Intermittency and the Value of Renewable Energy*

Gautam Gowrisankaran† Stanley S. Reynolds‡ Mario Samano§

March 18, 2011

Abstract

This paper develops an empirical approach to estimate the equilibrium value of renewable electricity technologies, and applies it to evaluate solar energy mandates in southeastern Arizona. Solar generation and other renewables suffer from intermittency because weather varies and is only partially forecastable. Intermittency imposes costs as a planner must maintain backup capacity and allocate operating reserves in order to avoid system failure. We model an electricity system where a system operator optimizes the amount of generation capacity, operating reserves, and demand curtailment in the presence of variable and partially forecastable demand and renewable production. We use generator characteristics, solar output, demand and weather forecast data to estimate most parameters, and use existing estimates of demand elasticity. Equilibrium costs of a 20 percent mandate are \$136.1/MWh of solar generation, out of which unforecastable intermittency accounts for only \$2.7/MWh. If CO_2 reductions are valued at \$25/ton then this mandate would be welfare neutral if solar capacity costs dropped from the current \$5/W to \$1.82/W. Our methods can be applied to examine the value of other technologies, such as wind power and storage, and electricity market changes, such as real-time pricing.

*We thank Joseph Cullen, Lucas Davis, Ken Gillingham, Ben Handel, Paul Joskow, Catherine Wolfram and seminar participants at Arizona, Berkeley, NYU, and Yale for helpful comments and Valentina Kachanovskaya for research assistance.

†University of Arizona, HEC Montreal and NBER

‡University of Arizona

§University of Arizona

1 Introduction

Electricity generation from fossil fuels is the largest source of greenhouse gas (GHG) emissions worldwide. Currently over 70 percent of electricity produced in the U.S. is generated from fossil fuels [see EPA, 2009]. Many U.S. states and foreign countries have enacted renewable portfolio standards (RPSs) that specify minimum percentages of electricity generation from renewable energy sources. For instance, California's RPS specifies 33% power generation for renewables by 2020 while Arizona has set a 15% RPS by 2025. Carbon tax and cap-and-trade policies that various jurisdictions have considered or are considering will also likely significantly increase renewable energy production.

Many observers consider solar energy to be a crucial part of future renewable energy growth in the U.S., and solar energy has recently attracted large amounts of venture capital funds, plant investment, and federal and state government subsidies [see Glennon and Reeves, 2010]. While solar facilities produce electricity at marginal costs close to zero, they also produce intermittently, with production only during daylight hours and by far the highest production levels during clear, sunny periods. Moreover, weather conditions, and hence production levels, are not perfectly forecastable. Electricity is very costly to store and involves scheduling production to meet demand in real time. The inability to meet demand can result in a system failure where no one receives power and accompanying large welfare losses. If solar generation is adopted on a large scale then grid operators may need to engage in costly precautions, such as investing in backup fossil fuel generation capacity and scheduling additional generation reserves to avoid system failure. Thus, intermittency may significantly limit the value of solar energy and other renewables.

The general issues of intermittency for renewable power are understood both by policy-makers[1] and academics. A number of recent studies seek to quantify the potential importance of intermittency for different technologies. Some studies deal with backup capacity investment [see Campbell, 2010, Hansen, 2008, Hoff et al., 2008, Skea et al., 2008]; others with

[1] For instance, a recent Texas state report [see SECO, 2011] notes intermittency, costs 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."

the time-varying generation profile of renewable energy [see Borenstein, 2008, Denholm and Margolis, 2007, Joskow, 2010, Cullen, 2010b]; and finally some with intermittency and its impact on operating reserves [see GE Energy, 2008, Helman et al., 2011]. These are all important aspects of renewable energy. Yet, prior studies have not arrived at an overall economic assessment of the value of large-scale renewable generation. The extent to which different issues associated with intermittency balance out is unknown and depends crucially on how a particular renewable energy source affects optimal choices of generator scheduling, operating reserves and backup capacity.

This paper develops an empirical approach to estimate the equilibrium value of renewable technologies. We model the decisions of the system operator who must schedule generation and reserves and invest in new fossil fuel generators under different levels of renewable capacity. We use the approach to evaluate solar energy mandates in southeastern Arizona. This 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 Joskow and Tirole [2007], who model a system operator of an electricity market who seeks to maximize the discounted present value (DPV) of welfare when faced with fossil fuel plants that can suddenly fail. Our model builds on this paper by modeling renewable energy intermittency as similar to the unexpected failure of a traditional generator as well as by modeling variability and uncertainty in demand.

In our model, at time 0 the operator chooses how many new fossil fuel plants to build and how high to set the price for "curtailment contracts." These contracts 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. The operator must then decide how many plants to schedule for generation and reserves and also how much demand to curtail, if any. Operator decisions depend crucially on five factors: (1) the retail price, which it takes as given; (2) the variability of renewable power sources;

(3) the extent to which this variability correlates with demand; (4) the extent to which this variability is forecastable; and (5) the costs of building backup generation.

Our model has two central limitations. First, we do not model any dynamic linkages from period to period, as would occur with start-up costs for plants, for instance.[2] Second, our model is consistent with competition in the wholesale generation market but not with strategic multi-firm oligopoly bidding.

We apply our model to the portion of the electric grid operated by Tucson Electric Power (TEP), whose coverage area roughly consists of southeastern Arizona. We obtain 2008 data for TEP generator characteristics, demand by time period, generator failure rates, and solar photovoltaic (PV) output from a site in Tucson. We also obtain day-ahead weather forecast data from the National Oceanic and Atmospheric Administration (NOAA). We estimate the predictable and unpredictable components of demand and renewable outputs at every hour by regressing demand and renewable outputs on the previous day weather forecast of the conditions at that hour,[3] using a seemingly unrelated regression (SUR) model that captures the fact that the unforecastable components of these processes may be correlated. We recover most of the other parameters of the model using detailed data that include plant heat rates, plant outages, fuel prices, prices of spinning reserves, and capacity costs. We assume that demand follows a constant elasticity up to some maximum reservation price and calibrate both the elasticity and reservation price from the literature.

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.[4]

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

[3]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.

[4]TEP is subject to rate of return regulation by the ACC and this form of regulation has the potential to introduce inefficiency. For example, TEP may have limited incentives to hold down costs of generation and of providing operating reserves, since the regulatory commission is is likely to approve rates that would allow the utility to recover their costs. This regulatory effect is known as X-inefficiency. Wolfram [2005]

Using the estimated parameters of the model, we solve for the optimal policies under different counterfactual scenarios, involving different levels of solar capacity and assumptions about forecastability. The optimal solution involves balancing very low probability but very costly system failures against the certain cost of additional operating reserves. Crude simulation of very low probability events can be computationally very time consuming. We develop a simulation procedure that oversamples multiple plant failures by summing over different number of failures and then simulating which plants fail given that a certain number fail.

Under the assumption that solar PV capacity costs are \$5/W, the DPV of average cost of solar generation in Tucson is \$193/MWh (19.3 cents/KWh) over the life of the installation. This is \$135/MWh higher than the average cost of generation for a new combined cycle natural gas unit. This cost gap is narrowed to \$123/MWh, after taking into account savings in transmission and distribution costs due to the the distributed nature of some solar generation. Our model yields expected welfare associated with different potential RPS policies, taking into account optimizing behavior of the system operator. Not accounting for the benefit of CO_2 reduction, RPS policies of 10, 20 and 30% that are implemented solely with solar PV would impose equilibrium costs of \$129.3, \$136.1, and \$141.1 per MWh of solar generation respectively, with the upward slope due to the increasing substitution from low cost plants and the increasing need to construct backup fossil fuel generators. These per unit welfare costs are higher than the (adjusted) average cost gap of \$123/MWh, but not dramatically higher. The net costs associated with variability and intermittency range from 3 to 9 % of the average cost of solar generation for RPS policies in the 10 - 30 % range. Without unforecastable intermittency, the equilibrium costs of the 20% RPS would drop by \$2.7/MWh and without the positive correlation between solar output and demand forecast errors, would rise by only \$0.2/MWh. If CO_2 reduction is valued at \$25 per ton, the 10% RPS would be welfare neutral with a capacity cost of \$1.97/W and the 30% RPS would be welfare neutral with a capacity cost of \$1.69/W.

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 failure.

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. Finally, Section 6 concludes.

2 Background on Electricity

2.1 System Operations

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. The system is set up to supply the power that customers wish to use at each point in time. 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. To ensure matching of supply and demand, the manager of an electricity grid engages in "system operations." System operations involve control of generation plants, 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 generation plants 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.[5]

Several characteristics of the electricity industry lead to concerns about the adequacy of electricity resources (supply) to meet customers' demand; Bushnell [2005] explains these issues in more depth. First, as noted above, most electricity systems have very limited storage capacity. Second, in most electricity systems the retail prices paid by customers are fixed over long periods of time. Thus, the price mechanism is not used to adjust consumption in the short-to-medium term in response to shortage or surplus. Third, both demand and available

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

supply can vary considerably from hour to hour. Demand varies by time of day, season, and weather conditions. Rapid changes in weather conditions can lead to unexpected changes in demand. Supply can vary quickly and unpredictably due to equipment malfunction or breakdown and due to intermittent renewable generation.

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. A total system collapse is a drastic situation in which demand and generation are shut off over a large area in an uncontrolled fashion. An example was the 2003 blackout in the Northeast U.S. and Ontario in which 50 million customers lost power.[6] 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. In a system collapse, the sudden failure of one or more components of the generation and transmission system leads to a complete inability to supply power over the grid.

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, 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.[7] 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

[6] Minkel [2008].
[7] See, NERC [2011].

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.[8]

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 both competition in the wholesale market and competition 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 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.

[8] 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.

2.2 Electricity Provision in southeastern Arizona

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 1. 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.[9] 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 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.[10] 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.[11]

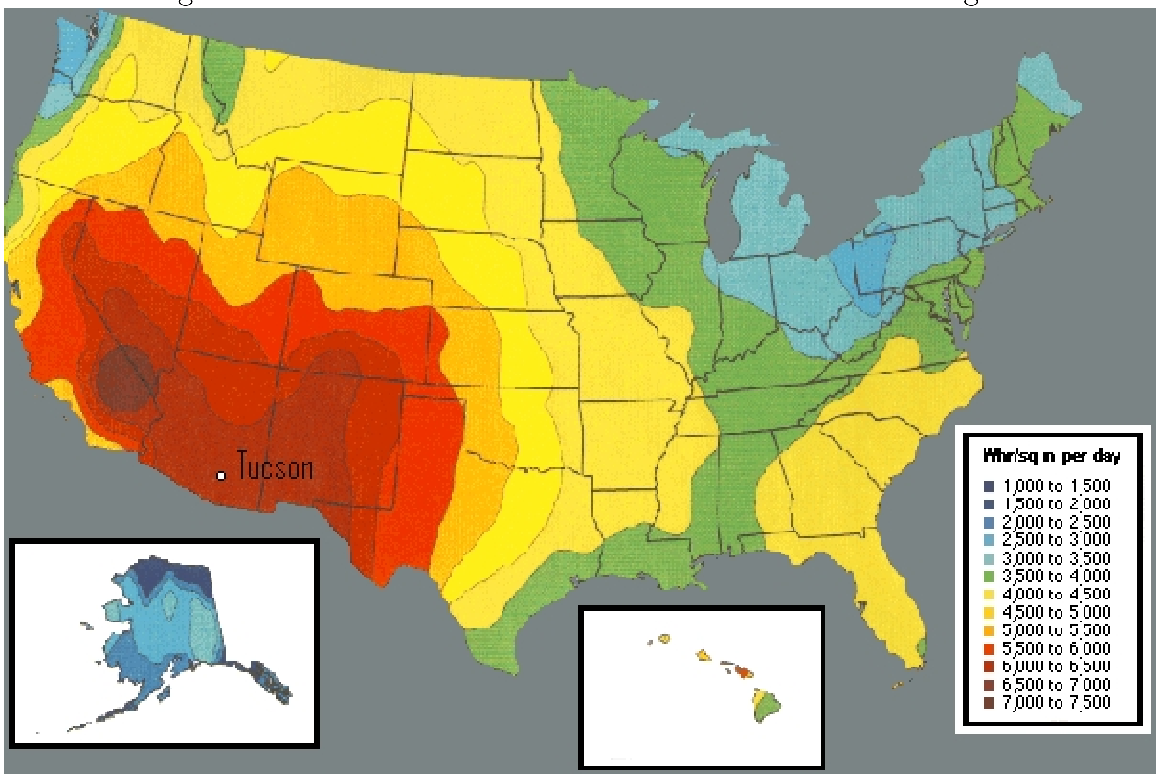
TEP is subject to a Renewable Portfolio Standard (RPS), mandated by the ACC, which calls for an increasing fraction of load to be generated from renewable sources until 15 percent of load is from renewables by 2025. For 2008 the RPS was 1.75 percent. TEP satisfies the RPS through a combination of its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

[9] 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].

[10] 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.

[11] TEP [2009].

Figure 1: Photovoltaic Solar Resource: Source www.NREL.gov



energy, distributed solar generation by its customers, and retirement of banked renewable energy credits.

2.3 Solar Energy and Intermittency

The nature of renewable energy intermittency depends on the characteristics of the renewable generation technology and the location of the equipment. For solar energy there are two main types of generation systems: solar photovoltaic (PV) and concentrating solar power (CSP), also known as solar thermal. 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. On the other hand, increases in temperature above 77 degrees Fahrenheit cause lower generation from crystal silicon solar PV panels, holding other factors constant.[12] The highest levels of solar PV generation in Tucson occur on sunny spring and fall days, rather than on the hottest days of summer. Most solar PV panels in the northern hemisphere 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 tilt adjusts according to the season (steep in winter, flatter in summer), and double-axis systems in which both the angle of panels and the direction in which panels face are controllable. The facing direction is adjusted over the course of each day to track the movement of the sun. Single and double-axis tracking PV systems yield more generation per unit of capacity but their capital costs and space requirements are greater than those of fixed tilt systems. Among fixed tilt PV systems, Borenstein [2008] notes that while south-facing panels will yield maximum total generation, west or southwest-facing panels may be more valuable because generation is shifted to later in the day.

Concentrating solar power (CSP) systems collect the sun's energy by using mirrors that focus sunlight on a heat-transfer fluid. The hot fluid then is used to boil water in a conventional steam-turbine generator to produce electricity or, in the case of a dish/engine CSP system, to drive pistons in an engine to create mechanical power which can run a generator. CSP systems are typically built at utility scale (50 MW capacity or more) whereas solar PV systems range from small rooftop residential installations to utility scale projects. Generation from CSP systems is less intermittent than with solar PV. The thermal energy storage feature of CSP yields less variability in generation during daylight hours and also permits generation to continue beyond daylight hours into the evening.[13]

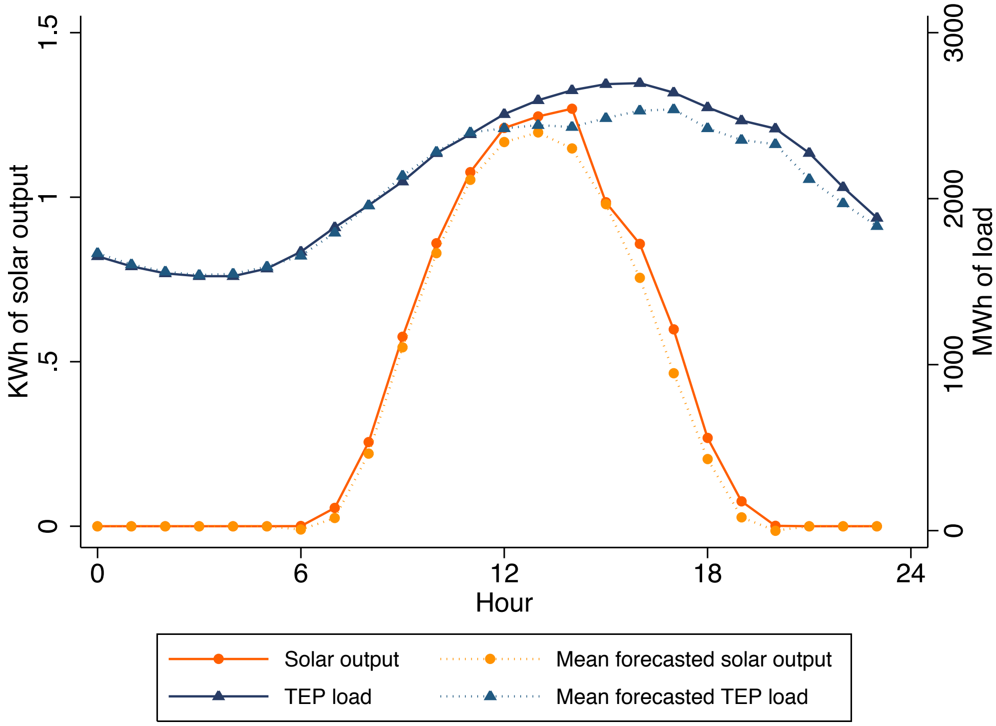
To illustrate the issues of intermittency, Figures 2 and 3 show southeastern Arizona demand and solar PV output in solid lines, for Jul. and Aug. 15, 2008 respectively.[14] Because southeastern Arizona power demand is driven by air conditioning, it peaks during hot and sunny periods; but sunny periods also have a lot of solar production. Thus, solar output correlates positively with demand during the daytime. This suggests a greater value for

[12] See Borenstein [2008]. The impact on PV performance depends on the particular PV technology. This effect is in contrast to CSP technology, for which higher temperatures improve performance.

[13] Glennon and Reeves [2010] provides an extensive discussion of CSP and solar PV systems, including issues involving water use, land use and siting, and environmental damage.

[14] The solar PV output is for a 1.536 KW test facility near the Tucson International Airport.

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



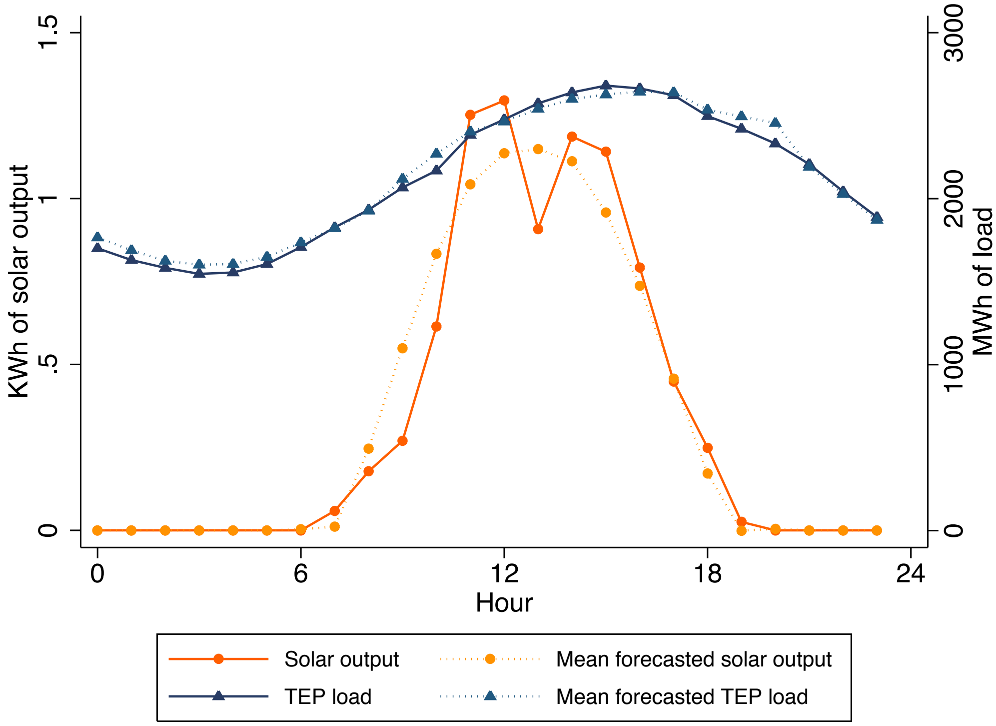
solar power than without a correlation since production occurs when demand, and hence marginal costs of generation, are high. Figures 2 and 3 also illustrate that the correlation of solar output and demand is not perfect; daily peak demand tends to occur later in the day than peak solar output. Moreover, although clear skies increase solar production, high temperatures decrease production.

With dotted lines, Figures 2 and 3 also show the mean forecasted demand and output using day ahead weather forecasts.[15] For Jul. 15, both actual output and load lie above their predicted values from 11AM onwards, suggesting that the day was sunnier and hotter than forecasted. However, solar PV output is particularly vulnerable to unforecastable intermittency relative to other technologies, because solar radiation drops precipitously when cloud cover appears.[16] On Aug. 15 at 1PM, actual output plummeted below forecasted output, illustrating this problem. Unforecastable intermittency is particularly costly because of the

[15] We provide details on our forecast methodology in Section 4.

[16] Glennon and Reeves [2010] consider solar PV to "present a major intermittency problem" (p. 97) relative to the older concentrating solar technology.

Figure 3: Predicted and actual southeastern Arizona load and solar output, Aug. 15, 2008



need to schedule backup reserves.[17] On one hand, the positive correlation in the unforecastable portions of solar output and load will increase the value of solar by decreasing the variability of load net of solar output. 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. We model a planner who makes optimal decisions regarding capacity investment for generation units, generator outputs, operating reserves and demand curtailment. The planner makes its decisions taking as given a level of solar capacity, as would occur if an RPS were binding; the current set of generators; and the retail price of electricity, which we assume to

[17] This point is generally understood by system operators. See, for instance, GE Energy [2008], who discuss this point for wind energy.

be a fixed number, $\overline{p}$, consistent with the relatively inflexible retail pricing observed in most U.S. electricity markets.[18] System reliability and total welfare are in turn functions of the planner's choice variables, retail price and solar capacity.

Our model builds on Section 4 of Joskow and Tirole [2007] and extends their model in several ways in order to arrive at a framework that can be taken to data. In particular, we model different types of power plants; price-responsive demand; uncertainty about demand forecasts; the possibility of demand curtailment contracts; reserves that are less costly than generation; and the possibility of renewable energy sources.

In our model, the future is discounted with discount factor β at the level of the year, and time runs from 0 to T, the lifetime of the generators. There are two stages of decision-making. In the first stage the planner chooses (a) the number of new fossil fuel generation units to construct, and (b) a price per MWh for compensating customers who have their demand curtailed, which we denote p_c. The second stage is composed of a sequence of short-run periods that span the lifetime of the generation units. In each second-stage period, the planner chooses which generators to schedule for production and operating reserves, and the level of demand curtailment. Because we endogenize the long-run choices of fossil fuel capacity investment, our model provides a framework with which the welfare impact of introducing significant amounts of renewable generation into the system can be numerically evaluated. We then evaluate policies that mandate different minimum amounts of solar PV capacity.

Each second-stage period corresponds to one particular hour, e.g. Jul. 15, 2008, 6PM, and is denoted t.[19] Periods are differentiated from each other, to account for the fact that electricity demand and solar output vary based on the weather, time of day, day of the week, and other available information, and that generators may be unavailable due to scheduled maintenance. In each period, the planner first makes its decisions, and then actual demand, solar generation, and generator breakdowns are realized. This then results in a level of utility

[18] It is possible to loosen this assumption to understand the relationship between real time pricing and the equilibrium value of renewables, among other questions.

[19] Hansen [2008] finds that solar cost calculations may be sensitive to the length of the observation period between a minute and an hour, suggesting the value of robustness analysis in our case.

as consumers obtain surplus from using electricity. In the unlikely event that total available generation (scheduled generation plus operating reserves) is less than the total demand, a system collapse occurs, with zero consumer surplus for a set number of periods, d^{fail}.

Part of the variation in demand and solar output is forecastable at the time that generation and reserves are scheduled while a portion will not be forecastable. For each second-stage period, we denote the state of the world as the set of variables that the planner uses when making its decisions. The state has two components: a vector w^t that represents information that might affect the distribution of solar output and/or demand for period t and that is observable at the time of production scheduling, and a vector m^t that indicates the scheduled maintenance status of each generation unit in t. Included in w^t are weather forecasts, the time of day, the day of the week, and the time since sunrise and sunset, all of which will predict load and/or solar output. Each state (w^t, m^t) thus implies a joint distribution of demand and solar generation for period t as well as a probability distribution for generator failures, or forced outages. Each generator failure is assumed to be an independent event and independent of (w^t, m^t).

3.2 Demand and Consumer Welfare

We specify demand for electricity to be a function of retail price p and forecast information w^t.[20] We choose a very parsimonious specification for demand in order to minimize the burden of identification of the demand parameters. 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^t, $\overline{D} \sim F^D(\cdot|w^t)$. Demand in period t 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 $F^D(\cdot|w^t)$ has a lower bound $\overline{D}^{min}(w^t)$.

[20] Although we consider a fixed-price regime, allowing for price-responsive customers is necessary to understand the welfare loss from system failure and hence to understand optimal decision-making.

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].

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 appendix for derivation.

Lemma 3.1 implies that we can calculate the reservation value v using estimates of VOLL and the price elasticity η.

Let $B_t(Q)$ be the gross consumer benefit function (area under the inverse demand curve) in period t as a function of quantity Q. If Q is equal to the quantity demanded at retail price $\overline{p}$ then $B_t(Q) = VOLL \times Q$. If there is a complete network collapse in period t then the opportunity cost of the collapse is $B_t(Q) \times d^{fail}$.

Our demand model also 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 per-unit price of p_c as compensation. In the first stage, all users with valuation below p_c will sign up for interruptible power contracts.[21] In each second stage period, knowing (w^t, m^t) (and hence knowing $F(\cdot|w^t)$), 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.[22] We assume that the set of known users has mass $\overline{D}^{min}(w^t)$.

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,

[21] 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.

[22] 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 $0 \leq z \leq \overline{D}^{min}(w^t)\,[\overline{p}^{-\eta} - p_c^{-\eta}]$, 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 appendix for derivation.

The welfare loss function $WLC(z, p_c)$ indicates the loss in consumer benefits relative to the amount of gross consumer benefit $B_t(Q)$ 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 failure. However, an increase in p_c also implies an increase in the average valuation of the curtailed user, which decreases welfare as it increases $WLC(z)$.

3.3 Generation from Fossil Fuel and Solar PV

We assume that there is a set of existing generation units indexed by $j \in \{1, ..., J\}$. Each unit has a maintenance status m_j^t at time t, with $m_j^t = 1$ implying that the unit is unavailable for production. Each available unit can be scheduled for production at full capacity or no production; let on_j^t denote a 0-1 indicator for scheduled production at time t. Note that $m_j^t = 1 \implies on_j^t = 0$.

Each unit uses a particular generation technology; coal, natural gas, etc. We use the following notation:

- k_j = production capacity in MW of unit j
- c_j = operating cost per MWh (constant) of unit j
- P_j^{maint} = probability of scheduled maintenance per period for unit j
- P_j^{fail} = probability of sudden failure per period for unit j
- c^s ratio of reserve costs to operating costs

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. 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.

Potential output from (non-solar) generation unit j at time t is given by

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

Our model of generation unit failures is based on the probabilities of losing discrete-sized units which allows us to take the model to the data on generation unit outages. Although generation unit start-up costs imply that operation decisions are dynamic [see Cullen, 2010a], we abstract away from this concern and treat all operations costs as static.

We allow for the planner to invest in new fossil fuel generation capacity. Specifically, we assume that there is a fixed capacity size k^{FF} for new fossil fuel generation units, with 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 plants, $n^{FF} \in \{0, 1, 2, \ldots\}$. Each of the new fossil fuel generation units have the same MC, maintenance probability, and failure probability. 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 plants, solar PV plants 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. Production from solar PV generation will then take on a state-contingent distribution $n^{solar}\overline{S}_t$, where $\overline{S}_t \sim F^S(\cdot|w^t)$. Let $F(\cdot|w^t)$ denote the joint distribution F^D, F^S of forecasted load and solar output. This formulation allows for the possibility of correlation between forecast errors for demand and for solar generation.

Finally, we abstract from transmission constraints and treat the system as a single zone. We assume a constant marginal cost of transmission and distribution (T&D) per MWh of delivered electric service; denote this cost as c_{TD}. One advantage of renewable energy is that the generation is often locally distributed, minimizing T&D costs. We assume that a fraction

$d^{TD,solar}$ of solar output is produced in a distributed environment and hence does not incur T&D costs.

Even though TEP imports and exports power from the Western Interconnection, we do not model this possibility.[23] Removing access to this market will tend to imply that the planner should construct more new generators than otherwise. However, we are not concerned with evaluating the absolute level of new construction. Rather, we evaluate the extent to which solar capacity *changes* the optimal number of new generators. We believe that this number will be relatively robust to not specifying the import and export market.

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}. In the first stage, the planner chooses n^{FF} and p_c. In each second-stage period t, the planner makes two decisions conditional on the state (w^t, m^t) and first-stage decisions: (1) generator scheduling decisions on^t and (2) amount of demand to be curtailed, z_t.

We model the choice of spinning reserves as a simplified version of how reserves are treated in unit commitment models.[24] Upon learning the state (w^t, m^t), the planner chooses on_j^t for each unit with $m_j^t = 0$. Then, the time t random variables are realized. Possibly, a complete system failure occurs, with no production. 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 highest marginal costs as reserves. Let $PC(\overline{D}, \overline{S}, x)$ denote the ex-post minimized costs of power generation and reserves for a given demand, solar output and actual generator output realization.

We illustrate the calculation of PC with a simple example. Consider a case with two scheduled generators each with capacity 1, with $c_2 > c_1$, realized demand is 1.6 and no

[23]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].

[24]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 can be thought of as a spinning reserve up.

generator failures or solar production. 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 $PC(1.6, 0, (1, 1)) = c_1 + 0.6 \times c_2 + 0.4 \times c_2 \times c^s$.

A system failure occurs when total generation is less than demand, taking into account both unit failures and demand curtailment. The probability of system failure, conditional on the state and actions is then

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

The planner's second-stage problem may now be defined as

$$\begin{aligned}
&W(w, m \mid n^{FF}, n^{solar}, p_c) = \max\nolimits_{z,on} \\
&\quad \big\{E\left[(1 - d^{Fail}SFP(z, on, w, n^{FF}, n^{solar}))\left(B(\overline{D}\overline{p}^{-\eta}) - WLC(z, p_c)\right)\right. \\
&\quad \left. -c_{TD}(\overline{D}(w)\overline{p}^{-\eta} - z - n^{solar}\overline{S}(w)d^{TD,solar}) - PC\left(\overline{D}(w), \overline{S}(w), x(on)\right) \mid w\right]\big\} \\
&\text{such that } m_j = 1 \implies on_j = 0.
\end{aligned} \tag{4}$$

From (4), the planner trades off the expected consumer welfare accounting for the possibility of system failure and demand curtailment (the first line) against the transmission, distribution and 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 the difference between production plus reserves and net demand:

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

Extra generation in the form of operating reserves provides a "cushion" in the event that one or more generation plants 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 (4) over all the periods in one year, and then discounting the expected annual welfare over the life of generation plants. Specifically, we suppose there is a distribution over states (w^t, m^t) over all the periods of a year. Using this distribution, define N to be the numer of hours in a year and define the annual expected welfare as

$$V(n^{FF}, n^{solar}, p_c) = N \times E[W_t(w^t, m^t) \mid n^{FF}, n^{solar}, p_c].$$

Then, expected total surplus is

$$TS(n^{FF}, n^{solar}, p_c) = \frac{1-\beta^T}{1-\beta} V(n^{FF}, n^{solar}, p_c) - FC^{FF} k^{FF} n^{FF} - FC^{solar} n^{solar}. \qquad (5)$$

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 (5). The amount of solar PV generating capacity, n^{solar}, is constrained via RPS regulations.

Although we have developed our model as the single-agent social optimum, it could be generated by a market-based model, similar to ERCOT, under the condition that each generator is run by a single firm. Specifically, we would model two auctions, one for the wholesale generation market and one for the operating reserves market. The system operator would submit hourly bid requests on the generation and operating reserves markets under uniform-price auctions where every firm is paid the lowest rejected bid. In the generation market, the planner would choose the total quantity of accepted bids to equal expected net demand less solar generation. In the operating reserves market, the planner would choose the quantity of accepted bids to equal the scheduled total quantity from the planner's problem in (4) minus the amount of generation from the generation market.

Under this system, single-generator firms would have the incentive to bid their marginal costs for production and reserves and hence all information would be revealed to the system operator. However, multi-unit oligopolistic firms will not bid their valuations in this type of auction as they will have the incentive to increase their bids above marginal costs on marginal units since this will benefit their infra-marginal units [see Ausubel and Cramton, 2002].

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 includes 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.

We create our data on generation units serving Tucson in 2008 by combining information 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 the Environmental Protection Agency (EPA) eGRID2007 report and from EIA Form 923. The EPA report 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. The EIA also has information about capacity investment cost for new generation units and average retail electricity price.

The EPA eGRID2007 report also 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. TEP units 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 AZ [see EPA, 2011]. SO_2 permit fees are from the EPA's annual advance auctions for years 2011 - 2017.

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 generation plants. 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.

Actual hourly solar generation data for 2008 is from a solar PV test site near the Tucson International Airport run jointly by TEP and the University of Arizona [see TEP, 2011]. This system has 24 solar PV modules with total rated capacity of 1.536 KW.[25] The modules

[25]This is a relatively small facility, somewhat smaller than the size of a typical residential installation.

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 at a fixed 30 degree tilt facing south. Summary statistics on solar output are given in Table 2. Actual mean hourly output per month is never more than 0.393 kWh even though the rated output of the system is 1.536 KW. No solar energy is generated between the hours of 9PM and 6AM. The maximum solar generation occurs in April, 2008. Unlike electricity demand, solar generation is relatively consistent throughout the year. If one assumes a 6% discount rate and a 25 year life for solar panels (as we do in our computations) then these data, coupled with our assumptions about the cost of solar panels, yield an average cost of \$193/MWh (19 cents/kWh) for solar PV generation. Note that the partially distributed nature of solar generation saves some T&D costs – in our case \$12/MWh as we discuss below – savings that should be taken into account when comparing solar PV generation cost to costs of other generation technologies. Also note that our average cost figure for solar PV is based on generation data for a particular set of panels during 2008. There are now solar PV panels available with higher efficiency, and hence lower average 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 Atmo-

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.

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

Month	Mean output (kWh)	Hour	Mean output (kWh)
Jan. 2008	0.282	6 AM	0.0005
Feb. 2008	0.325	7 AM	0.024
Mar. 2008	0.279	8 AM	0.190
Apr. 2008	0.393	9 AM	0.516
May 2008	0.373	10 AM	0.816
Jun. 2008	0.363	11 AM	1.026
Jul. 2008	0.334	12 PM	1.127
Aug. 2008	0.352	1 PM	1.141
Sep. 2008	0.389	2 PM	1.082
Oct. 2008	0.374	3 PM	0.931
Nov. 2008	0.320	4 PM	0.690
Dec. 2008	0.244	5 PM	0.380
		6 PM	0.114
		7 PM	0.013
Rated capacity: 1.536 kW		8 PM	0.0002
Average output: 0.344 kWh		9 PM – 5 AM	0

spheric 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. Our data 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].

We do not have data for generator outages or for costs of operating reserves used by TEP, and so we use information from ERCOT for these two measures. Specifically, to calculate the failure probability of fossil fuel plants, we use data from ERCOT on generation unit outages and resource plan data. Outages are relatively infrequent, so we compute average outage probabilities for two broad classes of generators: coal and natural gas. The outage data indicate which plants are unavailable both for scheduled maintenance and for unit failures

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

Forecast Variable	Average	Standard deviation
Cloud cover (%)	0.277	0.200
Temperature (°F)	70.4	16.9
Dew point (°F)	36.5	15.2
Relative humidity (%)	34.3	19.1
Wind speed (MPH)	8.53	4.06
Number of observations:	8,448	

(termed "forced outages" in ERCOT) [see ERCOT, 2011b]. The resource plan data indicate which plants are in operation at particular times. The ERCOT data on outages and resource plans do not overlap. Our approach is to pair outage data for August 2008 with resource plan data for August 2009. Our implicit assumption is that generator usage over the hours of a single month will be similar from one year to the next.

Finally, we use information from ERCOT on procurement costs for operating reserves and energy balancing market prices to provide a proxy for TEP operating reserve costs. ERCOT maintains ancillary services auctions for each of the four types of reserves described in Section 2.1. Market clearing prices for the four ERCOT operating reserves auctions for each hour of each day in 2008 for these ancillary service auctions were obtained from ERCOT [2011a]. We focus on up-regulation and responsive reserve services, as these are most like the spinning reserve concept we use in our analysis. Market clearing prices for the energy balancing market were collected for each hour of 2008. These prices should provide reasonable estimates of the hourly marginal cost of generation in ERCOT.[26]

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 [2004], but well within their range. Our value of $\overline{p}$ is based on EIA data for Arizona in 2008.

The reservation value can be recovered from (2) using numerical values for elasticity,

[26] See Hortacsu and Puller [2008] for a discussion of the extent to which exercise of market power by generators may drive a wedge between these wholesale market prices and marginal generation cost.

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
v	Demand reservation value per MWh	\$6,157	Computed so that VOLL is \$4,500/MWh
$F \equiv (F^D, F^S)$	Forecastable distribution of demand and solar output		Estimated

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 conservative estimate of VOLL=\$4,500/MWh which implies the listed reservation value. Note that a higher VOLL estimate or a lower demand elasticity would imply that the planner would want to prevent system failure more and hence maintain higher reserves. We can investigate the impact of higher VOLL numbers and demand elasticity on our results.

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. The unit of observation is the hour, for all daytime hours (defined as the hours after sunrise until the hour past sunset) in 2008. As solar output is zero outside these hours, we estimate a separate regression with just demand, for all the other hours in 2008. For 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 parametrizing the joint density of residuals, we directly simulate from this joint 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 and at the rated maximum of the system.

A large number of studies have constructed the marginal cost of operation for generation units. We follow the approach outlined in Cullen [2010a]. We compute the marginal cost of a fossil fuel generation unit as the product of the heat rate (MMBTU/MWh) and the cost of fuel (in $/MMBTU). The costs of emission permits for SO_2 also enter into MC of generation. The SO_2 emission rate for each unit is multiplied by the SO_2 average emission permit price for permits available for years 2011 - 2017.

Summary statistics for existing TEP generators are reported in Table 5. Except for a small 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 small size. Table 5 also lists characteristics of potential new generators, which we discuss below.

Table 6 lists the remaining supply parameters. For natural gas units we use a relatively small plant 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. The solar capacity cost includes the expected discounted present value of costs for 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 ERCOT data on the ratio of the average price in the auction markets for up-regulation and for responsive reserve services to the average price in the balancing market. The average price is $65.41/MWh in the balancing 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 0.32, while the average of the ratio of the up regulation to balancing market prices over all hours is 0.28. Our estimate of the reserve costs is the average of these two numbers.

We use the same constant per-unit cost for c^{TD} that is used in the real-time pricing simulations in Borenstein and Holland [2005]. The Arizona RPS states that 30% of solar energy must be generated in a distributed environment which motivates our choice of $d^{TD,solar}$. We

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)	20.57 (1.24)	3.92 (1.08)	2.35 (1.87)	2,163 (128)
Natural Gas – Combined Cycle	3	62 (20.7)	59.0 (0)	1.26 (0)	0.71 (0)	970 (0)
Natural Gas – Steam Turbine	3	59.3 (0)	89 (13.9)	3.90 (0)	6.44 (0)	1,955 (0)
Natural Gas – Gas Turbine	7	30.5 (18.5)	151.9 (109.5)	3.71 (1.48)	1.87 (3.12)	1,921 (47.2)
Potential New Natural Gas - Combined Cycle	By eqm.	60 (0)	38.5 (0)	1.26 (0)	0.71 (0)	970 (0)

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

estimate d^{fail} by examining EIA reports on "Major Disturbances and Unusual Occurrences" in the U.S.; see EIA [2010]. We identified outages due to equipment failure (not, for example, weather driven outages) that impacted more than 50,000 customers. For 2008-09 there were 10 such outages with an average duration of 2.4 hours. Finally, our assumption of a real discount rate of 6% and a lifespan of $T = 25$ for generation plants is equivalent to an annual real discount rate of 8% and $T = \infty$, in which case $\beta = 0.926$.

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. We proceed by maximizing the DPV of welfare over the first stage decisions of the number of new gas plants 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.

Table 6: Remaining supply parameters

Param.	Interpretation	Value	Source
d^{Fail}	Duration of system failure in hours	2.4	EIA
c_{TD}	T&D costs per MWh	$40	Borenstein and Holland [2005]
$d^{TD,solar}$	Fraction of solar generation that is distributed	0.3	Arizona RPS
c^{FF}	New gas generator MC per MWh	$38.44	EIA
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	0.3	Calculated from ERCOT data
P_j^{maint}	Scheduled maintenance prob.		Estimated
P_j^{fail}	Scheduled failure prob.		Estimated
k^{FF}	New gas generator capacity	60 MW	TEP average generator size
β	Discount factor	0.94	
T	Lifetime of generators in years	25	

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 plants in ascending order of MC when computing optimal generation for a second-stage period.[27] Although this point is intuitively reasonable, because of size differences across generators, it is possible that a planner would want to schedule a higher MC plant and not a lower MC one. Second, we assume that the planner curtails demand only if all available plants for which MC plus c_{TD} 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 coal units with planned outages; the natural gas units with planned outages; and the joint forecastable distribution of load and solar generation. We then choose the production and curtailment decisions, integrating over three remaining sources of uncertainty: forced outages of coal units; forced outages of natural gas units; and the realization of load and solar generation given the forecastable

[27] Coal plants have lower failure probabilities than gas plants, and thus this ordering will preserve the effective MC that accounts for intermittency.

distribution.

Given our above computational assumptions, we order generators by MC, and then loop over the number of generators to schedule. For each scheduling choice, the planner must choose the amount of demand to curtail, if the marginal generator has $MC > dWLC(z)/dz$. In this case, we compute a grid search with 200 points over the level of demand curtailment. If the marginal generator is sufficiently low, there is no further choice. Given the scheduling and curtailment choice, we integrate over the three dimensions of uncertainty, and then solve for the probability of failure 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 integrate over the joint distribution of load and solar generation conditional on a forecast with 20 discrete draws. Note that the planner's problem also involves simulation of generator failures. Outage probabilities for individual generation units are small, and probabilities of multiple outages – which might cause a system collapse – are very small, but the adverse consequences of a system collapse 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 n failures given N operating generators,[28] 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 Estimation Results

The estimated relationship from the SUR model of daytime load and solar output on weather forecasts is reported in Table 7. We estimate splines for each regressor. For cloud coverage,

[28] If the probability of generator failure is p, then this probability is $Bin(N, N-n)p^n(1-p)^{N-n}$.

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 on different temperature levels on solar output suggest that higher temperatures lower solar output though not significantly. 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 0.965 for load and 0.897 for output, suggesting that both levels are highly, though not perfectly, forecastable. The correlation in the residuals between load and solar output is 0.136 and statistically significant ($\chi^2(1) = 71.1, P < 0.01$). 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 gas generators report a higher rate of sudden failure (0.235%) than do coal generators (0.0052%).

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 adding higher RPS policies.

Without solar PV investment, the planner chooses 13 new natural gas generation units. Demand curtailment accounts for 0.3% of operating reserves although at peak times such as July at noon, the probability of some demand curtailment is over 10 percent. On average over

Table 7: Estimation of Daytime Load and Solar Output Forecasts

	Load (MWh)			Solar output (Wh)		
	Slope for			Slope for		
	1st decile	5th decile	10th decile	1st decile	5th decile	10th decile
Temperature	–18.091**	5.586*	48.832**	3.965	–3.390	–5.522
	(3.144)	(2.629)	(2.175)	(5.549)	(4.640)	(3.839)
Dew point	–4.863	2.148	–3.951	4.104	–13.22*	–1.234
	(2.989)	(3.527)	(3.265)	(5.276)	(6.27)	(5.763)
Relative humidity	29.483**	4.783	–4.039**	–10.440	–0.894	0.673
	(7.336)	(2.779)	(1.469)	(12.948)	(4.904)	(2.593)
Wind	–7.752	–10.129	–3.764**	45.56**	–5.753	5.752**
	(4.989)	(5.988)	(0.965)	(8.805)	(10.568)	(1.702)
	2–15%	38–60%	78–94%	2–15%	38–60%	78–94%
Cloud cover	48.124	–5.014	180.477	–776**	–1636**	–2559.4**
	(130.21)	(121.1)	(189.893)	(230)	(213.8)	
	Slope for hour			Slope for hour		
	1	4	6	1	4	6
Hours since sunrise, AM	–34.578*	3.499	21.55	65.87**	113.8**	123.94**
	(17.291)	(12.612)	(15.068)	(30.519)	(22.2)	(26.59)
Hours till sunset, PM	–30.110	44.793**	30.708**	95.796	138.1**	95.085**
	(43.876)	(10.378)	(12.021)	(77.439)	(18.316)	(21.218)
Temp × cloud	–0.378 (1.913)			9.195** (3.378)		
RH × cloud	3.313 (1.810)			8.582** (3.194)		
Wind × cloud	3.484* (1.541)			–11.712** (2.721)		
Dew × cloud	–3.811 (2.222)			0.895 (3.922)		
6AM dummy	1674.5** (232.1)			55.460 (103.7)		
...						
12PM dummy	1921.79** (234.0)			570.275** (134.1)		
...						
6PM dummy	2088.988** (240.0)			–167.96** (33.164)		
R–squared	0.965			0.897		
Correlation of residuals			0.136**			

Note: Model estimated with a SUR specification. Number of observations is 4,508. 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)		
	Slope for 1stdecile	5thdecile	10thdecile
Temperature	–13.194**	13.261**	59.729**
	(3.500)	(3.251)	(6.183)
Dew point	–4.898	–7.018	–4.547
	(4.079)	(3.981)	(3.396)
Relative humidity	23.051*	5.651	–0.671
	(10.033)	(3.327)	(1.307)
Wind	–7.345*	–2.670	–7.665**
	(3.340)	(3.707)	(1.978)
	2–15%	38–60%	78–94%
Cloud cover	214.582*	127.557	–47.148
	(103.611)	(97.937)	(124.773)
Temperature × cloud cover	–6.333** (1.906)		
Relative humidity × cloud	–1.067 (1.229)		
Wind × cloud cover	7.495** (1.541)		
Dew point × cloud cover	4.846* (2.092)		
9PM dummy	268.24** (4.668)		
...			
3AM dummy	–56.271** (4.428)		
R–squared	0.956		

Note: Model estimated with OLS. Number of observations is 3,723. 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 in %	Planned outage probability in %	Observations
Natural gas generator	0.236** (0.011)	0.124** (0.008)	184,118
Coal generator	0.0052* (0.0029)	0.014** (0.0047)	61,146

all hours, operating reserves are 18.9 percent of load. The 18.9 percent figure for operating reserves appears to be higher than average actual reserves for many systems. There are several factors that 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, our estimate for plant outages may overstate the expected failure probability, leading to relatively high operating reserves. Third, TEP has a relatively small number of generation units and its largest units comprise a significant fraction of load. 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. This output level would require 740 MW of solar PV capacity, with an investment cost of $3.7 billion. The solar PV panels would yield roughly 1.5 million MWh per year, which represents a capacity factor of 23%. Optimal investment in new fossil fuel capacity falls from 13 to 9 new generation units, reducing fossil fuel generation by 240 MW. This yields a capital cost offset of $236 million (about 6% of solar investment cost).

Although solar generation offsets an equal amount of fossil fuel generation, optimal operating reserves rise compared to the no-solar case. Thus, the sum of production and reserves does not fall on a one-to-one basis, but rather, the 1.497 million MWh of solar production reduce scheduled fossil fuel production plus reserves by 1.255 million MWh. This represents a ratio of 84%.

Interestingly, the probability of system failure falls slightly with the RPS but is very low in both cases. The drop appears to be due to the combination of greater total generation capacity (including solar capacity) and higher levels of operating reserves. The overall impact of a 10% RPS standard is to reduce the expected DPV of welfare over the life of the units by about $2.6 billion gross of the CO_2 emissions reduction. Put differently, the net welfare cost of the solar mandate is approximately 70 percent of the $3.7 billion investment cost for solar PV capacity. These results factor in the value of SO_2 and NO_x reductions, based on prices in EPA SO_2 permit auctions and the fact that NO_x harm is low enough that there are

Table 10: Outcomes with Different RPS Levels

RPS Policy	0%	10%	15%	20%	30%
Solar PV capacity (MW)	0	740	1,110	1,480	2,220
Solar production (1000 MWh/year)	0	1,497	2,245	2,993	4,490
Load (1000 MWh / year)	14,193	14,193	14,193	14,193	14,193
New 60MW natural gas generators (#)	13	9	8	8	8
Scheduled non-solar prod. + res. (1000 MWh/year)	16,866	15,611	15,064	14,570	13,606
Realized non-solar prod. + reserves (1000 MWh/year)	16,864	15,609	15,063	14,569	13,605
Reserves as % of production	18.9%	20.6%	22.1%	23.9%	27.6%
Average prob. of system failure	7.04e-5	6.80e-5	6.43e-5	5.39e-5	5.00e-5
Curtailment price p_c ($/MWh)	334	464	722	729	729
Total curtailment quan. (MWh/year)	14,260	14,368	18,144	15,820	13,360
Prob. of some curtailment Jul. 12PM	10.2%	0.006%	0.001%	0.0003%	–
Prob. of some curtailment Jul. 6PM	9.7%	25.9%	22.9%	19.7%	16.5%
Production costs (million $/year)	333.5	289.1	268.7	249.9	220.8
T&D costs (million $/year)	568.3	550.3	541.5	532.4	514.4
Reserve costs (million $/year)	29.7	29.4	30.5	31.3	32.1
Gas generator investment costs (mil. $)	768	531	472	472	472
Solar capacity investment costs (mil. $)	0	3,700	5,550	7,400	11,100
DPV of net surplus (million $)	849,559	846,945	845,518	844,054	840,998
DPV of future solar production (mil. MWh)	0	20.22	30.34	40.45	60.67
Loss in surplus per unit solar production ($/MWh)	–	129.3	133.2	136.1	141.1
NO_x emissions (1000 tons / year)	22.0	20.7	19.9	19.0	17.3
SO_2 emissions (1000 tons / year)	20.9	19.2	18.2	17.1	15.4
CO_2 emissions (million tons / year)	15.0	14.0	13.4	12.7	11.5

no NO_x permits in Tucson.

With the RPS, curtailment price also rises, as it is optimal to be able to curtail more demand in the event of low solar output. Moreover, the most common curtailment times shift from noon to 6PM, when solar output is low but load is still high.

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 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 plants, the welfare loss per MWh of solar generation rises monotonically from $129.3 to $141.1 as the RPS increases from 10% to 30%.

Comparisons between solar PV and conventional generation are often based on average cost over the life of the unit.[29] As noted above, the average cost of solar PV generation net of the T&D cost savings is $181/MWh. The average cost of generation for a new combined cycle generation unit is $58/MWh.[30] Thus, on the basis of a simple average cost comparison, solar PV is about 3 times more expensive than conventional generation, with an additional per unit cost of $123/MWh. 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. 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, we also consider system-wide factors associated with large-scale renewable energy, such as changes in operating reserves and changes in the amount of fossil fuel capacity. When these system-wide factors are taken into account along with time of day and seasonal differences in the value of solar generation, the equilibrium costs of solar actually exceed the $123 cost difference although the costs imposed by intermittency are generally less than those reported in the literature.[31]

[29]For average cost we employ the concept of the levelized cost over the lifetime of a generation unit. See `http://www.eia.doe.gov/oiaf/aeo/electricity_generation.html`.

[30]This is the levelized cost of energy, as reported by EIA [2011].

[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.

5.3 Equilibrium Costs to Solar from Unforecastable Variation

We now evaluate the equilibrium costs to solar from the fact that solar output is only partially forecastable and that the unforecastable part of solar correlates with demand. We present the results in Table 11. Columns 1 and 2 repeat the 0% and 20% RPS policies from Table 10 while columns 3 and 4 present the results of two hypotheticals.

Column 3 examines the optimal policy for the hypothetical and infeasible case where solar output at any time period was given by its forecastable mean value. We find that eliminating the unforecastable component of solar output results in one less gas generator under the optimal solution. Moreover, the costs of reserves drop by about 20%, to below their costs in the absence of any solar power. Because of these differences, the equilibrium cost of solar drops by \$2.7/MWh, from \$136.1 to \$133.4. In spite of the gains from not having any unforecastable variation, note that the drop is small compared to the overall additional equilibrium cost of solar generation.

Column 4 examines the optimal policy for another hypothetical case, where the unforecastable component of solar output at any time period had the same marginal distribution as estimated but where that distribution is not correlated with the unforecastable component of demand. Because we estimate a positive correlation between the two residuals, we would expect the value of solar capacity to be lower in the absence of a correlation. Indeed, we find this to be the case, but we also find the impact to be small: in the uncorrelated case, the cost of solar capacity is only \$0.5/MWh higher than in the feasible, correlated case.

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. It is beyond the scope of this study to analyze the environmental benefit or potential R&D outcomes. Thus, we proceed by choosing four

[2008] and Hansen [2008] evaluate similar costs for solar PV.

Table 11: Costs to Solar from Unforecastable Variation and Correlation with Demand

RPS policy	None	20%		
Experiment		Base	No var.	No corr.
Solar PV capacity (MW)	0	1,480	1,480	1,480
Solar production (1000 MWh/year)	0	2,993	2,993	2,992
Load (1000 MWh / year)	14,193	14,193	14,193	14,193
New 60MW natural gas generators (#)	13	8	7	8
Sched. non-solar prod. (1000 MWh/year)	16,866	14,570	14,002	14,654
Realized prod. + reserves (1000 MWh/year)	16,864	14,569	14,001	14,653
Reserves as % of production	18.9%	23.9%	19.9%	24.5%
Average prob. of system failure	7.04e-5	5.39e-5	5.78e-5	5.49e-5
Curtailment price p_c ($/MWh)	334	729	789	693
Total curtailment quantity (MWh/year)	14,260	15,820	14,753	16,718
Prob. of some curtailment Jul. 12PM	10.2%	0.0003%	1.5e-12%	1.1e-6%
Prob. of some curtailment Jul. 6PM	9.7%	19.7%	22.8%	23.0%
Production costs (million $/year)	333.5	249.9	250.2	250.0
T&D costs (million $/year)	568.3	532.4	532.4	532.5
Reserve costs (million $/year)	29.7	31.3	26.5	32.3
Gas generator investment costs (million $)	768	472	413	472
Solar capacity investment costs (million $)	0	7,400	7,400	7,400
DPV of net surplus (million $)	849,559	844,054	844,166	844,035
DPV of future solar production (million MWh)	0	40.45	40.44	40.44
Loss in surplus per unit solar ($/MWh)	–	136.1	133.4	136.6
NO_x emissions (1000 tons / year)	22.0	19.0	18.9	19.0
SO_2 emissions (1000 tons / year)	20.9	17.1	17.0	17.2
CO_2 emissions (million tons / year)	15.0	12.7	12.6	12.7

"No var." is a hypothetical solar facility without unforecastable variance.
"No corr." is a hypothetical solar facility where F^S is independent of F^D but otherwise as estimated.

levels for the cost of offset CO_2 (or, environmental damages from CO_2) that span the set of values suggested by most industry observers,[32] and by calculating 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 \$100/ton. At this emissions cost, solar capital costs would have to fall to \$3.48 for the 10% RPS to be welfare neutral, and \$3.31 for the 30% RPS to be welfare neutral. As one would expect, the target capital costs are decreasing in the value of offset CO_2 emissions. For instance, the 30% RPS welfare neutrality capital costs drop from \$3.31, to \$2.23, \$1.69 and \$1.14, as CO_2 emissions costs drop from \$100 to \$0.

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 generation plants, which will decrease its equilibrium value. On the other hand, the lower cost generation plants will tend to be coal instead of gas plants, and coal plants emit more than double the 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 \$25 CO_2 cost case, the welfare neutral capacity costs fall from \$1.97 to \$1.69 from the 10% to 30% RPS cases.

[32] The \$0 – 100/ton interval of CO_2 prices is within the range of estimates of marginal damage cost of CO_2 emissions surveyed by Tol [2005].

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	1.47	1.36	1.28	1.14
\$25	1.97	1.88	1.82	1.69
\$50	2.47	2.41	2.36	2.23
\$100	3.48	3.45	3.44	3.31

Note: solar capital costs are in millions of dollars per rated megawatt

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 develop 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 compute 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 and that we do not model firm market power. Moreover, other of our assumptions, notably our assumed T&D costs and 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 \$136.1/MWh, out of which unforecastable intermittency accounts for only \$2.7/MWh and that if CO_2 reductions are valued at \$25/ton, such an RPS would be welfare increasing if solar capacity costs dropped below \$1.82/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. Our approach calculates the costs if utilities 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 \$2.50 or \$2.00, 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 large 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; 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

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

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

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.

References

Ausubel, L. and Cramton, P. (2002). Demand reduction and inefficiency in multi-unit auctions. *Working Paper, University of Maryland.*

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.

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

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.

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.

Bushnell, J. (2005). Electricity resource adequacy: Matching policies and goals. *The Electricity Journal*, 18:11–21.

Campbell, A. (2010). Subsidized intermittent technologies and incentives for investment. 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 (2009). Executive summary – inventory of u.s. greenhouse gas emissions and sinks: 1990-2007. `http://www.epa.gov/climatechange/emissions/downloads09/GHG2007-ES-508.pdf`. Last accessed January 31, 2011.

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

ERCOT (2011b). Generation: archived zonal reports. `http://www.ercot.com/gridinfo/generation`. Last accessed January 21, 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.

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.

Hortacsu, A. and Puller, S. (2008). Understanding strategic bidding in multi-unit auctions: A case study of the texas electricity spot market. *RAND Journal of Economics*, 39:86–114.

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.

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