Intermittency and the Value of Renewable Energy*

Gautam Gowrisankaran[†] Stanley S. Reynolds[‡] Mario Samano[§]

December 7, 2011

Abstract

This paper develops an empirical approach to estimate the equilibrium value of intermittent renewable energy. We model an electricity system operator who optimizes the amount of generation capacity, operating reserves, and demand curtailment potentially in the presence of large-scale solar facilities. We use generator characteristics, solar output, demand and weather forecast data to estimate parameters for southeastern Arizona. The deadweight loss of a 20% solar mandate is 79% of its \$184/MWh average cost. Unforecastable intermittency accounts for \$12.5/MWh. At a \$21/ton social cost of CO_2 this mandate is welfare neutral if solar capacity costs decrease from \$5/W to \$1.38/W.

*We thank Allan Collard-Wexler, Alex Cronin, Joseph Cullen, Lucas Davis, Ken Gillingham, Ben Handel, Bill Hogan, Paul Joskow, David Keith, Derek Lemoine, Ariel Pakes, Catherine Wolfram, John Wooders, Gregor Zoettl, and seminar and conference participants at numerous institutions for helpful comments; and Valentina Kachanovskaya for research assistance. We acknowledge research support from Arizona Research Institute for Solar Energy.

[†]University of Arizona, HEC Montreal and NBER, P.O. Box 210108, Tucson, AZ 85721, gowrisankaran@eller.arizona.edu

[‡]University of Arizona, reynolds@eller.arizona.edu

[§]University of Arizona, msamano@email.arizona.edu

1 Introduction

Electricity generation from fossil fuels is the single largest source of greenhouse gas (GhG) emissions in the United States. Several U.S. states and foreign countries have enacted laws to replace large fractions of their generation with renewables. Many observers consider solar energy from photovoltaics (PV) to be a crucial part of future renewable energy.[1] However, a key potential problem with solar and other renewable energy sources is intermittency. Solar facilities produce electricity intermittently, with by far the highest production levels during clear, sunny periods. The intermittency from solar energy increases the variance of the energy supply. This increased variance in turn adds risk to the electricity system because the inability to provide adequate supply to meet demand can result in a system outage – where a large number of customers do not receive power – with large welfare losses.[2] The intermittency and cyclic nature of renewable energy are seen as among the biggest hurdles to their large-scale adoption.[3]

This paper develops an empirical approach to value renewable energy accounting for intermittency. In conjunction with assumptions on the social cost of greenhouse gases, our approach can be used to understand the optimal levels and types of renewable energy. A simpler technique to value solar energy would be to use "levelized cost," which is the present values of future average cost. The levelized cost of a solar plant is its installation and maintenance costs divided by the total power it produces, discounted over its lifetime. The social value of a solar plant would then be its environmental benefits minus the difference in levelized cost between it and a fossil fuel plant. Yet, levelized cost differences do not account for the fact that the extra variability from having large-scale solar capacity may require grid operators to engage in costly precautions. Specifically, in the short run, operators may need

[1]Solar energy has recently attracted large amounts of venture capital funds, plant investment, and federal and state government subsidies [see Glennon and Reeves, 2010].

[2]The impact of solar generation on electric system outage is similar to the impact of distressed financial institutions on financial system risk. See Adrian and Brunnermeier [2010] for an analysis of financial system risk.

[3]For instance, a recent Texas state report [see SECO, 2011] identifies intermittency, cost and surface area as the three big challenges for solar energy, stating that "the solar resource's intermittency and cyclical nature pose challenges for integrating solar at a large scale into the existing energy infrastructure." Joskow [2010] notes that the value of renewable energy may be very different accounting for intermittency.

to schedule additional reserve generation to avoid a system outage. In the long run, they may need to invest in backup fossil fuel generation capacity to be used at times when demand is high but solar output is low. These effects are counterbalanced by the fact that solar plants tend to produce during peak consumption periods, which increases their value, as they offset generation from high marginal cost plants.

The extent to which different aspects of intermittency balance out is unknown and depends critically on how a particular renewable energy source affects optimal choices of generator scheduling, operating reserves and backup capacity. These decisions in turn depend crucially on three factors: (1) the variability of the source including the extent to which the variability correlates with demand; (2) the extent to which output from the source is forecastable; and (3) the costs of building backup generation. These general issues of intermittency for renewable power are well understood. A number of recent studies seek to quantify the potential importance of intermittency by considering one of these issues in depth. Some studies focus on the time-varying generation profile of renewable energy [see Borenstein, 2008, Denholm and Margolis, 2007, Joskow, 2010, Cullen, 2010b]; others model intermittency and its impact on operating reserves [see GE Energy, 2008, Mills and Wiser, 2010, Helman et al., 2011] and finally some deal with backup capacity investment [see Campbell, 2011, Hansen, 2008, Hoff et al., 2008, Skea et al., 2008]. By combining these three central factors in one model, we derive an overall economic assessment of the value of large-scale renewable generation.

Our approach calculates the equilibrium social costs of intermittency from renewable sources by endogenizing optimal policies as a function of renewable capacity. We use this approach to evaluate solar energy mandates in southeastern Arizona. We argue that our optimizing approach has important implications for evaluating intermittency costs. In particular, we find that if the planner did not change decision-making policies for operating reserves, demand curtailment, and backup investment when large-scale solar is adopted then the welfare cost of solar would be almost 10 times as high as under our optimizing approach, mostly due to high system outage probabilities. Even if the planner adopted heuristic backup capacity and reserve strategies that have been proposed for solar, welfare costs would still

be over 3 times as high. Our method could also be used to examine the equilibrium value of other renewable technologies such as wind power, as well as how developments such as real-time pricing and improvements in energy storage technology affect the value of renewable technologies.

The starting point of our approach is Section 4 of Joskow and Tirole [2007], who model a system operator of an electricity market who seeks to maximize the expected discounted present value (DPV) of welfare when faced with fossil fuel generators that can suddenly fail.[4] Our model builds on this paper by modeling renewable energy intermittency as similar to the unexpected failure of a traditional generator; by modeling variability and uncertainty in demand; and by using the empirical distribution of generator characteristics.

In our model, the system operator is faced with a fixed retail price and level of solar generation capacity, as specified by a state-mandated renewable portfolio standard (RPS). At time 0 the operator chooses how many new fossil fuel generators to build. The operator also sets the price for "curtailment contracts," which allow certain flexible customers, typically industrial users, to be paid not to consume electricity in periods of high demand. Each period, which is one hour, the operator is faced with a distribution of demand, and in the presence of renewables, a joint distribution of demand and renewable output. These distributions are derived from the previous day's weather forecasts. Renewable output can fluctuate at high frequency within the hour. Observing the distribution, the operator must then decide how many generators to schedule for generation and reserves and also how much demand to curtail, if any.

We apply our model to the portion of the electric grid operated by Tucson Electric Power (TEP), a regulated, investor-owned utility whose coverage area roughly consists of southeastern Arizona. We obtained data for the Tucson area on generator heat rates, fuel prices, retail electricity prices, line losses, demand at the hourly level, solar output at the 5-minute level, and day-ahead weather forecasts; as well as national data on capacity costs, prices of spinning reserves, generator outages and system outage events. We estimate the predictable and unpredictable components of demand and renewable outputs by regressing

[4] We use the terms "system operator" and "planner" interchangeably.

demand and renewable outputs on the weather forecast for that hour made the previous day,[5] using a seemingly unrelated regression (SUR) model to capture correlations in the unforecastable components of these processes. To our knowledge, our study is the first to use weather forecast data to separate intermittency into its forecastable and unforecastable components; and the first to formalize reserve operations in an economic model that is taken to data. We assume that demand has constant elasticity up to some maximum reservation price and calibrate both the elasticity and reservation price from the literature. We recover most of the other parameters with simple estimation techniques.

Our model and application have four central limitations. First, we do not model any dynamic linkages from period to period, as would occur with start-up costs for generators, for instance.[6] Second, we solve for the planner solution and hence do not model the impact of oligopoly power in the generation market.[7] Third, we do not model imports or exports outside of the local market, which can serve to reduce the variability of net demand, and hence lower the costs of intermittency. Finally, we use data only from a single solar site.

Note also that we do not structurally estimate demand. One could potentially recover the demand curve by a structural estimation process that would match the actual level of operating reserves to predicted values. However, we believe that it would be somewhat problematic to assume that current TEP decisions reflect optimizing behavior within the context of our model and hence that it is more credible to take these parameters from the literature.[8]

Using the estimated parameters of the model, we solve for the optimal policies under several counterfactual scenarios, involving different levels of solar capacity and assumptions

[5]Electricity system operators commonly schedule operating reserves one day ahead. For example, the system operator for the Electric Reliability Council of Texas (ERCOT) obtains operating reserves for each hour in one-day-ahead procurement auctions.

[6]Cullen [2010a] estimates a dynamic model of start-up costs for generators. A similar model would hugely complicate our analysis.

[7]It is possible to design mechanisms that would generate the equilibrium of our model.

[8]TEP is subject to rate of return regulation by the Arizona Corporation Commission. This form of regulation has the potential to introduce inefficiency, for example due to weak incentives to hold down costs of generation and of providing operating reserves. This regulatory effect is known as X-inefficiency. Wolfram [2005] finds evidence that non-regulated merchant power producers operate generation units at lower cost than do regulated investor-owned utilities. Furthermore, the TEP system operator may act in a more risk averse manner than predicted by our optimization model, whether due to career concerns or in response to regulatory penalties for system outage.

about forecastability. The optimal solution involves balancing very low probability but very costly system outages against the certain cost of additional operating reserves. Simulation of very low probability events can be computationally very time consuming. We develop a simulation procedure that oversamples multiple generator failures by summing over different number of failures and then simulating which generators fail given that a certain number fail.

Under the assumption that solar PV capacity costs are \$5/W, the levelized cost of solar generation in Tucson is \$184/MWh (18.4 cents/KWh).[9] This is \$126/MWh higher than the levelized cost for a new combined cycle natural gas unit.[10] Our model, which incorporates optimizing behavior of the system operator, finds that the true social cost of renewable energy is higher than the levelized cost difference and increasing in the level of solar capacity. Specifically, RPS policies of 10, 20 and 30% implemented solely with solar PV would impose equilibrium costs of \$138.8, \$145.9, and \$150.5 per MWh of solar generation respectively, not accounting for the benefit of CO_2 reduction. The upward slope is due to the increasing substitution from low cost generators and the increasing need to construct backup fossil fuel generators. With solar capacity, demand-side management in the form of curtailment contracts increases in periods with high demand and low solar output, leading to a higher offered price on these contracts. Re-optimizing policies is very important in making large-scale solar feasible: a planner who adopts a common heuristic approach for choosing backup capacity and managing operating reserves in response to a solar mandate would face an exorbitantly high cost of \$677.2/MWh for a 20% RPS, due to the average system outage probability rising from 0.006% to 1.1%.

Without unforecastable intermittency, the equilibrium costs of the 20% RPS would drop from \$145.9 to \$133.4 per MWh. In contrast, if solar energy were fully dispatchable, costs would drop to \$89.8/MWh and if capacity costs fell to \$2/W, equilibrium costs would drop to \$35.5/MWh. If CO_2 reductions are valued at \$21 per ton as mandated by recent U.S. government regulatory rules,[11] the 10% RPS would be welfare neutral with a capacity cost

[9] Throughout this paper we use total installed cost for solar PV capacity. Total installed cost includes the cost of PV panels, installation cost, and lifetime balance-of-system costs.
[10] See EIA [2011].
[11] See EPA [2010].

of \$1.38/W and the 30% RPS would be welfare neutral with a capacity cost of \$1.26/W.

The remainder of the paper is divided as follows. Section 2 provides a background on the electricity market. Section 3 discusses the model; Section 4 the data, estimation and computation; and Section 5 the results. Section 6 concludes. On-line Appendix B provides background on system operations and the electricity market in Tucson.

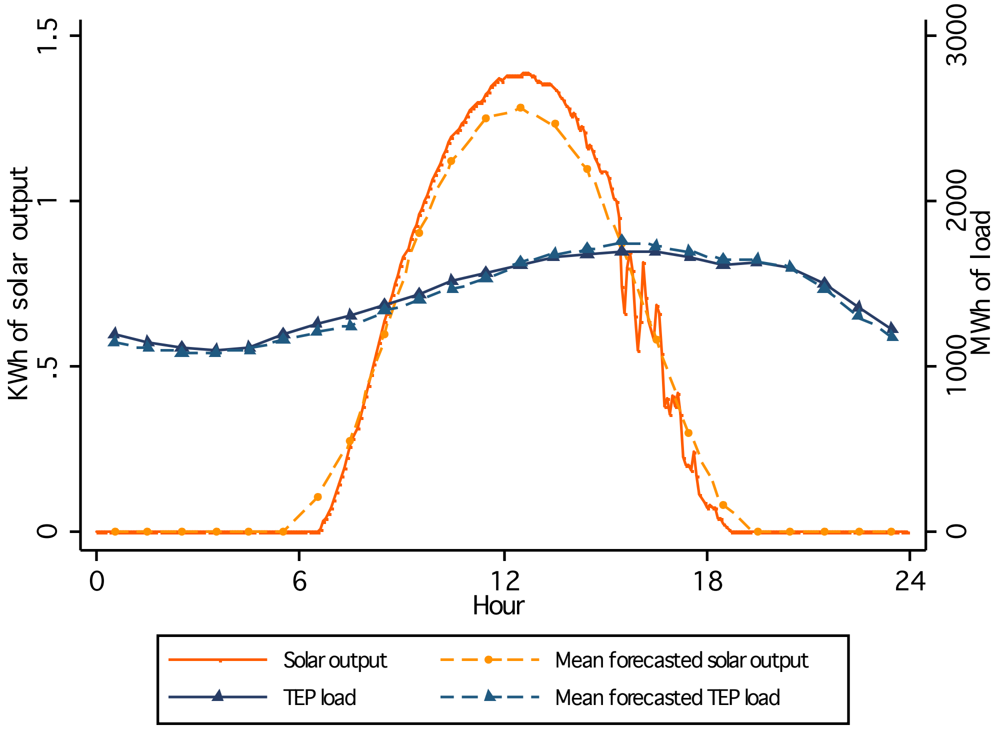
2 Solar Energy and Intermittency

Solar PV systems utilize panels of materials (such as silicon) that convert solar radiation into direct current (DC) electricity, coupled with inverters that convert DC current to alternating current (AC) that is used by customers [see NREL, 2011]. Electricity generation from solar PV panels varies with solar insolation, a measure of energy from sunlight. Higher solar insolation yields more PV generation, holding everything else constant. Most solar PV panels in the northern hemisphere – including those for which we have data – are mounted to face south at a fixed tilt, with the tilt based on latitude. There are also single-axis tracking systems in which the panels are typically rotated to track the motion of the sun through the day, and double-axis systems in which both the angle of panels and the direction in which panels face are controllable.

To illustrate the issues of intermittency, Figures 1 and 2 show southeastern Arizona demand and solar PV output in solid lines, for Apr. and Aug. 15, 2008 respectively.[12] Demand (in our data and in the figures) is at the hourly level and solar output at the five minute level. With dotted lines, Figures 1 and 2 also show the hourly mean forecasted demand and output using day ahead weather forecasts.[13] Because southeastern Arizona power demand is driven by air conditioning, it peaks during hot and sunny periods; but sunny periods also have a lot of solar production. Thus, solar output correlates positively with demand during the daytime. Figures 1 and 2 also illustrate that the correlation of solar output and demand is not perfect; daily peak demand tends to occur later in the day than peak solar output. Moreover, particularly in August, the solar output has large fluctuations that last only a

[12]The solar PV output is for a 1.536 KW test facility near the Tucson International Airport.
[13]We provide details on our forecast methodology in Section 4.

Figure 1: Predicted and actual southeastern Arizona load and solar output, Apr. 15, 2008



few minutes at a time, resulting from clouds. On one hand, the positive correlation between solar output and load will increase the value of solar production because production occurs when demand, and hence marginal costs of generation, are high. On the other hand, the unforecastable intermittency in solar production will lower its value.

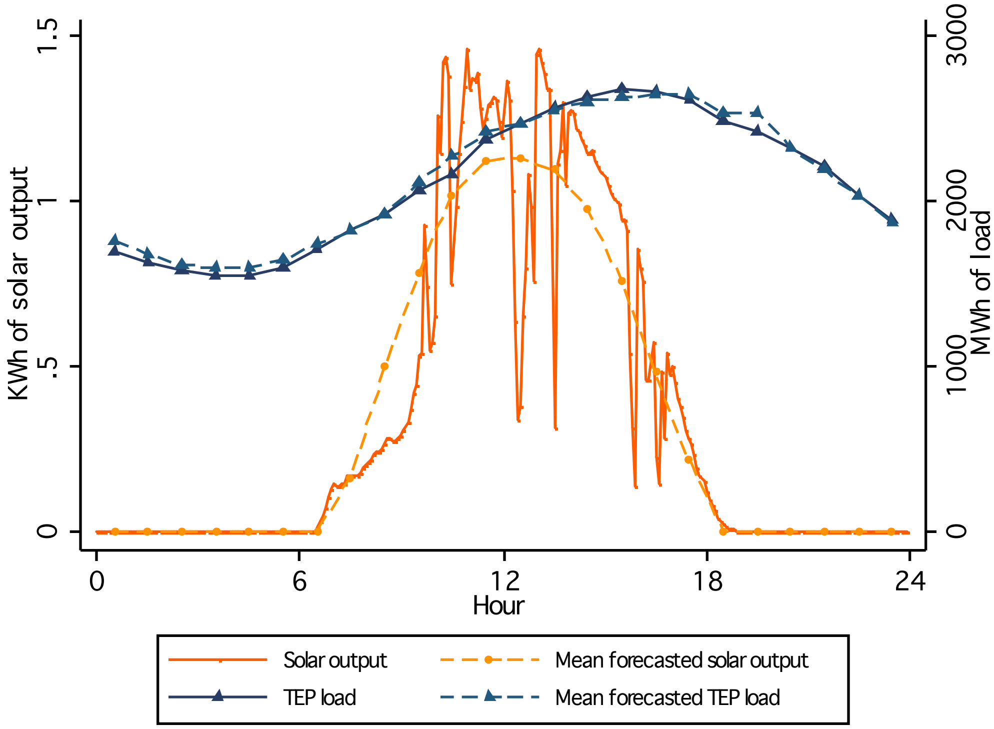
3 Model

3.1 Overview

We develop a model of electricity generation, system operations and the demand for power. In our model, at time $T = 0$, the regulator exogenously chooses the level of solar capacity and the retail price of electricity, $\overline{p}$.[14] Observing the solar capacity and existing fossil fuel generators, the system operator decides on capacity investment for new generation units. We assume that the new generation units, solar panels, and existing fossil fuel generators all last

[14] We assume $\overline{p}$ is fixed, consistent with the relatively inflexible retail pricing observed in most U.S. electricity markets. It is possible to relax this assumption to understand the relationship between real time pricing and the equilibrium value of renewables, among other questions.

Figure 2: Predicted and actual southeastern Arizona load and solar output, Aug. 15, 2008



until year $T = 25$. The system operator also chooses a price for curtailment contracts, which are fixed over a year.

Following the choice of curtailment contracts, the system operator is faced with a sequence of short-run production periods, each of which represents a particular hour of a particular day of the year. Each day the operator obtains 24 weather forecasts corresponding to the hours of the subsequent day. Using these weather forecasts, the operator computes the forecastable distribution of the solar output and load for each of these 24 periods. The operator also receives a report of which generators will be unavailable due to scheduled maintenance at each period. The operator then chooses which generators to schedule for production and reserves for each hour as well as how much load to curtail from the set of users who have signed up for curtailment contracts.

During each short-run period – and a day after the operator choice of generation and reserves – generator failures, solar output levels, and load are realized. Load takes on one value for each period, while solar output varies multiple times during each period, to max-

imally use our solar data which are at the 5-minute level. We can divide the realizations for each 5-minute interval into three cases. First, load could be less than the sum of the solar output and the realized output (from generators that have not failed) net of line losses. In this case, the system operator will reduce the rate of generation for one or more units to balance output with load. Second, load could be more than the output (also net of line losses), but less than output plus reserves. In this case, the system operator will move some of the capacity from reserves to production. Finally, load could be more than output plus reserves. In this (hopefully rare) case, a system outage occurs and results in a fraction of customers losing power, due to the system operator initiating an involuntary cut-off of power to some customers, to avoid a complete system collapse. Let d^{outage} denote the product of the fraction of customer who lose power times the number of periods for which they lose power.

The system operator makes all decisions in order to maximize the expected discounted value of future total surplus with a discount factor β. We assume that there are no linkages from period to period, as would occur with ramping constraints or start-up costs, for instance, implying that the short-term decisions of the operator have no dynamic consequences. Moreover, the same time periods in the year are repeated for each year of the T-year time horizon, implying that the price of curtailment contracts can be chosen once at the beginning.

3.2 Demand and Consumer Welfare

At the start of each hour-long period, the planner knows two vector-valued state variables: w, weather forecast information; and m, the scheduled maintenance status of each generation unit. Included in w are the time of day, the day of the week, and the time since sunrise and sunset, since these may predict load and/or solar output. Each state (w, m) thus implies a joint distribution of demand and solar generation as well as a probability distribution for generator failures. Although retail price is a constant $\overline{p}$, we need to specify the demand curve in order to quantify the cost of system outage and the response to curtailment contracts.

We choose a very parsimonious specification for demand in order to minimize the burden of identification. Specifically, we assume that demand has a constant price elasticity η for prices up to a reservation value, v. While the elasticity of demand is constant across states,

the level takes on a distribution that varies with w, $\overline{D} \sim F^D(\cdot|w)$. Demand is then

$$Q^D(p, \overline{D}) = \begin{cases} 0, & p > v \\ \overline{D}p^{-\eta}, & p \leq v. \end{cases} \tag{1}$$

We assume that $F^D(\cdot|w)$ has a lower bound $\overline{D}^{min}(w)$.

The term *value of lost load* (VOLL) is used in the electricity industry to describe the average value of electricity per unit for customers; see Cramton and Lien [2000]. Let $B(Q)$ be the gross consumer benefit function (area under the inverse demand curve) as a function of quantity Q. If Q is the quantity demanded at retail price $\overline{p}$ then $B(Q) = VOLL \times Q$. If there is a system outage during the period then the opportunity cost of the outage is $B(Q) \times d^{outage}$. VOLL and the reservation value v have a simple, monotonic relation given the price elasticity η:

Lemma 3.1. *With demand specified in (1) and retail price fixed at $\overline{p}$, VOLL is constant within and across states and satisfies:*

$$VOLL = \left(\frac{1}{1-\eta}\right) v^{1-\eta}\overline{p}^{\eta} - \left(\frac{\eta}{1-\eta}\right)\overline{p}. \tag{2}$$

Proof See Online Appendix A for derivation.

Our demand model allows for a system operator that offers interruptible power contracts, as described in Baldick et al. [2006]. In the first stage, the system operator chooses a curtailment price p_c and offers contracts whereby users would agree to have their power curtailed as necessary and be paid a net per-unit price of $p_c - \overline{p}$ as compensation. At this point, all users with valuation below p_c will sign up for interruptible power contracts. In each second stage period, knowing (w, m), the planner will choose the amount z of demand curtailment. When demand is curtailed, the planner randomly selects customers for curtailment from the set of customers who have signed up for interruptible power contracts and who are known to use power at that time.[15] We assume that the set of known users has mass $\overline{D}^{min}(w)$.

The amount by which the planner can curtail demand in any period is limited by $\overline{D}^{min}$, curtailment price p_c, and the price elasticity of demand. Specifically,

[15]We assume that it is not possible for the planner to curtail demand from the lowest valuation users. If possible, this would result in more efficient rationing.

Lemma 3.2. *If $p_c < v$, then curtailment z satisfies $0 \le z \le \overline{D}^{min}(w)\,[\overline{p}^{-\eta} - p_c^{-\eta}]$, with welfare loss of*

$$WLC(z, p_c) = \frac{\eta(\overline{p}^{1-\eta} - p_c^{1-\eta})z}{(\eta - 1)(\overline{p}^{-\eta} - p_c^{-\eta})}.$$

Proof See Online Appendix A for derivation.

The welfare loss function $WLC(z, p_c)$ indicates the loss in consumer benefits relative to the amount of gross consumer benefit $B(Q)$ when there is no curtailment. Note that there is a tradeoff from increasing p_c. An increase in p_c implies that the planner can curtail more demand, which increases expected welfare as it allows the planner to avoid system outage. However, an increase in p_c also implies an increase in the average valuation of the curtailed user, which decreases welfare as it increases $WLC(z, p_c)$.

3.3 Generation from Fossil Fuel and Solar PV

All new and existing fossil fuel and solar generators last from year 0 through year T. We assume that there is a set of existing dispatchable generation units indexed by $j \in \{1, ..., J\}$. At any given period, each unit has a maintenance status m_j, with $m_j = 1$ implying that the unit is unavailable for production. Maintenance probability is exogenous in our model and stochastic from the point of view of the system operator; let P_j^{maint} denote the probability of scheduled maintenance. We use data on generation unit maintenance outages to estimate the probability that a generation unit will be unavailable due to maintenance. Observing (w, m), the planner will schedule each available unit for production at full capacity k_j or no production; let on_j denote a 0-1 indicator for scheduled production. Note that $m_j = 1$ implies that $on_j = 0$.

Each unit uses a particular generation technology; coal, natural gas, etc. The marginal costs (MC) of generation for unit j are c_j. The MC of fossil fuel units depend on fuel cost, unit heat rate and costs associated with emissions. We assume that generator maintenance and failure events are *i.i.d.* across generators and independent from demand.

Generators can also be used to provide operating reserves which allows them to produce electricity in the case of the failure of another generator or load in excess of forecasted load.

For any generator, we assume that the marginal cost of reserves is a fraction c^s of the cost of producing electricity for whatever fraction of capacity of the generator is under reserve.[16]

Our model of generation unit failures is based on probabilities of losing the capacity of entire generating units. Potential output from (dispatchable) generation unit j is given by

$$x_j(on_j) = \begin{cases} k_j, & \text{with prob } (1 - P_j^{Fail})on_j \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

We allow the planner to invest in new fossil fuel generation capacity. Each of the new fossil fuel generation units is an identical combined cycle gas generator, which the same failure and maintenance probabilities as existing gas generators. We assume that each has fixed capacity size k^{FF}, investment cost of FC^{FF} per MW of capacity, and operating costs of c^{FF} per MWh. Knowing these values, the planner chooses the number of new generators, $n^{FF} \in \{0, 1, 2, \ldots\}$. We label the new fossil fuel units $j = J + 1$ through $j = J + n^{FF}$.

Similarly, we assume that solar PV capacity costs FC^{solar} per MW of installed capacity. Solar units have zero MC and maintenance and failure probabilities; scheduled maintenance costs are included in FC^{solar}. Unlike gas generators, solar PV generators are continuously scalable. We assume that the planner is faced with a fixed level of listed solar PV generation capacity n^{solar} as specified by an RPS-type mandate. Solar units are not dispatchable. Production from solar PV generation will take on a state-contingent distribution $n^{solar}\overline{S}$, where $\overline{S} \sim F^S(\cdot|w)$ is a vector, providing multiple measures of the rate of solar output during the hour, consistent with our high-frequency data. Denote $\overline{S} \equiv (\overline{S}^1, \ldots, \overline{S}^Y)$ where Y is the number of solar output observations in an hour; generally $Y = 12$ as we use 5-minute solar output data. Without loss of generality, order the solar output draws within a period so that $\overline{S}^1 \leq \ldots \leq \overline{S}^Y, \forall \overline{S}$. Let $F(\cdot|w)$ denote the joint distribution F^D, F^S of forecasted load and solar output. This formulation allows for the possibility of correlation between forecast

[16]Our model of the cost of reserves is a simplification of a much more complicated problem. In reality, there are several types of costs associated with maintaining operating reserves. Fossil fuel generation units typically have minimum and maximum generation rates and need to be operating at or above the minimum in order to provide operating reserve. Some high cost generation units may be operating at their minimum so that they have excess capacity from which reserves can be provided; this implies that some lower-cost generation units may be operating below their maximum which then implies an increase in production costs due to reserves. Also, it may be necessary to incur start-up costs for some of these higher cost units. Finally, units that are providing spinning reserves are not available for maintenance implying that there may be deferred maintenance costs associated with operating reserves.

errors for demand and for solar generation.

Even though TEP imports and exports power from the Western Interconnection, we do not model this possibility.[17] It would be problematic to model imports and exports since the import and export market will not necessarily stay constant as other jurisdictions may implement similar RPSs.

Finally, we discuss transmission costs. These costs may change with solar capacity because solar capacity can be partly locally distributed, e.g. on customers' rooftops. We assume that a fraction d^{solar} of solar capacity is installed in a distributed environment. We let the distributed nature of solar production affect transmission costs in two ways. First, we assume that the fixed costs of transmission (the DPV of equipment investment and maintenance costs) are a function of the *maximum* expected load across states net of distributed solar:

$$TFC(n^{solar}) = AFC^T max_w\{E[\overline{D}(w)\overline{p}^{-\eta} - \frac{1}{Y}\sum_{y=1}^{Y} d^{solar} n^{solar} \overline{S}^y(w)]\} \tag{4}$$

where AFC^T is the average transmission fixed cost per MW of non-distributed capacity.[18] To the extent that solar production is positive in periods with the highest load, solar capacity will lower the fixed costs of transmission.

Second, we model line losses using the identity that generation plus line losses must equal consumption. Following Borenstein [2008] and Bohn et al. [1984], we assume that line losses from transmission in any period are equal to a constant α times the square of non-distributed generation. Let Q be load minus demand curtailment minus distributed solar generation. If line loss is given by LL then the quantity $Q+LL$ is equal to total non-distributed generation. Line loss satisfies $LL = \alpha(Q + LL)^2$; let $LL(Q)$ be the function implicitly defined by this relationship,[19]

$$LL(Q) = (2\alpha)^{-1}(1 - 2\alpha Q - \sqrt{1 - 4Q\alpha}) \tag{5}$$

We model line losses using (5). From (5), distributed solar production will reduce line losses,

[17] This assumption of not allowing imports or exports has been used in the literature that uses electricity data from the Western U.S. An example is the analysis of real-time-pricing using California data; see Borenstein and Holland [2005].

[18] We assume that local distribution costs are not impacted by a change in the amount of distributed solar capacity. Local distribution costs tend to be driven by factors such as population and population density.

[19] There are two roots to the quadratic equation. Welfare is maximized at the smaller root.

especially when it occurs during periods with high load.

3.4 Planner's Problem

We seek to characterize the social planner's problem of maximizing expected discounted total surplus, subject to $\overline{p}$ and n^{solar}.[20] In the first stage, the planner chooses n^{FF} and p_c. In each second-stage period, the planner makes two decisions conditional on the state (w, m) and first-stage decisions: (1) generator scheduling decisions *on* and (2) amount of demand to be curtailed, z.

We model the choice of spinning reserves as a simplified version of how reserves are treated in unit commitment models.[21] Upon learning the state (w, m), the planner chooses on_j for each unit with $m_j = 0$. Then, the state-specific random variables are realized. Possibly, a system outage occurs, with a large fraction of customers not getting power. Otherwise, the planner will adjust actual generation to be exactly equal to demand. Observing actual demand and generator failure, the system operator can minimize costs by using the generators with the lowest marginal costs to satisfy demand, leaving the generators with the highest marginal costs as reserves. Let $PC(D, x)$ denote the ex-post minimized costs of power generation and reserves, where D denotes demand (net of curtailment) plus line loss minus solar production, and x denotes generator output realization vectors.

We illustrate the calculation of PC with a simple example. Consider a case with two scheduled generators each with capacity 1, with $c_2 > c_1$, $D = 1.6$ and no generator failures. Following the demand realization, the planner would partially shut down generator 2 as it has higher costs. Thus, the total production plus reserve costs would be $PC(1.6, (1, 1)) = c_1 + 0.6 \times c_2 + 0.4 \times c_2 \times c^s$.

A system outage occurs when fossil fuel generation is less than demand minus the mini-

[20] Although we have developed our model using a single-agent social optimum approach, our results would be replicated by a market-based model, similar to ERCOT. With multi-unit firms, Vickrey auctions for the generation and reserves markets could be used to induce efficient outcomes. A Vickrey auction specifies that a firm that sells k units is paid the lowest k losing offers submitted by rival firms [see Krishna, 2010].

[21] A unit commitment model would specify the cost of generation as well as costs of several types of reserves for each unit: spinning reserve up (to provide for an increased rate of generation), spinning reserve down (to provide for a reduced rate of generation), and non-spinning reserves. Bouffard et al. [2005] formulate and analyze a unit commitment model with stochastic demand. Our model is simplified in that we have a single type of operating reserve, which is essentially both a spinning reserve up and down.

mum solar production plus line losses, taking into account both generator failures and demand curtailment. The system outage probability, conditional on the state and actions is then

$$SOP(z, on, w, n^{FF}, n^{solar}) =$$

$$Prob\left[\sum_{j=1}^{J+n^{FF}} x_j(on_j) < \overline{D}(w)\overline{p}^{-\eta} - z + LL\left(\overline{D}(w)\overline{p}^{-\eta} - z - d^{solar} n^{solar} \overline{S}^1(w)\right) - n^{solar}\overline{S}^1(w)\right].$$

The planner's second-stage problem for a single period may now be defined as

$$\begin{aligned}
& W(w, m \mid n^{FF}, n^{solar}, p_c) = \max\nolimits_{z,on} \\
& \quad \Big\{E\left[(1 - d^{outage} SOP(z, on, w, n^{FF}, n^{solar}))\left(B(\overline{D}(w)\overline{p}^{-\eta}) - WLC(z, p_c)\right)\right. \\
& \quad \left. -\tfrac{1}{Y}\textstyle\sum_{y=1}^{Y} PC\left(\overline{p}^{-\eta}\overline{D}(w) - z + LL\left(\overline{p}^{-\eta}\overline{D}(w) - z - d^{solar} n^{solar}\overline{S}^y(w)\right) - n^{solar}\overline{S}^y(w), x(on)\right) \mid w\right]\Big\} \\
& \text{such that } m_j = 1 \implies on_j = 0.
\end{aligned} \tag{6}$$

From (6), the planner trades off the expected consumer welfare accounting for the system outage possibility and demand curtailment (the first line) against the production costs (the second line). Generators can only be operated if they are not undergoing scheduled maintenance (the third line).

The expected operating reserves associated with a decision are fossil fuel production plus reserves minus demand (net of curtailment) minus line losses plus solar:

$$E\left[\sum_{j=1}^{J+n^{FF}} x_j(on_j) - (\overline{D}(w)\overline{p}^{-\eta} - z) - \frac{1}{Y}\sum_{y=1}^{Y}\left[LL(\overline{p}^{-\eta}\overline{D}(w) - z - d^{solar} n^{solar}\overline{S}^y(w))\right] \mid w, m\right]$$

Extra generation in the form of operating reserves provides a "cushion" in the event that one or more generators fail, load exceeds forecast load, and/or renewable generation falls short of forecast renewable generation.

The planner rolls up the second-stage payoffs by taking the expected value of W in (6) over all the hours in one year, and then discounting the expected annual welfare over the life of generators. Specifically, defining N to be the number of hours in a year, expected discounted surplus is

$$\begin{aligned}
TS(n^{FF}, n^{solar}, p_c) = N \times \frac{1 - \beta^T}{1 - \beta} \times \\
E[W(w, m) \mid n^{FF}, n^{solar}, p_c] - FC^{FF} k^{FF} n^{FF} - FC^{solar} n^{solar} - TDFC(n^{solar}).
\end{aligned} \tag{7}$$

In the first stage the planner chooses the number of new fossil fuel units n^{FF} and compensation p_c per unit for demand curtailment to maximize $TS(n^{FF}, n^{solar}, p_c)$ in (7). The amount of solar PV generating capacity, n^{solar}, is constrained via RPS regulations.

4 Data, Estimation, and Computation

4.1 Data

In order to estimate and calibrate the parameters of our model, we use data from a variety of sources. These include the Energy Information Administration (EIA), the Environmental Protection Agency (EPA), ERCOT, TEP, FERC and NOAA. Our data pertain mostly to the Tucson area in 2008.

We use 2008 hourly load data for the Tucson service area from a FERC Form 714 filing by TEP. Summary statistics on load data are provided in Table 1. The peak month for electricity demand was August, due to hot weather and high air conditioning use. March was the month with the lowest electricity demand. Demand for electricity has grown by roughly one percent per year in southeastern Arizona.

We draw our data on generation units serving Tucson in 2008 from several sources. The EIA maintains a database on all existing generation units in the U.S. This database includes information about capacity, fuel source, and location. We obtain information on heat rates from EPA eGRID2010. This EPA database provides heat rates at the plant level, where a plant may have multiple generation units. We assume that each generation unit at a plant site has the same heat rate.[22] The EIA also has information about capacity investment cost for new generation units and average retail electricity price. We use information about total line losses for TEP from UniSource [2008].

The eGRID2010 database has average annual emission rates for CO_2, SO_2, and NO_x at the plant level. We apply the same emission rates for each generation unit at a plant. SO_2 permit fees are from the EPA's annual advance auctions for years 2011 – 2017. TEP units

[22] We make one exception to this assumption. We use data on heat rates at the individual generation unit level from the 2008 EIA form 923 report for the H. Wilson Sundt Generation Plant. This plant has several types of units with large differences in heat rates across units. Online Appendix C provides additional details about TEP generation facilities.

Table 1: Summary Statistics for TEP Hourly Load (MWh), 2008

Month	Average	Standard deviation
January	1,344	118
February	1,314	123
March	1,288	125
April	1,345	182
May	1,432	262
June	2,041	477
July	2,088	407
August	2,101	408
September	1,913	386
October	1,597	281
November	1,434	163
December	1,506	144
Number of observations:	8,784	

are not subject to NO_x permit fees. EPA's NO_x Budget Trading program, a cap and trade program for NO_x, applies to 20 eastern states, but does not apply to Arizona [see EPA, 2011].

Since our analysis is forward looking, we use information about projected future fuel costs. EIA Form 423 contains information about the terms of multi-year fuel contracts for each of the coal-fired generators. For natural gas we use NYMEX futures prices at Henry Hub in Louisiana [see CME, 2011]. We collect the last settlement price for each month for futures contracts in December 2010 for delivery from January 2011 through December 2015. Our natural gas price is the average of these prices.

We use actual solar generation data for 2008 from a solar PV test site near the Tucson International Airport run jointly by TEP and the University of Arizona [see TEP, 2011] and measured at the 5 minute level. This system has 24 solar PV modules with total rated capacity of 1.536 kW.[23] The modules are at a fixed 30 degree tilt facing south. Summary statistics on solar output are given in Table 2. Unlike electricity demand, solar generation is relatively consistent throughout the year. Our data include most, but not all, hours in 2008. There are now solar PV panels available with higher efficiency, and hence lower average

[23]These modules were produced by Uni-Solar and installed in 2003; each module has peak power of 64 W. A 1.536 kW facility is relatively small, somewhat smaller than the size of a typical residential installation. Solar PV panels generate electricity with roughly constant returns to scale, so we are able to use generation data from this facility to make generation projections for a much larger facility. Also, the fact that the panels were five years old at the time of sampling means that the data reflect to some degree the effect of aging on solar PV generation. PV cell production declines by approximately one percent per year over the life of the cell [see Borenstein, 2008].

generation cost, than the panels from which our data are drawn.

A novel aspect of this project is collection and use of weather forecast data which are used to determine the day-ahead forecasts of load and solar generation. We collect weather forecast data from the National Climatic Data Center of the National Oceanic and Atmospheric Administration (NOAA) [see NOAA, 2011]. The forecasts are generally at 3 a.m. for the next day at windows of 3 hours. We interpolate to convert to hourly forecasts. Information includes cloud cover, wind speed, temperature, relative humidity and dew point. All information is reported as a continuous measure except for cloud cover, which is reported as one of six discrete measures ("overcast" to "clear") each corresponding to an interval in terms of the numerical percent of sunlight passing through. We convert cloud cover to a continuous measure using the midpoint of the interval. Our weather forecast data is from the KTUS NOAA weather station, which is located at the Tucson International Airport. These data also include most, but not all, hours in 2008. Table 3 provides information on the variables used in the weather forecast. We supplement the NOAA weather information with data on sunrise and sunset times at the daily level [see Sunrise, 2011].

Data for generation unit outages come from the Generating Availability Data System (GADS) of NERC. GADS includes data for outages due to maintenance and for forced outages. We use GADS data for 2005-2009 for all U.S. generation units. An outage probability is calculated as the ratio of average number of occurrences to average total available hours. We compute maintenance outage and forced outage probabilities separately for coal units and for natural gas units.

We obtain data on system outage durations and number of affected customers from EIA reports on "Major Disturbances and Unusual Occurrences" in the U.S. [EIA, 2010]. Finally, we use ERCOT auction price data to define the costs of operating reserves.

4.2 Estimation and Calibration of Parameters

Table 4 lists the demand parameters. Short-run electricity demand is typically estimated to be quite price inelastic – see Espey and Espey [2004] for a survey and meta-analysis. Our value of $\eta = 0.1$ is somewhat lower than the median estimate reported in Espey and Espey

Table 2: Summary Statistics for Tucson Solar Test Site, 2008

Month	Mean output (kWh)	Hour	Mean output (kWh)
Jan. 2008	0.282	5 AM	0.0001
Feb. 2008	0.325	6 AM	0.015
Mar. 2008	0.376	7 AM	0.151
Apr. 2008	0.403	8 AM	0.471
May 2008	0.373	9 AM	0.792
Jun. 2008	0.363	10 AM	1.022
Jul. 2008	0.334	11 AM	1.145
Aug. 2008	0.352	12 PM	1.172
Sep. 2008	0.389	1 PM	1.128
Oct. 2008	0.374	2 PM	0.984
Nov. 2008	0.320	3 PM	0.752
Dec. 2008	0.243	4 PM	0.443
		5 PM	0.150
		6 PM	0.020
Rated capacity: 1.536 kW		7 PM	0.001
Average output: 0.345 kWh		8 PM – 4 AM	0

Table 3: Summary Statistics for Information Used in Weather Forecasts, 2008

Forecast Variable	Average	Standard deviation
Cloud cover (%)	27.7	20.0
Temperature (ºF)	70.4	16.9
Dew point (ºF)	36.5	15.2
Relative humidity (%)	34.3	19.1
Wind speed (MPH)	8.53	4.06
Number of observations:	8,448	

Table 4: Demand parameters

Parameter	Interpretation	Value	Source
η	Demand elasticity	0.1	Espey and Espey [2004]
$\overline{p}$	Retail price per MWh	\$95.6	EIA
g	Demand growth factor	1.2	Based on historical rate of demand growth
v	Demand reservation value per MWh	\$11,655	Computed so that VOLL is \$8,000/MWh
$F \equiv (F^D, F^S)$	Forecastable distribution of demand and solar output		Estimated

[2004], but well within their range. Our value of $\overline{p}$ is based on EIA data for Arizona in 2008. To be consistent with the loads projected during the middle part of the life of new generation units, we scale demand quantities by $g = 1.2$, based on historical rates of population and electricity consumption growth in Arizona. The 20% growth yields non-zero investment in new generators for all counterfactuals.

The reservation value can be recovered from (2) using numerical values for elasticity, average price and VOLL. Using mostly customer surveys, Cramton and Lien [2000] report estimates of VOLL that range from \$1,500/MWh to \$20,000/MWh. We choose a fairly conservative estimate of VOLL=\$8,000/MWh. Note that a higher VOLL estimate would imply a higher opportunity cost of system outage and an incentive for the planner to maintain higher reserves. We report the impact of a higher VOLL in sensitivity tests in the last part of the paper.

We estimate F^D, the relationship between day-ahead weather forecasts and load, jointly with F^S, the relationship between day-ahead weather forecasts and solar output. Specifically, we estimate a seemingly unrelated regression (SUR) specification with two dependent variables, Tucson load and solar output.[24] The time scale for data is five minute intervals, for all daytime hours (defined as the hours after sunrise until the hour past sunset) in 2008. We replicate the hourly load dependent variable for each of the 12 periods within the hour. We cluster standard errors at the one hour level. As solar output is zero outside of daytime hours, we estimate a separate regression with just load, for all the other hours in 2008. For

[24]The estimates are equivalent to OLS since we use the same regressors for both dependent variables.

all regressions, the regressors include the day-ahead weather forecasts and other factors that might affect load or solar output such as the day-of-the-week. The large number of observations allows for a flexible functional form for the regressors and hence we use linear splines. For our simulations, we need to predict the joint density of solar output and the demand constant $\overline{D}$ at any hour. Rather than parameterizing the joint density of residuals, we directly simulate from this density in order to predict the joint distribution of solar output and load at any hour. For each data element, we take 20 discrete draws from this distribution for use in the simulation procedure. For a given load level, we recover $\overline{D}$ by inverting the demand equation (1). For the minimum demand constant, $\overline{D}^{min}$, we use the lowest $\overline{D}$ recovered from the 20 discrete draws. We trim the solar output at 0.

A number of studies have constructed the marginal cost of operation for generation units. We follow the approach outlined in Wolfram [1999] and Borenstein et al. [2002]. We compute the marginal cost of a fossil fuel generation unit as the sum of fuel cost per unit plus emissions cost per unit. Fuel cost per unit is the product of the heat rate (MMBTU/MWh) and the cost of fuel (in \$/MMBTU). Emissions cost per unit is the product of the SO_2 emission rate and the average price for SO_2 emission permits available for years 2011 – 2017.

We report summary statistics for existing TEP generators in Table 5. Except for a 5.1 MW solar PV facility in Springerville, AZ, all of TEP's generation units are fossil fuel based. We treat this solar unit as though it were producing constantly at its mean output level of 0.756 MWh. We believe that the bias from not modeling the output of this unit more accurately will be small, given its relatively small size.

Table 5 also lists characteristics of potential new generators. We use a relatively small generator capacity size of $k^{FF} = 60$ MW, as the small size is close to the average size of 51.3 MW for TEP's gas generators and hence likely reflects the optimal generator size for a relatively small market such as southeastern Arizona. We assume that the other characteristics of new generators are the same as for TEP's existing combined cycle gas generators, which are all at the Luna Energy Facility. This facility began operating in 2006 and reflects modern technology.

Table 6 lists the remaining supply parameters. The solar capacity cost includes the

Table 5: Summary Statistics for TEP Generators, 2008

Unit Type	# Units	Mean Size (MW)	Mean MC $/MWh	Mean NOx (lbs./MWh)	Mean SO2 (lbs./MWh)	Mean CO2 (lbs./MWh)
Solar PV	1	0.756 (--)	0 (--)	0 (--)	0 (--)	0 (--)
Coal	10	155 (138)	22.25 (4.72)	4.07 (1.10)	1.95 (1.54)	2,337 (448)
Natural Gas – Combined Cycle	3	62 (20.7)	40.2 (0)	0.09 (0)	0.01 (0)	894 (0)
Natural Gas – Steam Turbine	3	89.0 (13.9)	58.9 (0)	3.33 (0)	0.01 (0)	1,433 (0)
Natural Gas – Gas Turbine	7	30.5 (18.5)	154.0 (107.9)	3.58 (1.76)	0.05 (0.02)	1,964 (136)
Potential New Natural Gas - Combined Cycle	By eqm.	60 (0)	40.2 (0)	0.09 (0)	0.01 (0)	894 (0)

Note: Standard deviations in parentheses. MC figures include emissions permits.

expected discounted present value of costs for maintenance including inverters over the life of the unit. We compute the ratio of the hourly reserve marginal cost to the hourly generation marginal cost, c^s, using data from the deregulated ERCOT market on the 2008 prices in the up-regulation and responsive reserve markets, which pay firms in exchange for giving ERCOT the option to force them to operate with short notice.[25] If they operate, they receive the price on the balancing market. The average price is \$65.41/MWh in the balancing (production) market; \$27.05 in the responsive reserve market; and \$22.71 in the up regulation market. The average of the ratio of the responsive reserve market to balancing market prices over all hours is 0.42, while the average of the ratio of the up regulation to balancing market prices over all hours is 0.40. Our estimate of the reserve costs is the average of these two numbers.

We estimate transmission cost savings from solar as follow. The Arizona RPS states that 30% of solar energy must be generated in a distributed environment which motivates our choice of d^{solar}. For line losses, we find the value of α that matches TEP's reported line losses of 6.6% of 2008 load, using $LL(Q)$ to calculate line loss as a function of α in

[25] We obtained the data from ERCOT [2011]. The up regulation market gives firms 3 to 5 seconds to adjust production while the responsive reserve market gives 10 minutes.

Table 6: Remaining supply parameters

Param.	Interpretation	Value	Source
d^{outage}	system outage hours times percent of affected customers	0.98	EIA
d^{solar}	Fraction of solar generation that is distributed	0.3	Arizona RPS
FC^{FF}	New gas generator capital cost per MW	$984,000	EIA
FC^{solar}	Solar capital cost per rated MW	$5,000,000	EIA
c^{s}	Ratio of MC for spinning reserves to production MC	0.41	Calculated from ERCOT data
α	Line loss constant	0.000035	Calculated fromTEP Form 10K
AFC^{T}	Average transmission fixed cost	$19.6	Borenstein and Holland [2005], Baughman and Bottaro [1976], TEP line loss cost
β	Discount factor	0.94	
T	Lifetime of generators in years	25	

any period and then summing across periods in 2008 to get total line losses as a function of α. To calculate AFC^T, we approximate average generating costs by $70/MWh, so that the 6.6 percent line loss represents $4.62/MWh. We difference the line loss from the total transmission and distribution (T&D) cost of $40/MWh from Borenstein and Holland [2005] to obtain an average T&D fixed cost of $35.38/MWh. Using information in Baughman and Bottaro [1976] we calculate that 55.3% of T&D fixed cost can be attributed to transmission. We obtain AFC^T in (4) by multiplying 55.3% of $35.38 by the discounted sum of expected load and then dividing by the maximum expected load.

To estimate d^{outage} we use major disturbances reported by the EIA whose causes were due to equipment failure (not, for example, due to storms) that impacted more than 50,000 customers. For 2008 there were 21 such disturbances for which we could find both the total number of customers and the number of affected customers, from which we calculate the percent of customers affected. For each of the 21, we multiplied the duration by the percent of customers affected. We estimate d^{outage} as the mean of this product.

4.3 Computation of Planner's Problem

We compute solutions to the planner's problem using the estimated and calibrated model parameters. We assume that the distribution of forecasted load for TEP remains constant at its 2008 level over time, adjusted by the 20% growth factor. We proceed by maximizing the DPV of welfare over the first stage decisions of the number of new gas generators and the curtailment price, taking as given the retail price of electricity and the solar output level. For each first stage decision vector, we compute the optimal policy for each second stage period, and the value that results from this optimal policy. The computation of the first stage involves a grid search over n^{FF}. For each value of n^{FF}, we search over p_c using the simplex method.

To compute the second stage optimal policy, we make two assumptions to ease the computational burden that we believe will not significantly bias the results. First, we assume that the planner schedules generators in ascending order of MC when computing optimal generation for a second-stage period. Although this point is intuitively reasonable, because of differences in sizes and failure probabilities across generators, it is possible that a planner would want to schedule a higher MC generator and not a lower MC one. Second, we assume that the planner curtails demand only if all available generators for which MC is below the marginal cost of curtailment, $dWLC(z)/dz$, are scheduled. Again, this point is intuitively reasonable but may not hold exactly because generators come in discrete chunks.

We now discuss our computation of the second-stage policies. At each second-stage period, we condition on the state (w, m), which encapsulates the units with planned outages and the joint forecastable distribution of load and solar generation. We then choose the production and curtailment decisions, integrating over remaining sources of uncertainty (forced outages and the realization of load and solar generation given the forecastable distribution) in order to solve for the probability of system outage and the associated expected welfare. We then maximize expected welfare over these choices. Finally, we integrate over the three ex-ante decisions to obtain the expected welfare associated with any first stage policy.

We perform the integration using simulation. Specifically, we use 20 discrete draws to

integrate over the joint distribution of load and solar generation conditional on a forecast. Each of the 20 draws includes one hourly load draw and 12 5-minute solar output draws. Note that the planner's problem also involves simulation of generator failures. Failure probabilities for individual generation units are small, and probabilities of multiple failures – which might cause a system outage – are very small, but the adverse consequences of a system outage are very large. Thus, our computation is challenging because integration using a direct simulation method would be very inefficient. Instead, for each type of generator, we integrate over the probability of a given number of failures given a total number of generators operating, and then simulate the identity of failed generators conditional on the number of failures. Similarly, at the first stage, we need to integrate over the distribution (w, m). We integrate over the forecastable weather distribution by simulating with replacement from the observed distribution and over generator scheduled maintenance with an analogous method to our simulation for sudden generator failure.

5 Results

5.1 Forecast Estimation Results

The estimated relationship from the SUR model of daytime load and solar output on weather forecasts is reported in Table 7. The unit of observation is the 5-minute level although load and all regressors are the same within an hour. Standard errors are clustered at the hour level. We estimate splines for each regressor. For cloud coverage, the knots of the splines correspond to the categorical cloud cover variable in weather forecasts. For other forecast variables, we use 10 splines where the knots are the deciles of the distribution. We report coefficients on the lowest, median and highest levels. We also include month, hour and day-of-week dummies, as well as interactions of cloud cover with other variables.

We find a U-shaped relation between forecasted temperature and load, as electricity is needed for both heating and cooling. Another important predictor for load is relative humidity, where the relation is inverse U-shaped. On the other hand, the coefficients of temperature on solar output suggest that increases in temperature in the upper deciles have no significant impact on solar output. Forecasted cloud cover variables have negative signs

and of increasing absolute value on solar output, as expected. Hours since sunrise before noon and hours until sunset after noon are also both strong positive predictors of solar output. The R^2 is 0.967 for load and 0.831 for solar output, suggesting that both levels are highly, though not perfectly, forecastable. The correlation in the residuals between load and solar output is 0.077 and statistically significant ($\chi^2(1) = 298, P < 0.01$). The nighttime impact of weather forecast on load is reported in Table 8. Temperature is an important predictor for nighttime demand as are hourly dummies.

The outage probabilities for gas and coal generators are reported in Table 9. Note that coal generators report a higher rate of sudden failure (0.123%) than do gas generators (0.054%).

5.2 Equilibrium Costs of Solar RPS Policies

Table 10 reports equilibrium computational results using the estimated and calibrated parameters, gross of the benefit from reduced CO_2 emissions (which we address below in Section 5.4). The first column reports results with no solar PV investment and other columns progressively add higher RPS requirements up to 30%.[26]

Without solar PV investment, the planner chooses 32 new natural gas generation units. Demand curtailment accounts for 0.8% of operating reserves although at peak times such as July at noon, the probability of some demand curtailment is over 13 percent. On average over all hours, operating reserves are 24.2 percent of load. The 24.2 percent figure for operating reserves appears to be higher than average actual reserves for many systems. Two factors might account for this. First, there may be room for improvement in our forecasting model for load and solar. A better forecasting model (with additional explanatory variables and/or a different specification) could yield a lower variance for forecast errors and lead to lower optimal operating reserves. Second, TEP has a relatively small number of generation units and its largest units comprise a significant fraction of load. For instance, TEP would need to have operating reserves amounting to 19% of average load to replace the output of its two largest coal units.

The second column of numbers in Table 10 reports results for a solar RPS of 10% of load.

[26] Arizona specifies a 15% RPS by 2025 while California specifies a 33% RPS by 2020.

Table 7: Estimation of Daytime Load and Solar Output Forecasts

	Load (MWh)			Solar output (Wh)		
	Coefficient on spline for			Coefficient on spline for		
	1^{st}decile	5^{th}decile	10^{th}decile	1^{st}decile	5^{th}decile	10^{th}decile
Temperature	–19.4**	5.3*	48.6**	12.9**	2.9	–0.6
	(2.6)	(2.3)	(2.3)	(4.5)	(4.1)	(2.9)
Dew point	–3.8	2.8	-3.5	–1.4	–23.3**	–7.9
	(2.7)	(3.1)	(3.8)	(3.8)	(5.8)	(5.6)
Relative humidity	27.8**	3.9	–4.4**	6.4	4.8	2.4
	(7.2)	(2.6)	(1.3)	(10.0)	(4.3)	(2.4)
Wind	–9.4*	–10.3	–3.6**	53.7**	11.6	6.0**
	(3.9)	(5.3)	(0.9)	(12.3)	(9.6)	(1.8)
	2–15%	38–60%	78–94%	2–15%	38–60%	78–94%
Cloud cover	28	–17	174	–731**	–1514**	–2718**
	(103)	(98)	(140)	(244)	(268)	(607)
	Coefficient on hour			Coefficient on hour		
	1	4	6	1	4	6
Hours since sunrise, AM	–36.3**	–2.2	17.3	179.5**	133.9**	112.9**
	(13.9)	(10.1)	(13.9)	(23.4)	(24.3)	(29.6)
Hours till sunset, PM	–154**	37**	26*	–289**	171**	110**
	(20)	(11)	(12)	(47)	(19)	(27)
Temp × cloud	–0.09	(1.6)		7.8*	(3.4)	
RH × cloud	3.5*	(1.5)		1.9	(3.6)	
Wind × cloud	3.1*	(1.4)		–12.1**	(3.9)	
Dew × cloud	–3.9*	(1.9)		6.4	(4.0)	
6AM dummy				–541**	(88)	
...						
12PM dummy	261**	(48)		–107**	(24)	
...						
6PM dummy	33*	(17)		–276**	(77.187)	
R–squared	0.967			0.831		
Correlation of residuals			0.077**			

Note: Model estimated with a SUR specification. Unit of observation is 5 minute interval in 2008. Standard errors are clustered at hour level. Number of observations is 50,124. We include as regressors day-of-week and month-of-year indicators and full sets of spline coefficients.
** Statistically significant at 1% level
* Statistically significant at 5% level

Table 8: Estimation of Nighttime Load Forecast

	Load (MWh)		
	Coefficinet on spline for		
	1st decile	5th decile	10th decile
Temperature	–14.7**	10.3**	55.0**
	(3.3)	(3.1)	(4.9)
Dew point	–3.6	–5.0	–4.7
	(3.8)	(3.7)	(3.4)
Relative humidity	27.3**	3.9	–1.1
	(10.3)	(3.1)	(1.2)
Wind	–7.6*	–3.8	–9.5**
	(3.5)	(3.8)	(1.9)
	2–15%	38–60%	78–94%
Cloud cover	228*	123	-38
	(101)	(98)	(129)
Temperature × cloud cover	–6.6**	(1.9)	
Relative humidity × cloud	–1.2	(1.2)	
Wind × cloud cover	9.3**	(1.5)	
Dew point × cloud cover	5.2*	(2.0)	
9PM dummy	272**	(5.1)	
...			
3AM dummy	–58.0**	(3.7)	
R–squared	0.959		

Note: Model estimated with OLS. Unit of observation is 5 minute interval in 2008. Standard errors are clustered at hour level. Number of observations is 48,852. We include as regressors day-of-week and month-of-year indicators and full sets of spline coefficients.
** Statistically significant at 1% level
* Statistically significant at 5% level

Table 9: Average Hourly Outage Probabilities

	Forced outage probability	Planned outage probability	Avg. number of units over period
Natural gas generator	0.0535%	0.0395%	358
	(0.122%)	(0.105%)	
Coal generator	0.123%	0.049%	866
	(0.119%)	(0.075%)	

Note: The time period covered is 2005-2009. Standard errors in parentheses.

Table 10: Outcomes with Different RPS Levels

RPS Policy	0%	10%	15%	20%	30%
Solar PV capacity (MW)	0	846	1,269	1,692	2,538
Solar production (1000 MWh/year)	0	1,701	2,551	3,401	5,102
Load (1000 MWh / year)	17,006	17,006	17,006	17,006	17,006
New 60MW natural gas generators (#)	32	29	29	28	27
Foregone new gas generators (#)	–	3	3	4	5
Scheduled non-solar prod. + res. (1000 MWh/year)	22,942	21,817	21,438	21,081	20,422
Realized non-solar prod. + reserves (1000 MWh/year)	22,924	21,799	21,420	21,064	20,404
Reserves as % of power consumed	24.2%	28.0%	30.8%	33.9%	40.2%
Average system outage prob.	4.45e-5	5.29e-5	5.26e-5	6.11e-5	6.94e-5
Curtailment price p_c ($/MWh)	408	577	551	717	722
Total curtailment quan. (MWh/year)	33,485	27,033	21,771	25,453	30,804
Prob. of some curtailment Jul. 12PM	13.0%	0.04%	8.0e-5%	7.2e-7%	8.9e-8%
Prob. of some curtailment Jul. 6PM	13.0%	25.9%	29.1%	22.7%	29.3%
Production costs (million $/year)	549.2	481.6	450.7	422.0	375.6
Reserve costs (million $/year)	76.4	85.2	91.3	97.5	106.9
Gas generator investment costs (mil. $)	1,889	1,712	1,712	1,653	1,594
Solar capacity investment costs (mil. $)	0	4,229	6,344	8,458	12,688
Transmission FC (million $/year)	332.7	326.0	324.5	323.0	319.9
Transmission line losses (1000 MWh/year)	1,482	1,386	1,342	1,298	1,217
Loss in surplus relative to baseline (million $/year)	–	236.1	364.4	496.2	768.0
Loss in surplus per unit solar production ($/MWh)	–	138.8	142.9	145.9	150.5
NO_x emissions (1000 metric tons / year)	20.9	20.5	20.1	19.6	18.4
SO_2 emissions (1000 metric tons / year)	14.3	13.9	13.7	13.2	12.5
CO_2 emissions (mill. metric tons / year)	18.9	18.0	17.4	16.9	15.8

This output level would require 846 MW of solar PV capacity, with an investment cost of \$4.4 billion. The solar PV panels would yield roughly 1.7 million MWh per year, which represents a capacity factor (i.e., average output as a percent of capacity) of 23%. Optimal new fossil fuel generators falls by 3, reducing fossil fuel generation capacity by 180 MW. This yields a capital cost offset of \$177 million (about 4% of solar investment cost). Optimal operating reserves rise compared to the no-solar case, from 24.2% to 28.0% of electricity consumed.

Under the no RPS case, the planner chooses a curtailment price of \$408/MWh, which is 4.3 times the retail price.[27] With the 10% RPS, the optimal curtailment price rises, as it is optimal to be able to curtail more demand in periods of low solar output and high demand. The times when curtailment occur shift as a result of the RPS. For instance, the probability of demand curtailment at noon in July drops, as solar output is high, but the probability of curtailment at 6PM in July – when solar output is low but load is still high – rises. Overall, total curtailment is lower than without solar.

Interestingly, the system outage probability rises by only a modest amount as the RPS standard is increased. The overall impact of a 10% RPS standard is to reduce expected welfare over the life of the units by about \$236 million per year, gross of the value of CO_2 emissions reduction. Dividing by solar production, the reduction in welfare is \$138.8/MWh of electricity produced, again gross of the CO_2 emissions reduction. Put differently, the net welfare cost of the solar mandate is approximately 75 percent of the \$4.2 billion investment cost for solar PV capacity. These results factor in the value of SO_2 emission reductions to the extent that SO_2 permit prices reflect marginal environmental damages.

Columns 3 through 5 in Table 10 report results for solar RPS policies of 15%, 20%, and 30% of load, respectively. There is little or no offset in fossil fuel capacity investment as the RPS is increased above 10% but otherwise, the welfare and cost results move in the same direction as the change from 0 to 10%. Because of the lack of fossil fuel capacity offset and the fact that solar generation will increasingly substitute from low cost fossil fuel generators, the welfare loss per MWh of solar generation rises monotonically from \$138.8 to \$150.5 as

[27] Baldick et al. [2006] note that compensation per MWh for curtailed demand in interruptible power contracts ranges from about 1.5 to 6 times higher than average retail price.

the RPS increases from 10% to 30%.

Comparisons between solar PV and conventional generation are often based on levelized cost, which is the average cost over the life of the unit.[28] The realized solar output and our assumption about the cost of solar panels together yield a levelized cost of $184/MWh for solar PV generation. The levelized cost for a new combined cycle generation unit is $58/MWh.[29] Thus, on the basis of a simple average cost comparison, solar PV imposes an additional per unit cost of $126/MWh. The levelized cost difference does not account for the time-profile of generation or endogenize the choice of planner policies in response to a solar mandate.

Borenstein [2008] makes the point that valuing solar PV generation using wholesale prices at the time of generation narrows the gap between solar PV and conventional generation due to the time-profile of solar generation. Our analysis takes into account the value of solar generation at different times of day and in different seasons, just as in Borenstein [2008]. However, in the case of large-scale renewable mandates, it is also important to consider the impact of the mandate on optimal system-wide policies, including operating reserves, backup fossil fuel capacity and demand-side management through curtailment contracts.[30]

Our analysis endogenizes these three factors and finds that, together with the Borenstein effect, they imply that, for a 20% RPS standard, solar is $19.9/MWh more costly than the levelized cost difference. In contrast, if the planner did not re-optimize in response to a 20% RPS (and instead used the same operating reserve quantity, backup fossil fuel capacity, and curtailment quantity and price), we calculate that the mandate would have a welfare effect of $1,378/MWh. This very high figure is due to the fact that the 20% RPS would then result in a 3% probability of system outage. Thus, with sub-optimal policies, the costs of renewable energy can indeed be much higher.

Table 11: Costs Associated with 20% RPS

RPS policy	Foregone new gas generators	Loss in surplus per MWh solar
Feasible solar	4	$145.9
Solar cost drop from $5 to $2/W	4	$35.5
No unforecastable variance	7	$133.4
Equal generation profile	7	$131.4
Fully dispatchable	27	$89.8
VOLL increased to $12,000	3	$146.2
No demand curtailment	0	$149.4
Rule of thumb policy	1	$677.2

"Equal generation profile" is a hypothetical solar facility which produces equally at every hour.
"Fully dispatchable" is a hypothetical solar facility which can be dispatched based on the demand forecast.
"No unforecastable variance" is a hypothetical solar facility that produces at the forecastable mean.

5.3 Components of Equilibrium Costs for Solar

Table 11 evaluates further the causes of the equilibrium costs of solar for the 20% RPS case. Each row corresponds to one scenario (which was column 5 in Table 10). We omit most of the detail from Table 10, presenting only two columns for each experiment: the foregone new gas generators and loss in surplus.

For reference, the first row repeats the baseline 20% case from Table 10. The second row simulates dropping the cost of solar from $5 to $2/W as many industry observers believe will occur. Note that optimal policies for the system operator remain unchanged with this drop in capacity cost; the reduction in surplus loss per MWh of solar is due only to reduced up-front capacity cost. The loss in surplus from solar drops dramatically in this case, from $145.9 to $35.5.

The next four rows examine the impact of different components of intermittency on the equilibrium costs of solar. First, we find that eliminating the unforecastable component of solar output raises value by $12.5/MWh relative to the base case, from $145.9 to $133.4. The planner foregoes 3 extra generators in this case. This drop is small compared to the overall additional equilibrium cost of solar generation and is less than drops reported in the

[28] See EIA [2011]

[29] See EIA [2011].

[30] The envelope theorem implies that it is not necessary to re-optimize policies with only a small amount of solar on the grid.

literature.[31] This may be in part because other studies have not fully endogenized policies in response to renewable energy mandates. Relatedly, utilities often express concern over the high cost of intermittency with large-scale solar. We believe that this may be because utilities do not fully understand the optimal policies in the case of large-scale solar.

Second, we find that a hypothetical solar facility that always produced at its mean output level would be better than the baseline – with a loss of \$131.4 instead of \$145.9/MWh. The planner foregoes 3 extra generators in this case. Moreover, the equal generation case would be better than even the case without unforecastable intermittency, by \$2.0. Thus, even though solar facilities tend to produce during peak periods, even the forecastable part of their intermittency reduces their value.

Third, we examine the value of solar if its energy could be dispatched at the times with the highest demand, as might occur with perfect storage mechanisms. In this case, the value of solar would rise by \$56.1/MWh relative to the base case – a substantial amount, but much less than the over \$100 rise from reducing solar capacity costs to \$2/W.

The next two rows test the robustness of our findings to the demand curtailment specification and VOLL measures, respectively. A 50% increase in VOLL has virtually no effect on the value of solar PV. However, if the system operator is constrained to zero demand curtailment then the optimal number of new fossil fuel generators decreases by three instead of four from the no solar to the solar case. In effect, the system operator replaces demand curtailment with the additional generators. Eliminating demand curtailment reduces the value of solar by \$3.5/MWh.

The final row of Table 10 compares the optimal policy with a rule of thumb policy similar to what has been proposed in some studies of renewable intermittency.[32] Our rule of thumb policy starts with the no-solar policy and modifies it in two ways. First, we reduce fossil

[31] For 20% wind power penetration in Great Britain, Skea et al. [2008] calculate that the back-up generation capacity required to address intermittency would add roughly 15% to the cost of wind generation. Hoff et al. [2008] estimates that back-up generation capacity required to address solar intermittency would represent about 15% of solar PV average cost. These studies do not consider costs of additional operating reserves that might be required.

[32] Madaeni and Sioshansi [2011] examine the impact of changes in wind forecast accuracy. Operating reserves are constrained to be greater than or equal to a fixed fraction of load throughout their analysis. Mills and Wiser [2010] examine how operating reserve costs change depending on the degree of variability of solar generation. They set operating reserves equal to a fixed percentile of the distribution of period-to-period changes in solar generation.

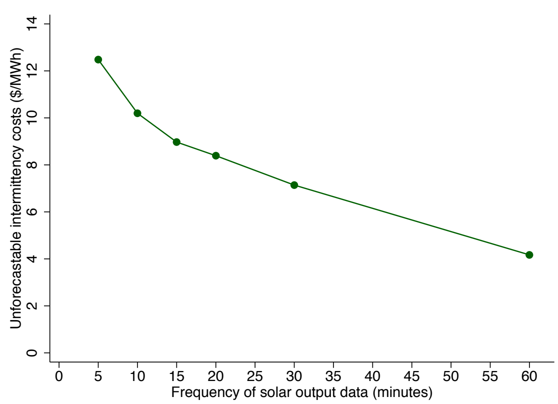
fuel capacity by a "capacity credit" for solar PV capacity of 388 MW, which is equal to the solar PV capacity factor (23%) times solar capacity of 1,692 MW. Second, we increase operating reserves by a fixed amount to address solar intermittency, with a corresponding addition to fossil fuel capacity. We adopt the heuristic for operating reserves proposed by Mills and Wiser [2010], who suggest increasing reserves by three standard deviations of the change in solar from one period to the next.[33] The welfare cost of this rule of thumb policy is \$482.5/MWh; 2.8 times higher than the welfare cost associated with optimal policies. This comparison illustrates the importance of policy optimization.

Figure 3 shows how the intermittency costs decrease as the time scale for solar output observations increases. As reported in Table 11, the 5 minute data imply unforecastable intermittency costs of \$12.5/MWh. With 60 minute data, we calculate the costs to be only \$4.2/MWh.[34] Thus, it is the high frequency nature of the solar intermittency that is the most problematic. The results also suggest that data at an even finer level than 5 minutes may give an even higher cost to solar.

[33] We compute the standard deviation of 5 minute changes in solar generation. Three standard deviations amounts to 287 MW. Thus, we add 287 MWh of operating reserves to optimal baseline operating reserves for all daylight hours, and adjust fossil fuel capacity downward by 388 − 287 MW.

[34] Our results on the impact of the solar generation time scale are roughly consistent with results in Mills and Wiser [2010]. Using a heuristic decision rule for operating reserves, they find that operating reserves required to address variability of solar PV generation for a single site at a one minute time scale are three times as costly as operating reserves required to address solar PV generation variability at a one hour time scale. Mills and Wiser [2010] also find that geographic dispersion of solar PV sites significantly reduces intermittency costs.

Figure 3: Unforecastable intermittency costs and frequency of solar output data



5.4 RPS Policies and Benefits from CO_2 Reductions

Finally, we analyze whether RPS policies would increase or decrease social welfare, when one accounts for the reduction in CO_2 emissions that would be caused by the RPS. The policy impact of an RPS depends crucially on two elements: first, on the environmental benefit per unit reduction in CO_2 emissions; and second, on the impact of ongoing R&D in reducing the costs of renewable power generation. Note that we can derive the cost of CO_2 emission reductions under current technological conditions from Table 10. For example, CO_2 is reduced by 2.3 million tons per year at a cost of \$216/ton for the 20% RPS case.

The U.S. government has recently set social cost of carbon values to be used in the regulatory approval process. They specify a range of values corresponding to different assumptions about discount rates, the impact of temperature change, and loss functions. The central value is \$21/(metric) ton in 2007 dollars. We use the four values developed by the U.S. government as well as a baseline of \$0.[35] Using these values, we calculate the "target" cost of solar capacity generation at which the RPS policy would be welfare neutral. The RPS will be welfare increasing if and only if solar capacity costs are lower than the target costs.

Table 12 presents the results, which can be derived without recomputing the model, since solar capital costs enter linearly into welfare. At the current cost of \$5/W, any RPS would reduce welfare even if CO_2 emissions are valued at the highest reported figure of \$65/ton. At this emissions cost, solar capital costs would have to fall to \$2.18 for the 10% RPS to be welfare neutral, and \$2.00 for the 30% RPS to be welfare neutral. As one would expect, the target capital costs are increasing in the value of offset CO_2 emissions. For instance, the 20% RPS welfare neutral capital costs rise from \$1.04 to \$2.10, as CO_2 emissions costs increase from \$0 to \$65.

Less evident is the impact of an increase of an RPS on the welfare neutral capacity cost. On one hand, with a higher RPS, solar capacity will substitute more from lower cost generators, which will decrease its equilibrium value. On the other hand, the lower cost generators will tend to be coal instead of gas, and coal generators emit more than double the

[35]These values appear in EPA [2010]. See also Greenstone et al. [2011] for details on this policy. Tol [2005] provides a survey of estimates of the social cost of carbon and obtains similar numbers.

Table 12: Welfare Neutral Solar PV Capital Costs with Benefits from CO_2 Reductions

RPS Policy Benefit per ton of CO_2 reduction	10%	15%	20%	30%
\$0	1.23	1.12	1.04	0.91
\$5	1.30	1.20	1.12	1.00
\$21	1.54	1.45	1.38	1.26
\$35	1.74	1.67	1.61	1.50
\$65	2.18	2.14	2.10	2.00

Note: solar capital costs are in dollars per rated watt

CO_2 per unit energy output than combined cycle natural gas units (see Table 5), which will increase its value. Under the social optimum, the generation cost effect dominates, but not by very much. For instance, for the central \$21 CO_2 cost case, the welfare neutral capacity costs fall from \$1.54 to \$1.26 from the 10% to 30% RPS cases.

6 Conclusions

A variety of current and potential policies are intended to stimulate investment in renewable energy generation. Intermittency of renewable generation may have a significant impact on electric grid reliability, system operations, and requirements for back-up generation capacity. Because a grid operator must make different long- and short-run decisions in response to intermittent renewable output, we believe that the costs of intermittency can best be understood in the context an optimizing or equilibrium model. Thus, we developed an empirical approach to estimate the equilibrium costs of renewable energy accounting for their intermittent nature. Our approach has three parts: (1) a theoretical model that is based on the work of Joskow and Tirole [2007]; (2) a process to estimate and calibrate the parameters of this model using publicly-available data; and (3) a computational approach to evaluate the impact of counterfactual RPS and other policies. We believe that the biggest limitations of our approach are that we do not allow for dynamic linkages from period to period; that we consider only the social optimum; and that we do not model imports and exports outside the local area. Moreover, other of our assumptions, notably our assumed spinning reserve costs, are at best approximations of reality.

Using our approach, we examined the impact of a renewable portfolio standard (RPS) on

Tucson Electric Power, the public utility that serves southeastern Arizona. We find that the equilibrium cost of a 20 percent solar PV RPS would be $145.9/MWh. Perfectly dispatchable solar energy would lower costs by $56.1/MWh. Unforecastable intermittency accounts for $12.5/MWh and high frequency unforecastable intermittency within an hour is the most costly part of this. If CO_2 reductions are valued at $21/ton, a 20% RPS would be welfare increasing if solar capacity costs dropped below $1.38/W from their current level of $5/W.

We believe that our study has a number of broader implications beyond the results for solar generation in Arizona. First, our finding that the costs of intermittency for solar energy are lower than many industry observers believe may be important. In particular, costs associated with intermittency are a relatively small component of the overall welfare cost of a solar RPS mandate; the bulk of the welfare cost is simply the high installation cost of solar. Our results on intermittency stem from the fact that our approach calculates the costs if utilities use forecasts to optimally schedule reserves, design demand curtailment contracts, and build capacity in response to solar PV mandates. It is possible that utilities need to obtain knowledge about how these decisions should change in the presence of substantial renewable generation, and our study provides a framework that can be used to guide utilities along this dimension.

Second, we believe that our study has implications about the optimality of different potential RPS policies. While we find that an immediate RPS with 2008 technology would reduce welfare, we also find that once solar capacity costs drop below $2.00 or $1.50, solar PV generation becomes welfare increasing. More surprisingly, at this point, capacity costs do not have to drop much further before it is optimal for solar to account for a significant proportion of generation in Arizona.

Finally, we believe that our approach can be used to analyze a variety of other energy policies many of which might also have important equilibrium impacts. These policies include understanding the impact of real-time pricing on reducing GhG emissions and intermittency costs; the relative costs of reducing emissions from an RPS versus a carbon tax; how geographically disparate wind or solar installations might lower intermittency costs; and how technologies such as battery storage and electric cars which change the effective time pattern

of demand can change the value of renewable mandates.

Appendix A Proofs

For Online Publication

Proof of Lemma 3.1

$$VOLL = \frac{\int_{\overline{p}}^{v} D(p,\overline{D})dp + pD(p,\overline{D})}{D(p,\overline{D})} = \frac{\overline{D}(\frac{1}{1-\eta})(v^{(1-\eta)} - \overline{p}^{(1-\eta)}) + \overline{p}\overline{D}\overline{p}^{-\eta}}{\overline{D}\overline{p}^{-\eta}}$$
$$= \frac{\overline{D}(\frac{1}{1-\eta})(v^{(1-\eta)} - \eta\overline{p}^{(1-\eta)})}{\overline{D}\overline{p}^{-\eta}}.$$

Dividing through by $\overline{D}\overline{p}^{-\eta}$, we obtain the expression in the statement of the lemma.

Proof of Lemma 3.2

Let $P(q,\overline{D})$ denote the inverse demand curve. Then, the welfare cost of z is

$$WLC(z,p_c) = \left(\frac{z}{D(\overline{p},\overline{D}) - D(p_c,\overline{D})}\right)\int_{D(p_c,\overline{D})}^{D(\overline{p},\overline{D})} P(q,\overline{D})dq$$
$$= \frac{z\eta(\bar{p}^{1-\eta} - p_c^{1-\eta})}{(\eta-1)(\bar{p}^{-\eta} - p_c^{-\eta})}.$$

Note that $\overline{D}$ drops out of the welfare cost, which depends on the state only through the quantity z of rationing chosen at that state.

Appendix B System Operations

For Online Publication

The electricity system is a multi-nodal network that connects a number of different types of generation plants to load centers (e.g., cities) via high-voltage transmission lines and ultimately delivers power to customers via lower voltage distribution lines. Since storage is very limited on most systems, the supply of power must equal (almost exactly) the demand for power, called load, on a real time basis. Moreover, load can vary unpredictably over the course of a day (e.g., due to weather changes) and available supply can vary quickly and

unpredictably due to equipment malfunction or breakdown and due to intermittent renewable generation. To ensure matching of supply and demand, the manager of an electricity grid engages in "system operations." System operations involve control of generators, decisions about rationing power to customers, and control of backup systems. The system operator insures reliability in part by having generators available on a stand-by basis so that customers can continue to be served in the event that one or more generators fails and/or load exceeds forecast. Operating reserves consist of generation capacity that is scheduled by the system operator over and above the amount required to serve forecasted load. Operating reserves are part of a set of ancillary services used by the system operator to regulate voltage and maintain stability of the system. It is common for ancillary network-support services to require scheduling generation capacity equal to 10-12 percent of load at any point in time.[36]

If available electricity supply is not sufficient to meet demand then a system operator will typically shut off power to some customers or some geographic areas, resulting in a partial blackout of the system. This is what we refer to as a system outage. The system operator initiates an involuntary cut-off of power to some customers, so as to avoid a complete system collapse. A total system collapse is a rare outcome in which demand and generation are shut off over a large area in an uncontrolled fashion.[37] A system collapse should not occur if the system operator responds appropriately to a power shortage by cutting off power to some customers.

In the absence of coordination by a system operator, the operator of a generation unit may impose externalities on other suppliers and on consumers. This is because a power generator may not face the additional cost of being the marginal producer that is causing the system to have to shut out users or, in some cases, completely collapse [see Joskow and Tirole, 2007]. This externality problem is potentially larger with more intermittency problems, suggesting that the role of the system operator may be more important with more renewable energy.

The North American Electric Reliability Corporation (NERC), an industry trade group,

[36] Joskow and Tirole [2007], p. 78.

[37] An example was the 2003 blackout in the Northeast U.S. and Ontario in which 50 million customers lost power [Minkel, 2008]. In this case, a transmission line fault led to deviations in network frequency, causing generators and transmission lines to trip out in a cascading fashion, which led to a blackout over a large area.

has developed a set of standards for safe and reliable operation of the electric grid. These standards cover many aspects of grid operations, including management of operating reserves.[38] NERC standards are aimed at achieving a level of reliability such that a loss of load occurs no more than one day in ten years. NERC Standard BAL-002-0 deals with what is termed "Disturbance Control Performance." This standard dictates the amount of reserve capacity that is to be available in the event of a loss of supply (typically from failure of a generator). Two key provisions of this standard are:

1. The Balancing Authority shall carry at least enough reserve to cover the most severe single contingency (e.g., failure of the largest generation unit in operation).

2. The maximum amount of time permitted for recovery from a disturbance is 15 minutes.[39]

The NERC standards were approved by the Federal Energy Regulatory Commission (FERC) in 2007 and are now mandatory for electric utilities in the U.S.

All electric grids have operating reserves, although there is variation in their management across grids. We have limited information about how TEP manages operating reserves but more information from the Electricity Reliability Council of Texas (ERCOT), which covers most of the state of Texas. ERCOT operates in a deregulated framework in which there is competition both in the wholesale market and among retail service providers. Wholesale electricity service is traded via bilateral contracts and in an energy balancing spot market. However, even in ERCOT's deregulated framework, there is a system operator that is responsible for managing operating reserves so as to maintain reliability.

The ERCOT system operator runs auctions to procure operating reserves from generation suppliers for several categories of reserves. ERCOT utilizes four main types of ancillary services [see Baldick and Niu, 2005]: (1) Up Regulation Service; (2) Responsive Reserve Services; (3) Non-spinning Reserves; and (4) Down Regulation Service. The first three of these services pay firms in exchange for giving ERCOT the option to force them to operate

[38] See, NERC [2011].

[39] The recovery period is defined as the amount of time it takes to return the area control error to the minimum of zero and its pre-disturbance value.

with short notice. If they are forced to operate, they then receive the market price on the balancing market. These three services differ mostly in the length of time which they have to increase production. The shortest is the Up Regulation, which allows firms 3 to 5 seconds to adjust production, and the longest is non-spinning reserves, which allows an hour to adjust. Down Regulation service pays firms that are operating generation units for giving ERCOT the option to reduce their rate of generation. ERCOT would exercise this option when demand is lower than expected. ERCOT conducts these ancillary service markets one day ahead and operates one auction for each service category for each hour.

Appendix C Electricity Provision in southeastern Arizona

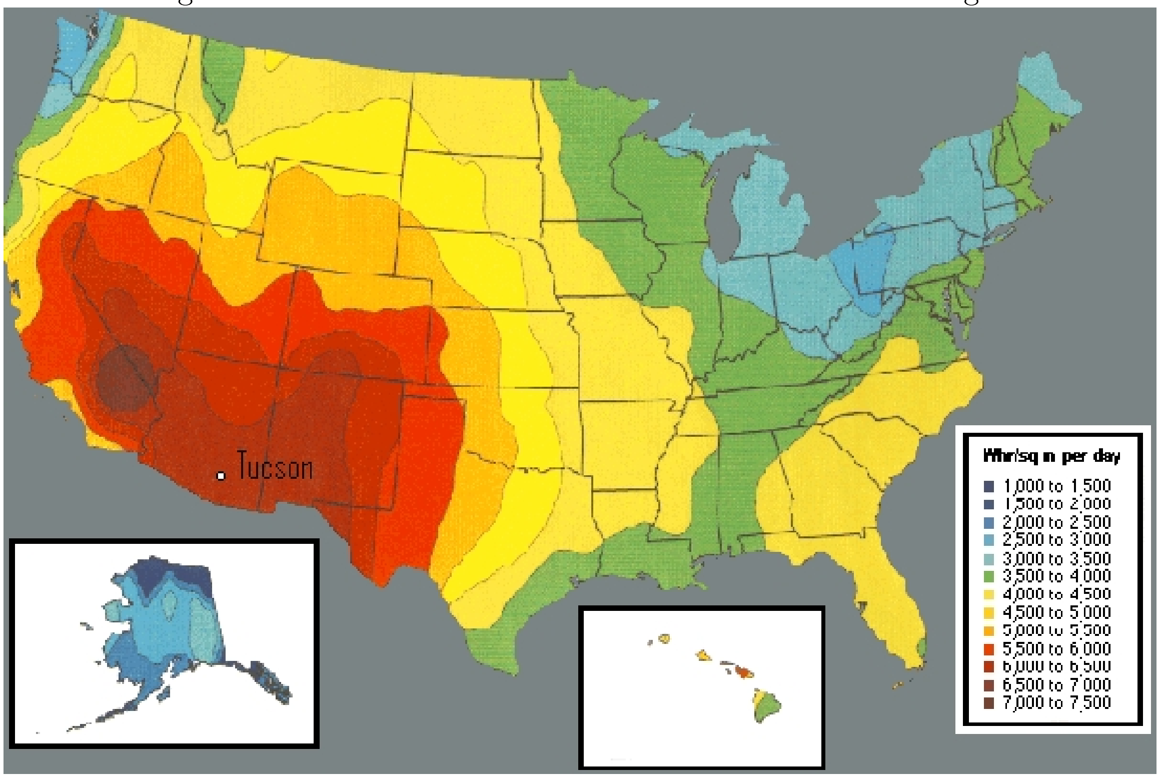
For Online Publication

Most people in southeastern Arizona live in the Tucson metropolitan area, which is one of the best locations in the U.S. for solar electricity generation, as evidenced by the solar radiation map in Figure 4. Electricity service is provided by Tucson Electric Power (TEP), a vertically integrated, investor-owned utility that is regulated by the Arizona Corporation Commission (ACC). TEP's service territory covers 1,155 square miles and includes a population of approximately one million in the greater Tucson metropolitan area.[40] Retail energy consumption by customer class in 2008 was distributed as follows: 41 percent residential, 21 percent commercial, and 38 percent industrial and public. Copper mining is the largest industrial user of electricity, accounting for about one-third of industrial consumption. Tucson is a summer peaking system, with very hot summers and high usage of air conditioning. The highest load in 2008 was 3,063 MWh for 3-4 p.m., August 1.

Tucson is situated within the Western Interconnection, the electrical grid that encompasses the Western U.S. and part of Western Canada. TEP is responsible for system operations and for scheduling generation and transmission power flows within its balancing authority area, which covers most of southeastern Arizona. At different times, TEP both

[40] Detailed information about TEP customers and operations are found in the 2008 10-K annual report for UniSource Energy Corp., TEP's parent company; [see UniSource, 2008].

Figure 4: Photovoltaic Solar Resource: Source www.NREL.gov



imports and exports power over the Western Interconnection. As of the end of 2008, TEP owned or leased generation units with total capacity of 2,222 MW. This capacity is virtually all powered by fossil fuel.[41] The primary sources for data for TEP generators are described in Section 4.1. Most of these generators are wholly owned and controlled by TEP. However, TEP has a partial ownership stake in the Luna Energy natural gas plant and in the Navajo, Four Corners, and San Juan coal plants [see TEP's FERC Form 714 filing and UniSource, 2008]. For our analysis, we specify the generation capacity for each unit at a jointly owned plant as total unit capacity times TEP's plant ownership share.

TEP is subject to a Renewable Portfolio Standard (RPS), mandated by the ACC, which calls for an increasing fraction of load to be generated from renewable sources until 15 percent of load is from renewables by 2025. For 2008 the RPS was 1.75 percent. TEP satisfies the

[41] Other utilities in Arizona own and operate non-fossil fuel generation plants. The Salt River Project has several hydroelectric plants. Arizona Public Service operates the nation's largest nuclear generator, Palo Verde. There is some wind generation in Arizona. However, wind is not expected by be a major source of renewable generation in the state.

RPS through a combination of its own solar PV generation, wholesale purchases of renewable energy, distributed solar generation by its customers, and retirement of banked renewable energy credits. Many TEP customers have solar PV panels at their business or residence. However, total distributed solar PV capacity in TEP's service territory was only 2.7 MW as of the end of 2008.[42]

References

Adrian, T. and Brunnermeier, M. (2010). CoVaR. Working Paper.

Baldick, R., Kolos, S., and Tompaidis, S. (2006). Interruptible electricity contracts from an electricity retailer's point of view: Valuation and optimal interruption. *Operations Research*, 54:627–642.

Baldick, R. and Niu, H. (2005). *Lessons Learned: The Texas Experience*, pages 182–226. University of Chicago Press.

Baughman, M. and Bottaro, D. (1976). Electric power transmission and distribution systems: Costs and their allocation. *IEEE Transactions on Power Apparatus and Systems*, 95(3):782–790.

Bohn, R., Caramanis, M., and Schweppe, F. (1984). Optimal pricing in electrical networks over space and time. *RAND Journal of Economics*, 15:469–493.

Borenstein, S. (2008). The market value and cost of solar photovoltaic electricity production. CSEM Working Paper 176, University of California, Berkeley.

Borenstein, S., Bushnell, J., and Wolak, F. (2002). Measuring market inefficiencies in California's restructured wholesale electricity market. *American Economic Review*, 92:1376–1405.

Borenstein, S. and Holland, S. (2005). On the efficiency of competitive electricity markets with time-invariant retail prices. *RAND Journal of Economics*, 36:469–493.

[42] TEP [2009].

Bouffard, F., Galiana, F., and Conejo, A. (2005). Market-clearing with stochastic security - part I: Formulation. *IEEE Transactions on Power Systems*, 20(4):1818–1826.

Campbell, A. (2011). Government support for intermittent renewable generation technologies. Working Paper, Yale University School of Management.

CME (2011). Henry hub natural gas futures. `http://www.cmegroup.com/trading/energy/natural-gas/natural-gas.html`. Last accessed February 6, 2011.

Cramton, P. and Lien, J. (2000). Value of lost load. Working Paper, University of Maryland.

Cullen, J. (2010a). Dynamic response to environmental regulation in electricity markets. Working Paper, Harvard University.

Cullen, J. (2010b). Measuring the environmental benefits of wind generated electricity. Working Paper, Harvard University.

Denholm, P. and Margolis, R. (2007). Evaluating the limits of solar photovoltaics (PV) in traditional electric power systems. *Energy Policy*, 35:2852–2861.

EIA (2010). Electric power monthly, tables b.1 and b.2. `http://tonto.eia.doe.gov/ftproot/electricity/epm/02261003.pdf#page=141`. Last accessed February 11, 2011.

EIA (2011). Levelized cost of new generation resources in the annual energy outlook 2011. `http://www.eia.doe.gov/oiaf/aeo/electricity_generation.html`. Last accessed February 13, 2011.

EPA (2010). Technical support document – social cost of carbon for regulatory impact analysis. `http://www.epa.gov/oms/climate/regulations/scc-tsd.pdf`. Last accessed August 21, 2010.

EPA (2011). NOx budget trading program, compliance and environmental results, 2007. `http://www.epa.gov/airmarkt/progress/docs/2007-NBP-Report.pdfl`. Last accessed February 1, 2011.

ERCOT (2011). Day-ahead ancillary services market clearing prices for capacity. http://www.ercot.com/mktinfo/prices/mcpc. Last accessed January 29, 2011.

Espey, J. and Espey, M. (2004). Turning on the lights: A meta-analysis of residential electricity demand elasticities. *Journal of Agricultural and Applied Economics*, 36:65–81.

GE Energy (2008). Analysis of wind generation impact on ERCOT ancillary services requirements. Report for Electric Reliability Council of Texas.

Glennon, R. and Reeves, A. M. (2010). Solar energy's cloudy future. *Arizona Journal of Environmental Law & Policy*, 1(1):91–137.

Greenstone, M., Kopits, E., and Wolverton, A. (2011). Estimating the social cost of carbon for use in U.S. federal rulemakings: A summary and interpretation. Working Paper 16913, National Bureau of Economic Research.

Hansen, T. N. (2008). Utility solar generation valuation methods. Department of Energy, Solar America Initiative Project.

Helman, U., Rothleder, M., Loutan, C., Ventakaraman, S., and Guo, T. (2011). Integration of wind and solar energy in the California power system: Results from simulations of a 20% renewable portfolio standard. Working Paper.

Hoff, T., Perez, R., Ross, J., and Taylor, M. (2008). Photovoltaic capacity valuation methods. Solar Electric Power Association, SEPA Report no. 02-08.

Joskow, P. (2010). Comparing the costs of intermittent and dispatchable electricity generating technologies. Working Paper, Center for Energy and Environmental Policy Research, Massachusetts Institute of Technology.

Joskow, P. and Tirole, J. (2007). Reliability and competitive electricity markets. *RAND Journal of Economics*, 38:60–84.

Krishna, V. (2010). *Auction Theory*. Elsevier, second edition.

Madaeni, S. and Sioshansi, R. (2011). The impacts of stochastic programming and demand response on wind integration. working paper.

Mills, A. and Wiser, R. (2010). Implications of wide-area geographic diversity for short-term variability of solar power. *Ernest Orlando Lawrence Berkeley National Laboratory*. Working paper no. LBNL-3884E.

Minkel, J. (2008). The 2003 northeast blackout – five years later. *Scientific American*, 11.

NERC (2011). Reliability standards for the bulk electric systems of North America. `http://www.nerc.com/files/Reliability_Standards_Complete_Set.pdf`. Last accessed January 15, 2011.

NOAA (2011). Nndc climate data online. `http://www7.ncdc.noaa.gov/CDO/cdo`. Last accessed February 2, 2011.

NREL (2011). Solar photovoltaic technology. `http://www.nrel.gov/learning/re_photovoltaics.html`. Last accessed February 1, 2011.

SECO (2011). Energy efficiency: Texas' newest energy resource. `http://www.seco.cpa.state.tx.us/publications/renewenergy/solarenergy.php`. Last accessed February 8, 2011.

Skea, J., Anderson, D., Green, T., Gross, R., Heptonstall, P., and Leach, M. (2008). Intermittent renewable generation and the cost of maintaining power system reliability. *IET Gener. Transm. Distrib.*, 2(1):82–89.

Sunrise (2011). Sunrise and sunset in tucson. `http://www.timeanddate.com/worldclock/astronomy.html?n=393`. Last accessed January 15, 2011.

TEP (2009). Tucson Electric Power 2009 renewable energy standard and tariff compliance report. `http://images.edocket.azcc.gov/docketpdf/0000095556.pdf`. Last accessed January 15, 2011.

TEP (2011). TEP solar test yard data page. `http://www.physics.arizona.edu/~cronin/Solar/TEPweb/`. Last accessed February 8, 2011.

Tol, R. (2005). The marginal damage costs of carbon dioxide emissions: an assessment of the uncertainties. *Energy Policy*, 33:2064–2074.

UniSource (2008). 10-K report for 2008. `http://ir.uns.com/secfiling.cfm?filingID=1362310-09-2982`. Last accessed January 15, 2011.

Wolfram, C. (1999). Measuring duopoly power in the British electricity spot market. *American Economic Review*, 89:805–826.

Wolfram, C. (2005). *The Efficiency of Electricity Generation in the U.S. After Restructuring*. University of Chicago Press.