakos [9] and Mauritzen [14] show how hydropower in Norway and Sweden helps balance intermittent wind power production in Denmark.

To the best of our knowledge, studies of how RPS policies affect hydropower investment are absent from the literature. We aim to fill this gap.

3 Data

Table 1 gives a summary of RPS laws across states. Many states have restrictions on new hydropower development – especially those built with dams. Other restrictions on RPS generation include the size of hydropower plants, the location of the plants, and the age of the plants. Massachusetts, Montana, New Hampshire, New York, North Dakota, Oregon, Utah, and Washington all explicitly allow generation from upgrades to existing hydropower plants to count towards meeting RPS goals. We can find no state whose RPS rules explicitly exclude hydropower upgrades. For this reason, we focus on upgrades – defined below – to existing hydropower facilities.

Our main source of data is Form 860 from the U.S. Energy Information Administration (EIA). Form 860 includes information about all power plants; annual filing is required. We analyse all hydropower plants, thus avoiding a possible selection bias toward plants that consider responding to RPS incentives. The form 860 data also

Table 1: Renewable Portfolio Standards (RPS) vary substantially across states, ranging from approximately 5% in Texas to 40% in Maine. The RPS standards often vary in specifics, requiring differing targets for investor-owned (IOU) and versus public utilities. Footnotes on these details are provided when appropriate. (MW = megawatts)

	State	Total RPS	RPS Goal Year
Mandatory	Arizona	15%	2025
	California	33%	2020
	Colorado	30%	2020
	Connecticut	27%	2020
	DC	20%	2020
	Delaware	25%	2026
	Hawaii	40%	2030
	Illinois	25%	2025
	Iowa	105 MW	NA
	Kansas	20%	2020
	Maine	40%	2016
	Maryland	20%	2022
	Massachusetts	21%	2020
	Michigan	10%[a]	2015
	Minnesota	25%[b]	2025
	Missouri	15%	2021
	Montana	15%	2015
	Nevada	25%	2025
	New Hampshire	24.8%[c]	2025
	New Jersey	20%	2020
	New Mexico	20%	2020
	New York	29%	2015
	North Carolina	12.5%	2021
	Ohio	12.5%	2024
	Oregon	25%[d]	2025
	Pennslyvania	18%	2021
	Rhode Island	16%	2019
	Texas	5,880 MW	2015
	Washington	15%	2020
	Wisconsin	10%	2015
Voluntary	Indiana	10%	2025
	North Dakota	10%	2015
	Oklahoma	15%	2015
	South Dakota	10%	2015
	Utah	20%	2025
	Vermont	20%	2017
	Virginia	15%	2025
	West Virginia	25%	2020

[a] And 1100 MW.
[b] Xcel Energy will be required to generate 31.5% of its power from renewables by 2020.
[c] Hydropower specific goal of 1.5% by 2015.
[d] Applies only to large utilities, small utilities must meet 5–10%.

contains proposed changes to existing plants. We use planned upgrades to existing hydropower plants for 2012–2014. We consider three types of planned plant upgrades in our analysis.

- Uprate: The capacity of a plant is increased by upgrading an existing generator.
- Repowering: An existing generator is replaced with a new generator.
- Generator addition: The capacity of a plant is increased by investing in a new generator.

In our main analysis, we do not distinguish between these three outcomes, but instead aggregate them into a binary variable which equals one if a given plant in a given year was planning an upgrade and zero otherwise. A complementary variable exists that provides the size of the uprate, but these data are only available for a small subsample of planned uprates. Capacity data are not available for repowering and generator additions.[4]

Planned upgrades are distributed broadly across the United States, as shown in Figure 1. Table 2 provides summary statistics. The data include a total of 1424 hydropower plants, 87% of which are subject to RPS laws. In total we have 4199 hydropower plant-year observations, 123 (approximately 3%) of which indicated planned investment. The plants are distributed across 234 TAs. The average penetration of intermittent (solar and wind) power for these TAs was 4.3%, with a range of 0–89%.

[4]Section C in the Appendix presents the results of an analysis which uses the size of the proposed upgrade as the dependent variable rather than a simple indicator.

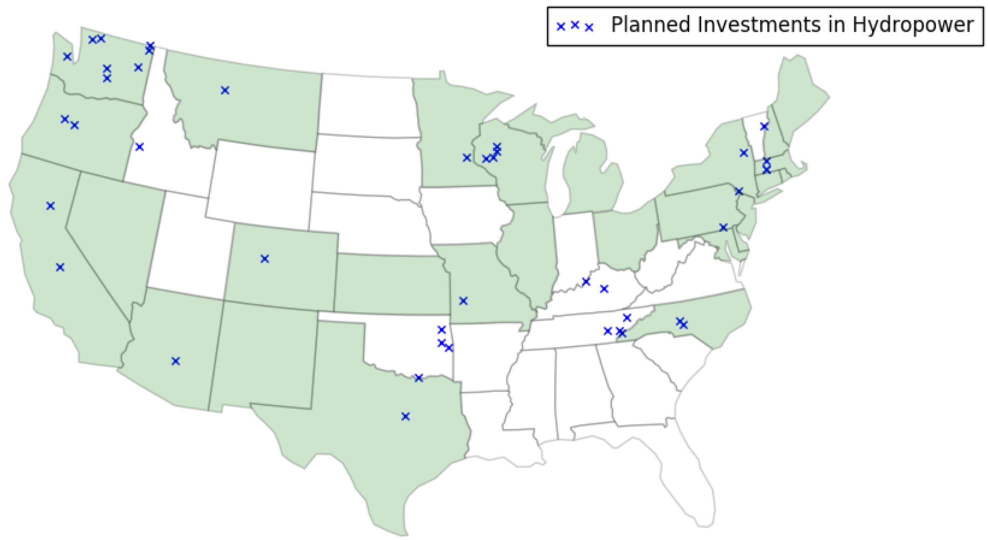

Figure 1: Geographic distribution of planned capacity investments in existing hydropower plants as reported in the 2012–2014 EIA Form 860. The states colored green represent those that have RPS targets.

To measure the size of plants, we use total nameplate capacity, which is the official maximum rated power output of the plant. Most of the hydropower plants are medium sized, with an average nameplate capacity of 73 MW. Approximately 20% are smaller than 2 MW, while about 2.5% are larger than 600 MW.

Washington, Oregon, and California have the largest amount of capacity, though significant amounts of capacity can also be found in New York and the southeast. The upper panel of Figure 2 displays the geographic distribution of hydropower plants. The size of the blue circles is in proportion to the size of the hydropower plants. The lower panel of Figure 2 shows the geographic distribution of solar and wind power plants. Most of the solar power distribution is found on the coasts. Wind power is

Table 2: Summary Statistics of Hydropower Plants. (MW = megawatts)

Total # observations	4199
% of observations with planned upgrade	2.9%
Total # plants	1424
Average age	58 years
Average nameplate capacity	73 MW
Utility owned plants	62%
Plants covered by RPS laws	87%
Plants with reservoir	12%
Plants that have pumped storage	2%
# of TA's	234
Average intermittent % of TA's generation	4.3%

most heavily concentrated in the central United States - from Texas through Iowa and Minnesota, as well as the western coastal states.

4 Logit Regression Analysis

We begin by conducting a simple logit regression analysis with the indicator variable for a planned upgrade as the dependent variable. The independent variables include an indicator for TAs which operate in RPS states, existing wind and solar capacity as a the percentage of TA total capacity, and five variables to control for characteristics of the plants. The regression specification follows.

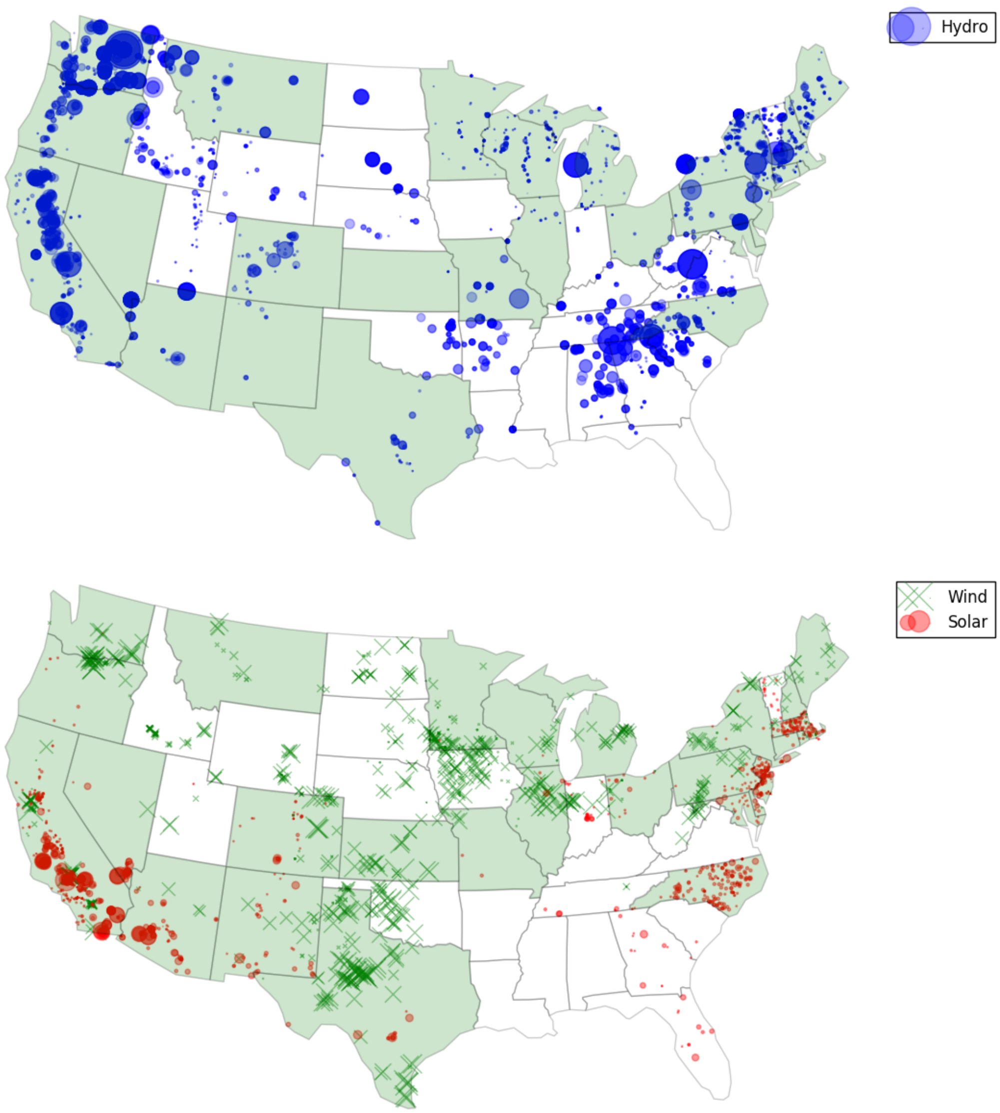

Figure 2: The top panel shows the geographic distribution of hydropower plants; the size of the blue circles is proportional the size of the hydropower plants. The bottom panel shows the distribution of solar and wind plants. The states with green background indicate the existence of RPS laws.

$$
\begin{aligned}
investment_{y,p} &= \alpha + (\beta_1 * rpsto_{y,j}) + (\beta_2 * intermittent_{y,j}) \\
&\quad +(\beta_3 * age_{y,p}) + (\beta_4 * irsrvr_p) + (\beta_5 * nplate_p) \\
&\quad +(\beta_6 * ips_p) + (\beta_7 * iutil_p), \qquad (1)
\end{aligned}
$$

where

$investment_{y,p}$ is an indicator variable which is equal to one if plant p had a planned upgrade in year y and zero otherwise,

$rptso_{y,j}$ is an indicator variable which is equal to one if TA j (in which plant p in located) operates in states with RPS in place in year y and zero otherwise,

$intermittent_{y,j}$ is the percent of total capacity in TA j in year y which is intermittent in nature,

$age_{y,p}$ is the age in years of plant p in year y,

$irsrvr_p$ is an indicator variable which is equal to one if plant p includes a reservoir and zero otherwise,

$nplate_p$ is the nameplate capacity of plant p, in MW,

ips_p is an indicator variable which is equal to one if plant p utilizes pumped storage and zero otherwise, and,

$iutil_p$ is an indicator variable which is equal to one if plant p is owned by a vertically integrated utility and zero otherwise.

We divide the results in Table 3 into three sections. The first section of results, labeled *Base*, presents the results for the specification in equation (1), which we label regression I. There are two main takeaways.

First, the *rpsto* coefficient is positive and significant at the 5% level, suggesting that upgrades are more likely for hydropower plants located in TAs which operate in RPS states. Second, the *intermittent* coefficient is negative and significant at the 1% level, suggesting that hydropower is a substitute for solar and wind, not a complement.

In the *Individual Interactions* section we interact the indicator for TAs which operate in RPS states with each of the other independent variables. The results are inconsistent. From regression II, when we interact *rpsto* with *intermittent*, we find no effect for either the stand alone *intermittent* coefficient or the interaction term.

In the Base results the *isutil* coefficient is positive and significant at the 1% level, indicating that plants owned by utilities are more likely to be upgraded. From regression VII, when we interact the RPS indicator with the utility indicator, neither the stand alone *iutil* coefficient nor the interaction term are significant. Notice also that the *rpsto* coefficient is not significant in this regression. Similar issues arise in the *Full* results, regression VII.

4.1 Hierarchical structure of the data

The inconsistencies discussed above may be because the simple logit regression does not account for the hierarchical (aka, multilevel) structure of the data. Our data are structured in four levels.

(i) For every plant, p, we have yearly observations, y, for 2012, 2013, and 2014.

(ii) Each plant has its own characteristics, e.g., size, age, hydropower type, and corporate ownership structure.

(iii) The plants are grouped by service area via their Transmission Area. Some TAs are vertically integrated utilities that also own power plants, while others only own the transmission infrastructure.

(iv) Finally, we group TA's based on those which operate in states with renewable portfolio standards and those which do not.

Accounting for the hierarchical structure of the data is important because observations from the same level, or cluster (e.g., from the same TA), are likely to be correlated. Ignoring this clustering can lead to unreliable inference.[5]

Multilevel mixed-effects logistic models which include random intercepts and random coefficients deal with exactly this kind of problem. However, our attempts to fit traditional multilevel models[6] would not converge due to the large number of plants and TAs. This leads us naturally to the use of Bayesian methods.

[5] One way to handle such clustering is using fixed effects, e.g., including an indicator variable for a particular TA. However, in this case it is not possible to separate the effects of observed and unobserved TA level characteristics. This problem is particularly important in our case as the effects of the observed variables are interesting in their own right.

[6] We used the ***meqrlogit*** command in STATA.

5 Bayesian Multilevel Model

We fit a multilevel (or hierarchical) regression model under a Bayesian framework, using Markov chain Monte Carlo (MCMC) simulation techniques. Bayesian models only recently have begun to make inroads in the economics and finance literature. In large part this is because the software and computing power required to estimate sufficiently realistic models only recently have become available. The main benefit of a Bayesian multilevel model is that it can accommodate many hundreds of parameters without overfitting.

In Bayesian analysis the model parameters themselves are considered to be uncertain and are defined by probability distributions rather than point estimates, with uncertainty estimates derived from sampling assumptions. Each parameter is assigned an initial prior distribution. In practice, prior distributions are often set to be non-informative or weakly informative – placing most of the initial probability density on a range of sensible magnitudes. The purpose of defining a weakly informative prior is to aid the MCMC software in converging by limiting the probability space the algorithm searches.

Once defined, the model is updated by the likelihood function derived from the data. A Bayesian model is then a weighted average of the joint prior distribution and joint likelihood function of the model, with the weights defined by the amount of data available.

5.1 Meta-parameters and partial pooling

In a multilevel model, lower level parameters within a certain group are derived from a distribution characterized by meta-parameters. The estimated group-level parameters are a weighted average of both the observations within the group as well as the full set of observations in the dataset – a feature of multilevel models called partial pooling.[7] Partial pooling also serves as a natural form of parameter shrinkage, which mostly eliminates the need for using corrections for multiple comparisons in inference [7].

The partial pooling feature of hierarchical Bayesian models is particularly useful for our dataset with relatively few positive responses (planned upgrades) spread among a large number of categories. Consider, for example, a planned upgrade of a hydropower plant in an area (TA) with few other hydropower plants. A traditional cross-sectional model might erroneously suggest that plants in this area are particularly likely to be upgraded and therefore bias the result. With partial pooling, all the local probabilities of upgrades are pulled towards the national average, helping to avoid such a bias.

[7] For intuition, consider the case of an epidemiological study of a rare disease across all counties of the US. The average incidence of the disease by county will likely show that small rural counties have the highest incidence of the disease, not because of any causal relationship, but simply because by pure chance, some small counties with several hundred or thousand residents will have one or two cases of the disease. A hierarchical model with partial pooling will correct the bias of these statistical outliers by pulling the county-averages towards the national average of the disease. [13]

5.2 Intuition and description of the Bayesian model

With a Bayesian multilevel models, the researcher has the flexibility in specifying the model without relying on assumptions such as constant group-level variances and Gaussian noise distributions. Because Bayesian simulation techniques result in an estimate of the full joint probability distribution of the model, the inference has the potential to be more informative than the typical point estimates and p-values of the standard hypothesis testing frameworks [10].

5.2.1 Likelihood and fitted response data

Let $\mathcal{L}(data|\theta)$ be the likelihood, where θ represents a vector of parameters to be estimated and *data* represent all the data available for estimation. MCMC is then used to sample θ from its posterior distribution,

$$\mathcal{P}(\theta|data) \propto \mathcal{L}(data|\theta) \times \pi(\theta) \tag{2}$$

where $\pi(\theta)$ is the prior distribution of the data.

The likelihood of the response data - whether or not an upgrade is planned in a certain year, y, for a given plant, p – denoted $investment_{y,p}$ – is modeled as a Bernoulli random variable transformed to the unbounded logit scale.

$$investment_{y,p} \sim bernoulli(logit^{-1((\hat{Y}_{y,p}))}) \tag{3}$$

In turn, the fitted data, $\hat{Y}_{y,p}$, is assumed to be comprised of two sources of variation. Reporting-year random effects, b_y^{ry}, capture year-to-year variation that will affect all plants reporting in a given year. In addition to the reporting-year random effects, we also model the fitted data as a function of plant-level observable effects, b_p^{obs}, a coefficient which we interpret as a score indicating how likely plant p is to have a planned upgrade.

5.2.2 Plant level scores

$$b_p^{obs} = A_{ta} + \mathbf{B_{ta}^T} \times \mathbf{X_p}. \tag{4}$$

The plant level scores are modeled as a vector of K plant-level variables, $\mathbf{X_p}$. The matrix of coefficients on the K plant-level variables, $\mathbf{B_{ta}}$, as well as the plant-level intercept, A_{ta}, are allowed to vary by TA.

The inclusion of the TA-level groupings and associated random effects is important in establishing identification of the effect of the RPS-policy and the role of existing intermittent capacity. Consider, for example, a TA in a region with an unobserved variable that affects both the propensity for a state to have a RPS-law and the likelihood of a planned hydropower upgrade. This could, for example, be an industrial presence in the region that leads to both more investment in hydropower upgrades as well as political pressure for increased renewable energy. Such idiosyncratic variation is controlled for by the TA random effects while still allowing for the estimation of an average treatment effect of RPS-policy.

5.2.3 Meta-parameters

There are 234 vectors $\mathbf{B_{ta}}$, one for each TA, and each contains $K = 5$ coefficients. Thus there are a total of 1170 coefficients to be estimated.[8] Our Bayesian multilevel model manages the large number of coefficient estimates by assigning them higher-level meta-parameters. These meta-parameters also have an important economic interpretations as the contrast between plants in RPS and non-RPS TAs.

In particular, we define the meta-parameters as in (5).

$$
\begin{aligned}
A_{ta} &= \alpha^{A}_{rps} + \beta^{A}_{rps} \times intermittent_{ta} + A^{re}_{ta} \\
\mathbf{B_{ta}} &= \alpha^{\mathbf{B}}_{\mathbf{rps}} + \beta^{\mathbf{B}}_{\mathbf{rps,k}} \times \mathbf{intermittent_{ta}},
\end{aligned} \tag{5}
$$

where the prior distributions on the coefficients can be written as in (6).[9]

$$
\begin{aligned}
\alpha_{rps} &\sim Cauchy(0, 5) \\
\beta_{rps,k} &\sim Cauchy(0, 5) \\
A^{re}_{ta} &\sim Cauchy(0, 5)
\end{aligned} \tag{6}
$$

[8] The plant-level variables are the same as in the logistic regression from Section 4, (1) whether or not the plant uses a reservoir as a water source, (2) whether or not the plant has pumped storage facilities (a pumped storage facility by definition has at least two reservoirs, thus plants are a subset of the hydropower), (3) whether or not the plant is owned by a vertically integrated utility, (4) total nameplate capacity (size), and (5) operating life (age). Note here that the plant-level variables do not change over time for a given plant, thus we estimate one vector of coefficients $\mathbf{B_{ta}}$ for each transmission area (234) and not each plant (1424).

[9] Both the meta-parameters for A_{ta} and $\mathbf{B_{ta}}$ are given the same weakly informative priors.

We follow Gelman, Jakulin, Pittau, and Su [8] and Gelman [5] in using weakly informative Cauchy prior distributions on many of the parameters. The higher level random effects and coefficient terms are assigned Cauchy priors with location zero and estimated scale terms σ. The σ terms are assigned half-Cauchy prior distributions with location parameter[10] of 0 and scale parameter of 2.5.[11]

The TA-level intercept terms, A_{ta} are modeled as a combination of an average mean effect which is allowed to vary based on whether the TA is in an RPS state or not, α^A_{rps}, and an interaction effect, β^A_{rps}, with the percent of intermittent capacity in the TA, $intermittent_{ta}$, which is also allowed to vary based on the RPS-status of the TA. Finally, TA-level random effects, A^{re}_{ta}, are estimated for each TA to capture idiosyncratic TA-level variation. The difference between the RPS and non-RPS estimates of *alpha*: $\alpha_{rps} - \alpha_{not-rps}$ gives the estimated average treatment effect of RPS on the probability of investing in a hydropower plant.

5.2.4 Interpretation: Meta-parameters contrasts

The coefficients β^A_{rps} represent the main effects of the intermittent variable by RPS-status. The contrasts, $\beta^A_{is-rps} - \beta^A_{not-rps}$, represents the average interaction treatment effect of RPS with the intermittent variable. This contrast can be interpreted as the effect of installed capacity of renewable energy on the intention of investing between RPS and non-RPS regimes. A positive contrast would be evidence for a

[10]In a normal distribution, the location would be the mean, for example.

[11]The use of the Cauchy distribution reflects the fact that a large in magnitude effect of around 5 or greater on the logit scale is highly unlikely in logistic regressions in which all non-binary data has been transformed to have mean zero and standard deviation 0.5, as we have done. In addition, the Cauchy distribution will give answers even under complete separation, and avoids computational problems inherent in assigning completely non-informative priors in multilevel models [8].

complimentary relationship between planned hydropower upgrades and intermittent generation. A negative contrast is evidence that planned hydropower upgrades serve as a substitute for intermittent generation.

The TA-level vectors of K plant-level coefficients, $\mathbf{B_{ta}}$, also have meta-parameters that can be given direct economic interpretations. The vectors $\alpha^{\mathbf{B}}_{\mathbf{rps}}$ represent the average main effects of the $K = 5$ plant-level variables. The k-vectors $\beta^{\mathbf{B}}_{\mathbf{rps}}$ represent the interaction effects with the intermittent variable. These meta-parameters are also allowed to vary by RPS-status. Thus the contrast $\alpha^{B}_{is-rps,k} - \alpha^{B}_{not-rps,k}$ can be interpreted as the interaction treatment effects of the plant-level variables with RPS policy. For example, the contrast $\alpha^{B}_{is-rps,reservoir} - \alpha^{B}_{not-rps,reservoir}$ indicates whether the intention to invest is more likely in a hydropower plant with a reservoir under a RPS regime. A summary of the model and hierarchy is presented in Figure 3.[12]

5.2.5 Aggregating responses across years

We use panel data in which not all positive responses – a plant owner announcing a planned investment in a given reporting year – indicates a unique event. A plant owner who announces an intent to invest in 2012 and 2013 is likely referring to the same planned investment. While a naive analysis that includes all the data would give too much weight to a single plant with multiple positive responses, simple solutions like restricting the analysis to a single year would discard data and degrees of freedom and worsen the power of the inference.

[12] We estimated the model with four chains and 1000 iterations, after an initial "warm-up" phase. An "R" -statistic [6] of one indicates that the chains converged to the target posterior distribution. We use the Stan default of extracting 2000 samples from the estimated target posterior distribution.

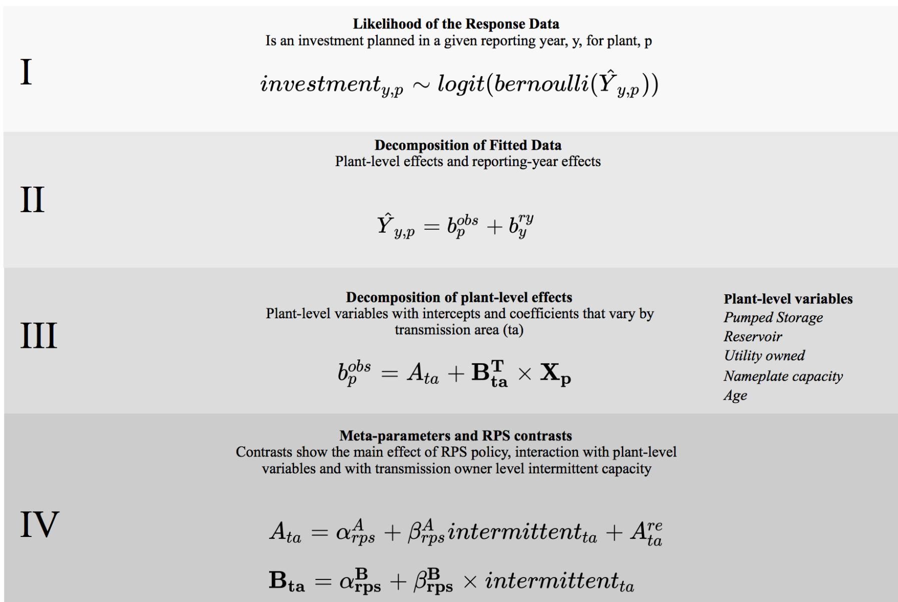

Figure 3: The diagram illustrates the structure of the multilevel model used to estimate the probability of investing in upgrades to hydropower plants. The likelihood of the data is modeled as a logit-transformed bernoulli random variable. The fitted values, $\hat{Y_{y,p}}$, are in turn modeled as a sum of plant-level effects, b_p^{obs}, and reporting year effects b_y^{ry}. The plant-level effects are further modeled as a TA-level intercept terms, A_{ta} and a vector of K plant-level covariates $\mathbf{X_p}$ and a corresponding vector of their estimated coefficients $\mathbf{B_{ta}}$. The TA-level intercept terms and coefficients on plant-level covariates are in turn modeled as a combination of an intercept term, α_{rps}, and an intermittency term with a coefficient β_{rps}. Both α and β are allowed to vary by whether the TA has a presence in a RPS state.

Instead of arbitrarily discarding data, the three yearly observations are grouped together. Each plant is then modeled explicitly as a single instance with three observations. The weight given to each observation is determined endogenously in the estimation of the full probability model.

6 Results

We follow the Bayesian practice of presenting the full marginal posterior distribution for parameter values of special interest in the form of histograms of the samples. Table 4 in the Appendix shows summary statistics of the main parameters. Presenting summary statistics of the thousand-plus TA-level parameters is space-prohibitive. Instead we present visual summaries of TA-level coefficients.

The meta-parameters α^A and β^A have clear and important economic interpretations as the main average treatment effects of RPS-policy and installed intermittent generation, respectively. In the first panel of Figure 4, the estimated distribution for the α^A parameters is displayed as a contrast $\alpha^A_{rps} - \alpha^A_{not-rps}$, which can be interpreted as the treatment effect of RPS-policy on the probability of planning a hydropower upgrade. The results indicate an approximately 90.6% probability that RPS-policy increases the likelihood of a planned upgrade.[13] These results provide evidence that RPS policies are effective in promoting upgrades in existing hydropower facilities.

The estimated distributions on the meta-parameters β^A_{rps} and $\beta^A_{not-RPS}$ provide strong evidence on the main average effects of installed intermittent capacity under

[13]In a Bayesian frameworks, probabilities can be interpreted directly as a degree of belief, as opposed to a hypothesis-testing frameworks where evidence leads to either reject or fail-to-reject conclusions.

RPS and non-RPS regimes. The second panel of Figure 4 shows that under the RPS regime, over 98% of the probability mass is less than zero with a center around -3. This is evidence of that large penetration of intermittent capacity tend to reduce the probability of planned upgrades of hydropower plants in TA's under RPS-regimes. In other words, hydropower upgrades are a substitute for other forms of renewable generation, rather than acting as complementary regulating capacity.

In contrast, in non-RPS regimes, the evidence for an effect of intermittent capacity is substantially weaker. The distribution of the parameter $\beta^{A}_{not-RPS}$, shown in the third panel, is centered around zero.

Figure 5 presents the estimated distributions of the $\alpha^{\mathbf{B}}$ parameters, which represent the main effects on the the plant-level coefficients. For the non-RPS plants, the coefficients are centered around zero. However, for plants under RPS, the nameplate capacity variable is centered around positive 0.5. In other words, under RPS large plants are more likely to plan upgrades. The simple regression in Section 4 was unable to generate these results.

The coefficient on pumped-storage is distributed over mostly positive values, suggesting that under RPS laws, plants with pumped storage are more likely to plan an upgrade. The increased ability to regulate production that pumped storage provides, and resulting increased revenue appears to lead to a higher probability of planned investments. However the estimation of the interaction term with installed capacity of intermittent generation, shown in Figure 6, does not provide evidence that planned investments in pumped storage have a complimentary relationship to investments in wind and solar generation.

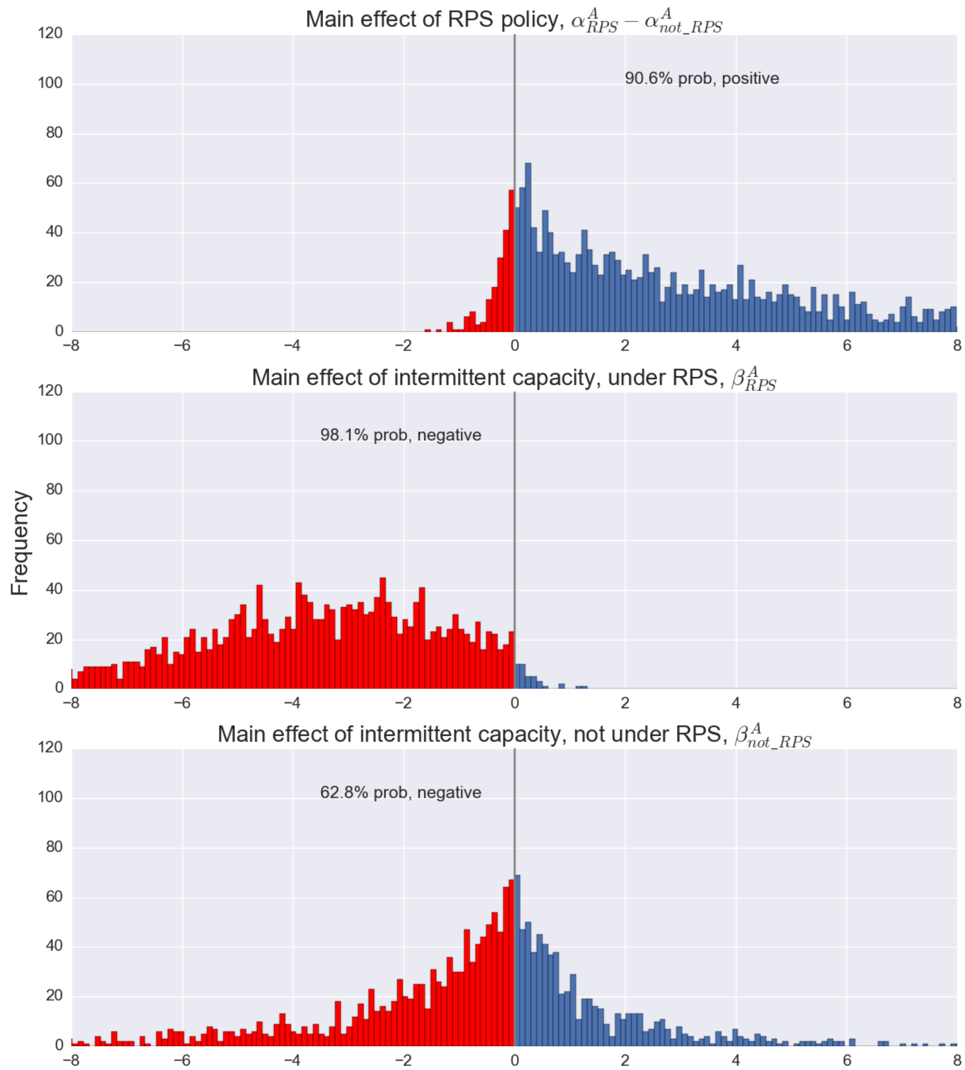

Figure 4: The first panel shows the estimated distributions of the contrast, $\alpha^A_{rps} - \alpha^A_{not-rps}$, which can be interpreted as the main effect of RPS policy. The estimation gives an 90.6% probability that the contrast is positive: That is that RPS policy increases the probability of planned investment. The second and third panels represent the main effects of the TA-level variable *intermittent* in RPS TA's and non-RPS TA's. The estimation indicates a high probability of over 98% that RPS plants are less likely to invest in new hydropower generation if there already exists substantial amounts of solar and wind power.

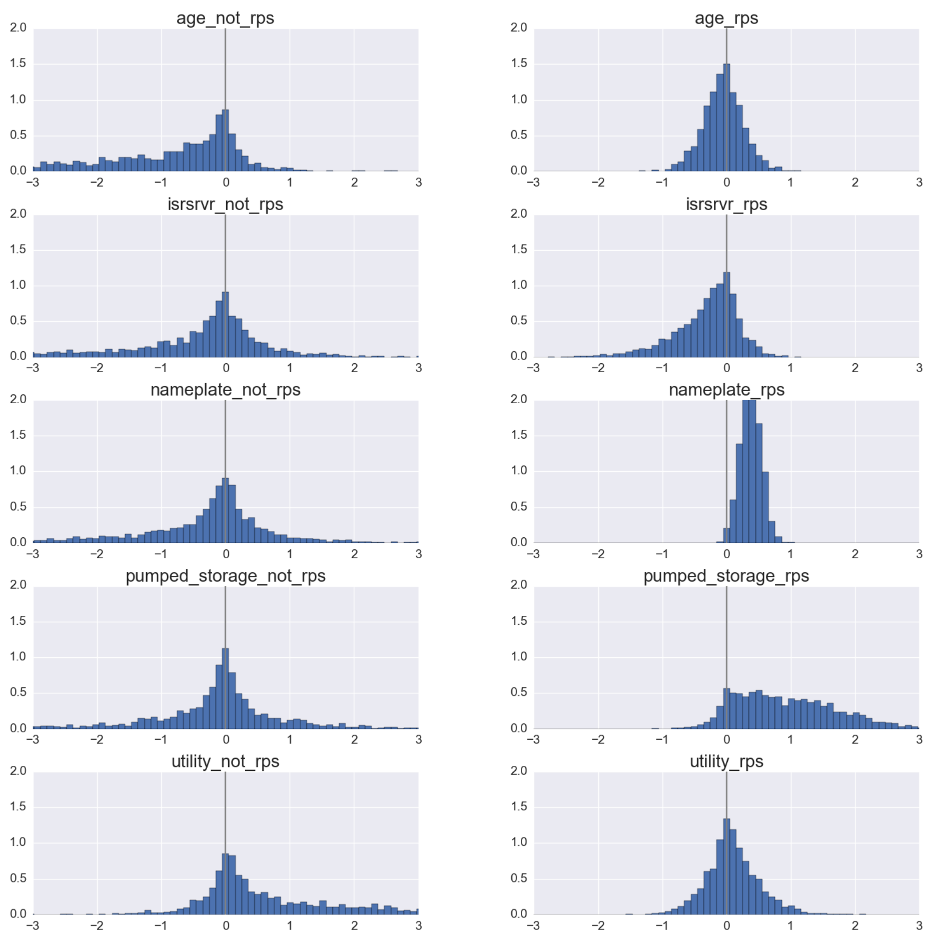

Figure 5: The estimated distributions of the $\alpha^{\mathbf{B}}$ coefficients. These represent the main effects of the plant-level variables in non-RPS and RPS TA's. Larger plants in RPS TA's are more likely to plan on upgrading hydropower plants. Plants with pumped storage facilities are also more likely to plan on upgrading. In non-RPS areas, no such plant-level effects are found.

These findings can be interpreted in the context of the state RPS laws. The positive values on the nameplate variable is most likely a reflection of the common-sense idea that a larger plant has more generators that could potentially be upgraded. However, the fact that this effect is only observed in RPS-states provides evidence of an association between the existence of RPS-laws and planned hydropower upgrades.

The results on the plant-level coefficients provide evidence that added capacity in hydropower is being used as a (presumably) cost effective way of meeting RPS standards, rather than as a complementary balancing generation to extra intermittent generation.

Figure 6 shows that most of the β^B terms, which represent the interaction effects between the plant-level variables and the TA-level intermittent capacity variable, are also centered around zero. The distributions of the parameters for age and nameplate capacity are centered around slightly negative values. In areas under a RPS regime with large amounts of wind and solar capacity, older and larger hydropower plants are less likely to plan an upgrade, relative to the main effects displayed in Figure 5.

A presentation of all the estimated parameters on the 234 TA areas would be space prohibitive. However, Figure 7 in the appendix gives a visual summary of the parameters A_{ta}, which represent the intercept terms that vary by TA. The figure shows the a 95% confidence interval of the coefficients ordered by whether the TA was under an RPS regime and the penetration of intermittent energy generation in the TA. The figure shows how the estimated A_{ta} coefficients are on average slightly larger under a RPS regime. The figure are also shows how A_{ta} falls with high penetrations of intermittent generation capacity. These results were captured by the

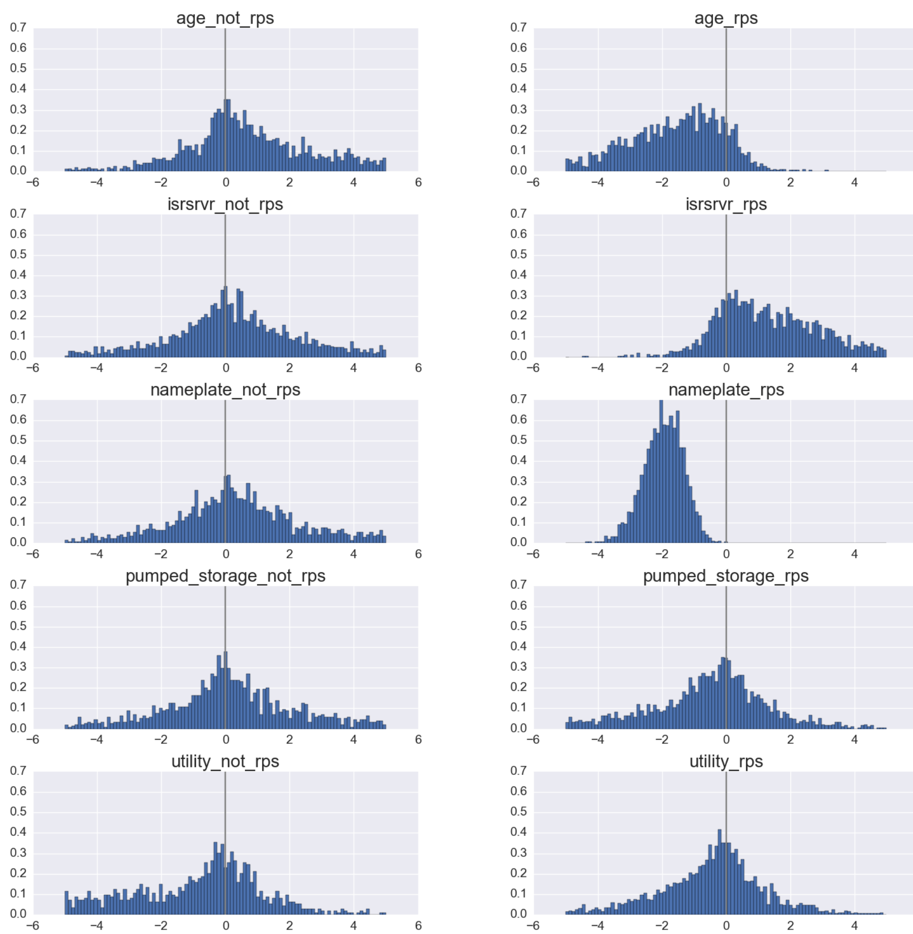

Figure 6: The estimated distributions on the β^B coefficients. These represent the interactions between the plant-level variables and the TA-level intermittent variable. Planned hydropower capacity additions are more likely in smaller and newer plants when the purpose is as a substitute for intermittent generation to meet a renewable portfolio standards.

meta-parameters discussed above.

RPS standards vary substantially by state. Because our observations are relatively evenly distributed between 23 states, attempting a state-by-state analysis becomes prohibitive due to a lack of data and relatively scarce positive outcomes. However, state-by-state variation is taken into account in the multilevel model by way of the TA-level random effect terms, A_{re}^{ta}. A visual summary of the distributions of these random effects do not show any discernible pattern when ordered by the RPS level, as shown in Figure 8 in the Appendix.

As an additional robustness check, we also explore models where proposed capacity of uprates are used as the dependent variable. The results can be found in Section C in the Appendix. The results are consistent with those we have presented here.

7 Discussion and Conclusion

Our main findings are as follows.

- Renewable portfolio standards (RPS) are associated with higher probabilities of planned investment in hydropower upgrades.

- Plants with pumped storage facilities tend to have a higher probability of planned capacity upgrade.

- Planned hydropower upgrades are negatively correlated with the amount of installed intermittent generation, though only where RPS laws apply.

Our results indicate that planned hydropower capacity additions are being used as a way of meeting RPS laws as a substitute for intermittent solar and wind power rather than as a complementary balancing investment. The pattern of investment decisions tends to match restrictions that many states have in place for hydropower, and areas with large amounts of solar and wind power tend to see less probability of investment in hydropower capacity.

We account for two sources of unobserved variables that could bias our results. Multilevel models allow us to control for unobserved geographic variables. The use of planned investment rather than realized investment avoids the influence of confounding regulatory factors and other unobserved variables that affect the timing of completed investments. However, states with large amounts of hydropower capacity could be motivated to implement RPS laws because they have renewable generation in place. Figure 2 displays no clear relationship between (initial) hydropower resources and existence of RPS laws.

Despite the model's attention to possible sources of endogeneity, it falls short of estimation based on random placement into treatment and control groups, and the results should be interpreted with this in mind. That said, we believe we have controlled for the major sources of endogeneity.

We note that, despite the rapidly falling costs of solar and wind power and the many restrictions put on hydropower investments, investments in hydropower expansion are still competitive - at least in certain forms and in certain areas. Perhaps, just as in financial markets, utilities are interested in having a diversified portfolio of generating assets.

Table 3: Logit Estimation Results

The dependent variable $investment_{y,p}$ is an indicator variable which is equal to one if plant p had a planned upgrade in year y and zero otherwise. $rptso_{y,j}$ is an indicator variable which is equal to one if TA j operates in states with RPS in place in year y and zero otherwise. $intermittent_{y,j}$ is the percent of total capacity in TA j in year y which is intermittent in nature. $age_{y,p}$ is the age in years of plant p in year y. $irsrvr_p$ is an indicator variable which is equal to one if plant p includes a reservoir and zero otherwise. $nplate_p$ is the nameplate capacity of plant p, in MW. ips_p is an indicator variable which is equal to one if plant p utilizes pumped storage and zero otherwise. $iutil_p$ is an indicator variable which is equal to one if plant p is owned by a vertically integrated utility and zero otherwise. The table presents the average marginal effects $\left(\partial Prob(investment = 1)/\partial x\right)$ of each independent (x) variable. For the indicator variables, the table presents the change in the probability of investment when the variable changes from zero to one.

	Base	*Individual Interactions*						*Full*
	I	II	III	IV	V	VI	VII	VIII
rpsto	0.014**	0.012**	-0.001	0.017***	0.011*	0.013**	-0.000	-0.017
intermittent	-0.354***	-11.222	-0.353***	-0.355***	-0.354***	-0.353***	-0.354***	-0.351***
$intermittent * rps$		10.878						
age	0.000	0.000	-0.000	0.000	0.000	0.000	0.000	0.000
$age * rps$			0.000					0.000
irsrvr	-0.016**	-0.016**	-0.017**	-0.005	0.017**	-0.017**	-0.016**	0.033
$irsrvr * rps$				-0.016				-0.025**
nplate	0.017***	0.017***	0.017***	0.017***	-0.051	0.017***	0.017***	0.002
$nplate * rps$					0.017***			0.015
ips	0.023	0.024	0.023	0.028	0.068	-0.005***	0.022	-0.045***
$ips * rps$						0.949		0.952
iutil	0.018***	0.018***	0.018***	0.018***	0.018***	0.017***	0.003	0.001
$iutil * rps$							0.018	0.010
pseudo-R^2	0.082	0.084	0.083	0.083	0.083	0.083	0.082	0.088
Log-likelihood	-510.0	-509.4	-509.3	-509.4	-509.6	-509.3	-509.7	-506.8

References

[1] Amundsen, E. S., and J. B. Mortensen, 2001, "The Danish Green Certificate System: some simple analytical results," *Energy Economics*, 23(5), 489–509.

[2] Bird, L., C. Chapman, J. Logan, J. Sumner, and W. Short, 2011, "Evaluating renewable portfolio standards and carbon cap scenarios in the U.S. electric sector," *Energy Policy*, 39(5), 2573–2585.

[3] Böhringer, C., and K. E. Rosendahl, 2010, "Green promotes the dirtiest: on the interaction between black and green quotas in energy markets," *Journal of Regulatory Economics*, 37(3), 316–325.

[4] Fischer, C., 2010, "Renewable Portfolio Standards: When Do They Lower Energy Prices?," *The Energy Journal*, 31(1), 101–119.

[5] Gelman, A., 2006, "Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper)," *Bayesian Analysis*, 1(3), 515–534.

[6] Gelman, A., J. Carlin, H. Stern, D. Dunson, A. Vehtari, and D. Rubin, 2013, *Bayesian Data Analysis, Third Edition*. Chapman and Hall/CRC, Boca Raton, 3 edition edn.

[7] Gelman, A., and J. Hill, 2006, *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press, Cambridge ; New York, 1 edition edn.

[8] Gelman, A., A. Jakulin, M. G. Pittau, and Y.-S. Su, 2008, "A weakly informative default prior distribution for logistic and other regression models," *The Annals of Applied Statistics*, 2(4), 1360–1383.

[9] Green, R., and N. Vasilakos, 2012, "Storing Wind for a Rainy Day: What Kind of Electricity Does Denmark Export?," *The Energy Journal*, Volume 33(Number 3).

[10] Kruschke, J. K., 2014, *Doing Bayesian Data Analysis, Second Edition: A Tutorial with R, JAGS, and Stan*. Academic Press, Boston, 2 edition edn.

[11] Linares, P., F. J. Santos, and M. Ventosa, 2008, "Coordination of carbon reduction and renewable energy support policies," *Climate Policy*, 8(4), 377–394.

[12] Linnerud, K., A. M. Andersson, and S.-E. Fleten, 2014, "Investment timing under uncertain renewable energy policy: An empirical study of small hydropower projects," *Energy*, 78, 154–164.

[13] Manton, K. G., M. A. Woodbury, E. Stallard, W. B. Riggan, J. P. Creason, and A. C. Pellom, 1989, "Empirical Bayes Procedures for Stabilizing Maps of U.S. Cancer Mortality Rates," *Journal of the American Statistical Association*, 84(407), 637–650.

[14] Mauritzen, J., 2013, "Dead Battery? Wind Power, the Spot Market, and Hydropower Interaction in the Nordic Electricity Market," *The Energy Journal*, 34(1).

[15] Menz, F. C., and S. Vachon, 2006, "The effectiveness of different policy regimes for promoting wind power: Experiences from the states," *Energy Policy*, 34(14), 1786–1796.

[16] Palmer, K., and D. Burtraw, 2005, "Cost-effectiveness of renewable electricity policies," *Energy Economics*, 27(6), 873–894.

[17] Tsao, C. C., J. E. Campbell, and Y. Chen, 2011, "When renewable portfolio standards meet cap-and-trade regulations in the electricity sector: Market interactions, profits implications, and policy redundancy," *Energy Policy*, 39(7), 3966–3974.

[18] U.S. Department of Energy, 2016, "Hydropower Vision: A New Chapter for America's 1st Renewable Electricity Source," .

[19] Wiser, R., and M. Bolinger, 2007, "Can deployment of renewable energy put downward pressure on natural gas prices?," *Energy Policy*, 35(1), 295–306.

[20] Yin, H., and N. Powers, 2010, "Do state renewable portfolio standards promote in-state renewable generation?," *Energy Policy*, 38(2), 1140–1149.

Appendix

A Bayesian model summary statistics

Table 4 shows summary statistics of the marginal posterior distribution on the upper level estimated parameters. The table shows the 2.5, 15, 50, 85, and 97.5 percentiles of the simulated distribution on the parameters. The probability mass between the 15th and 85th percentile is approximately equal to +/- 1 standard deviation from the median in a standard normal distribution. The probability mass between the 2.5th percentile and 97.5th percentile is approximately equal to 2 standard deviations from the median in a standard normal distribution.

B TA-level coefficients

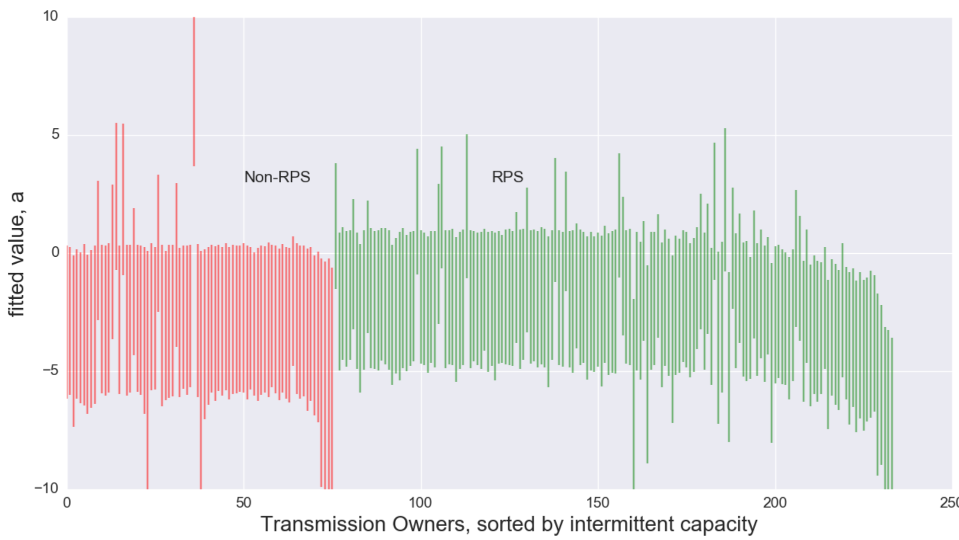

Figure 7: The estimated 95% confidence intervals of TA-level coefficients, A_{ta}, ordered by whether the TA is under a RPS and then by the amount of intermittent capacity in the TA area. The figure shows how the average A_{ta} coefficient is on average slightly larger under an RPS regime. The figure also shows how A_{ta} falls with high penetrations of intermittent energy.

Table 4: The table presents summary statistics of the estimated marginal posterior distributions on the upper-level coefficients of the model. The α^A represents the intercept terms. The α^B parameters represent the main effects on the plant-level variables. The β^A variables represent the the main effects on the intermittent capacity variable. β^B parameters represent interaction effects between the plant-level variables and the intermittent capacity variable.

		2.5%	15%	50%	85%	97.5%
param	var					
α^A	not_rps	-7.652	-5.776	-1.785	-0.001	0.414
	rps	-4.921	-4.440	-0.712	0.325	1.196
α^B	isrsrvr_not_rps	-2.004	-0.790	-0.056	0.474	1.149
	isrsrvr_rps	-2.379	-1.669	-0.916	-0.294	0.013
	pumped_storage_not_rps	-9.916	-2.220	-0.260	0.285	1.381
	pumped_storage_rps	-0.774	-0.276	0.099	0.605	1.254
	utility_not_rps	-1.706	-0.551	-0.012	0.620	1.738
	utility_rps	-0.383	-0.117	0.129	0.466	0.741
	nameplate_not_rps	-4.062	-1.006	-0.066	0.582	1.799
	nameplate_rps	0.326	0.535	0.845	1.233	1.579
	age_not_rps	-1.591	-0.757	-0.132	0.293	0.801
	age_rps	-0.337	-0.111	0.102	0.346	0.610
β^A	not_rps	-11.309	-3.314	-0.607	0.244	1.139
	rps	-3.902	-2.889	-1.858	-0.989	-0.276
β^B	isrsrvr_not_rps	-4.625	-1.559	-0.067	1.324	3.354
	isrsrvr_rps	-2.381	-0.773	0.175	1.459	2.974
	pumped_storage_not_rps	-4.340	-1.071	0.315	4.954	22.835
	pumped_storage_rps	-3.631	-1.173	0.049	1.328	3.713
	utility_not_rps	-12.295	-3.108	-0.418	0.643	2.926
	utility_rps	-1.105	-0.396	0.321	1.474	2.846
	nameplate_not_rps	-5.350	-1.276	0.157	2.287	13.098
	nameplate_rps	-4.528	-3.470	-2.315	-1.318	-0.721
	age_not_rps	-2.616	-0.847	0.134	1.369	4.073
	age_rps	-2.986	-1.979	-0.806	0.019	0.793

Figure 8: The estimated 95% confidence intervals of the TA-level random effects coefficients, A_{ta}^{re}, ordered first by whether the TA is under a RPS, and then by the RPS level, in percent. The blue text refers to the RPS level. The numbers under the non-RPS side refer to states with voluntary RPS laws. No pattern is apparent between the random effects terms and the RPS-target.

C Robustness: Proposed uprate capacity

The structure and availability of the data suggested a binary outcome variable in our main analysis. As a robustness check, we use the 2014 data to complete an analysis in which the outcome variable is a continuous variable of proposed upgrade capacity. We have available data for only those plants that were planning an uprate, and even here, some data were missing. In total we have data for proposed capacity of approximately half of the positive outcomes.

Replicating the full Bayesian probability model with capacity as the dependent variable is infeasible. There are not enough positive responses in order to estimate a complete multilevel model which fully takes into account heterogeneity across a large number of geographies and technologies.

Instead, we present results from two simpler models in Table 5. In these regressions, the plant-level variables are included identically in each model with fixed coefficients. TA-level variables indicating whether a plant is under an RPS-regime and how much intermittent capacity is present in the TA are also present in both models. In the first model, these TA-level variables are entered as fixed coefficients with an interaction term. In the second model, the intercept term and the *intermittent* variable have coefficients that are allowed to vary between RPS and non-RPS regimes.

We focus on the results from the latter model as they are most in line with the full model. In particular, we find that the intercept term under RPS-regime is significantly positive. On the other hand the intercept term is not found to be significantly different from 0 in the non-RPS regime. This is consistent with our earlier results that found RPS-regimes did have a positive impact on planned hydropower upgrades.

We also find that the *intermittent* coefficient is significantly negative under RPS regimes, but insignificant under non-RPS regimes. This is also consistent with our results from the complete model that hydropower upgrades play a role as substitutes for wind and solar power, rather than as complementary balancing power.

	fixed	varied
Intercept	-0.063	0.074
	(0.177)	(0.158)
isrsrvr	0.859^{***}	0.860^{***}
	(0.145)	(0.144)
nameplate_100mw	0.016^{**}	0.016^{**}
	(0.005)	(0.005)
age_10	-0.043^{**}	-0.042^{**}
	(0.015)	(0.015)
utility	0.215^{*}	0.210^{*}
	(0.103)	(0.102)
intermittent	0.116	
	(1.257)	
isRPS	0.365^{*}	
	(0.145)	
intermittent:isRPS	-0.765	
	(1.345)	
Intercept, RPS		0.268^{**}
		(0.054)
Intercept, not RPS		-0.029
		(0.108)
intermittent, RPS		-0.403^{**}
		(0.113)
intermittent, not RPS		0.216
		(0.226)
AIC	8460	19830
Num. obs.	3996	3996

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table 5: Results from a model with proposed capacity uprate as the dependent variable. The first column shows results where the RPS variable and *intermittent* variable enter with fixed coefficients. In the second column, the intercept and *intermittent* variables enter with coefficients that vary by RPS status. The results are consistent with those found in the complete probability model.