NBER WORKING PAPER SERIES

INTERMITTENCY AND THE VALUE OF RENEWABLE ENERGY

Gautam Gowrisankaran
Stanley S. Reynolds
Mario Samano

Working Paper 17086
http://www.nber.org/papers/w17086

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 2011

We thank Rob Clark, Allan Collard-Wexler, Joseph Cullen, Lucas Davis, Meredith Fowlie, Ken Gillingham, Ben Handel, Bill Hogan, Paul Joskow, David Keith, Derek Lemoine, Ariel Pakes, Philipp Schmidt-Dengler, Catherine Wolfram, John Wooders, Mo Xiao, Gregor Zoettl, the referees, Editor, and seminar and conference participants at numerous institutions for helpful comments; Alex Cronin and the University of Arizona Photovoltaics Research Lab for providing solar PV data; and Valentina Kachanovskaya and Anatolli Kokoza for research assistance. We acknowledge research support from Arizona Research Institute for Solar Energy. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Intermittency and the Value of Renewable Energy
Gautam Gowrisankaran, Stanley S. Reynolds, and Mario Samano
NBER Working Paper No. 17086
May 2011, Revised September 2014
JEL No. Q2,Q4

ABSTRACT

A key problem with solar energy is intermittency: solar generators only produce when the sun is shining. This adds to social costs and also requires electricity system operators to reoptimize key decisions with large-scale renewables. We develop a method to quantify the economic value of large-scale renewable energy. We estimate the model for southeastern Arizona. Not accounting for offset CO2, we find social costs of \$138.4/MWh for 20% solar generation, of which unforecastable intermittency accounts for \$6.1 and intermittency overall for \$45.9. With solar installation costs of \$1.48/W and CO2 social costs of \$36/ton, 20% solar would be welfare neutral.

Gautam Gowrisankaran
Department of Economics
University of Arizona
P.O. Box 210108
Tucson, AZ 85721-0108
and HEC Montreal
and also NBER
gowrisankaran@eller.arizona.edu

Stanley S. Reynolds
Professor of Economics
Department of Economics
P.O. Box 210108
Tucson, AZ 85721
reynolds@eller.arizona.edu

Mario Samano
Department of Economics
University of Arizona
P.O. Box 210108
Tucson, AZ 85721
msamano@email.arizona.edu

1 Introduction

> Like just about everyone who has looked at the numbers on renewable energy, solar power in particular, I was wowed by the progress. Something really good is in reach.
>
> Paul Krugman, *The New York Times*, Apr. 21, 2014

> For weeks now, the 1.1 million solar power systems in Germany have generated almost no electricity. ... As is so often the case in winter, all solar panels more or less stopped generating electricity at the same time. To avert power shortages, Germany currently has to import large amounts of electricity [including by] powering up an old oil-fired plant in the Austrian city of Graz.
>
> The German newsmagazine *Der Spiegel*, Jan. 16, 2012

Across the world, renewable energy capacity has increased dramatically due to falling prices, policies favoring renewable energy, and concern over greenhouse gas (GhG) emissions from fossil fuel generators. A key problem with generation from solar and other renewable sources is intermittency: solar generators only produce when the sun is shining.[1] Intermittency has the potential to hugely affect the economic value, or equivalently the social costs, of renewable energy. The above *Der Spiegel* quote suggests that electricity system operators use costly backup generation to manage the intermittency of large-scale solar. More generally, the implementation of large-scale solar may require significant changes in key decisions of investment, operations, and demand-side management to mitigate its social costs.

This paper develops a method to quantify the social costs and reductions in carbon emissions from large-scale renewable energy generation. We estimate social costs by solving for the decisions that maximize total surplus under different levels of renewable energy capacity. Social costs depend crucially on (1) the variability of the source including the extent to which the variability correlates with demand; (2) the extent to which output from the source is forecastable; and (3) the costs of building backup generation required to maintain system

[1] Notably, intermittency is a big issue for wind power as well and wind power is the largest source of renewable energy besides hydroelectric power [EIA, 2012].

reliability. We use our model to estimate the social costs of solar power for one grid area as well as to understand how the social costs vary under complementary policy environments such as greater real-time pricing. Our method can also be used to examine the social costs of different sources of renewable energy and in different contexts.

What does our economic analysis bring to the table? While there exists a systems engineering literature that evaluates the costs of renewable energy, our welfare-maximizing economic analysis is fundamentally different. We reoptimize operator decisions as a function of renewable energy penetration, which requires balancing the consumer welfare lost from system outages and demand-side management against the extra costs of maintaining and using backup capacity. Without reoptimization, output fluctuations from large-scale solar will lead to suboptimal decisions – e.g., overly high probabilities of system outage, or excess investment in backup capacity investment – which add to social costs. Moreover, an economic model such as the one we develop is necessary to understand differences in social costs across complementary policy environments.

The starting point for our analysis is Section 4 of Joskow and Tirole [2007], who model a system operator of an electricity market who seeks to maximize expected welfare when faced with fossil fuel generators that can suddenly fail. Our framework builds on Joskow and Tirole by modeling renewable energy intermittency as similar to the unexpected failure of a traditional generator, and by modeling a variety of factors (stochastic and variable demand, duration of system outages, generation and reserve costs for existing and new generators) that allow us to use available data in relatively simple ways. A contribution of our paper is to show how to take this theoretical model to data and use it to evaluate the social costs of renewable energy.

In our model, the system operator is faced with a fixed level of solar generation capacity as specified by an exogenous policy mandate. Initially, the operator chooses how many new fossil fuel generators to build and sets the price for contracts that curtail demand for flexible customers in periods of scarcity. Following this, in each (one hour) period the operator is faced with a distribution of demand, and in the presence of renewables, a joint distribution of demand and renewable output. Observing the distribution, the operator must then decide how many generators to schedule for generation and reserves and how much demand to curtail. Finally, demand and solar output are realized, resulting in a system outage if generation and reserves are insufficient to meet demand.

We apply our model to the Tucson, Arizona grid area. We obtained solar output data from 58 geographically dispersed sites in the Tucson area. We complement these data with a variety of cost and failure information on generators, day-ahead weather forecast data, and demand information. We forecast the demand and renewable output distributions that are observable to the operator with regressions of demand and renewable outputs on the weather forecasts made the previous day. We calibrate the demand elasticity from the literature.[2]

It is well understood that the average accounting cost of solar is higher than for fossil fuel generators. Not surprisingly, we find that the gross social costs of solar are high and increasing in penetration, going from \$126.7 per MWh with 10% penetration to \$138.4 per MWh with 20% penetration. The increase in cost from increased solar penetration is caused by solar generation substituting for cheaper fossil-fuel plants as penetration rises. Dividing the social costs by the CO_2 reductions, a ton of CO_2 would have to have an environmental cost of \$275 or more to justify immediate large-scale investment in solar, a figure that is much higher than the U.S. government figure of \$36 [see EPA and DOE, 2013].

Perhaps more surprising than the high cost is a breakdown of its sources. Focusing on the 20% case, if solar were perfectly forecastable, the social costs would drop very little, by \$6.1 per MWh. With perfect storage of solar, social costs would drop by \$45.9 per MWh, implying that intermittency overall is quantitatively much more important than unforecastable intermittency. In contrast, the bulk of the cost of solar comes from its relatively high fixed costs: if the fixed costs of solar dropped from their current level of \$4.41 per watt to \$2 per watt (as is reflected by the Krugman quote and as many industry observers believe will occur), then the social cost of solar would drop by \$99 per MWh.

[2]Alternatively, we could structurally estimate demand by matching operating reserves from data to predicted values assuming optimizing behavior. However, we believe that it may be problematic to assume that current TEP decisions reflect optimizing behavior within the context of our model. For instance, Wolfram [2005] finds evidence that regulated investor-owned utilities (such as TEP) operate generation units at higher cost than do non-regulated utilities. Furthermore, the TEP system operator may act in a more risk averse manner than in our model, due both to career concerns and to regulatory penalties for system outage.

We find that geographic dispersion of solar sites is important in lowering the social costs of solar generation: if all solar facilities were located within 10 kilometers of the center of Tucson, the social costs of 20% solar would rise by $18 per MWh relative to the base case of sites located within a 40 kilometer radius. In contrast, other changes to the environment – notably allowing imports and exports, having forecasts made less than one day ahead, and placing a substantial percent of customers on real-time pricing contracts – have very limited impact on the social costs of solar.

Our work builds on both a systems engineering and economics literature on valuing renewable energy. One strand of the systems engineering literature focuses on the variability of renewable output and allows for changes in utility operations to be based on rules of thumb, rather than reoptimizing policies in response to the renewable energy.[3] A second strand uses optimization models from power systems engineering.[4] The analyses in these papers specify the cost-minimizing combinations of generation units necessary to meet forecasted demand and operating reserve requirements.[5] However, this strand limits optimization to the generator scheduling decision, with changes to reserve operations and fossil fuel capacity still based on rules of thumb. We find that a system operator who followed rules of thumb from this literature would face social costs that are 41% higher than with optimizing policies. Importantly, while the accounting costs incurred by not endogenizing these factors will not change much, the social costs will, due to high system outage probabilities.

An economics literature also seek to quantify the importance of intermittency.[6] One strand of this literature focuses on how the time-varying generation profile of renewable energy affects its value.[7] A natural starting point of valuation for some of these papers is the equilibrium price of renewables, but since prices are informative only for valuations at the margin, a price-based approach cannot easily be used to evaluate policies that lead to large-

[3]See, for instance, Fabbri et al. [2005], Mills and Wiser [2010], Lueken et al. [2012].

[4]See, for instance, Tuohy et al. [2009], Madaeni and Sioshansi [2013], Mills et al. [2013], GE Energy [2008].

[5]An advantage of this strand over our paper is that it often models dynamic linkages (e.g., ramping constraints and startup costs) in operation decisions within a day. We provide some robustness checks on additional generator startup costs.

[6]See Baker et al. [2013] for a review of this literature.

[7]See Denholm and Margolis [2007], Borenstein [2008], Cullen [2010b], Joskow [2011].

scale investment in renewables. A second strand addresses how fossil fuel capacity should be adjusted in response to large-scale renewables, but often does not focus on the time-varying generation profile of renewable energy [see Lamont, 2008, Skea et al., 2008, Campbell, 2011].[8] Our model combines insights from both these strands. To our knowledge, our study is the first to formalize reserve operations in an economic model that is taken to data, and the first to use weather forecast data to separate intermittency into its forecastable and unforecastable components.

The remainder of the paper is organized as follows. Section 2 provides background on the electricity market and solar energy. Section 3 discusses the model; Section 4 the data, estimation and computation; and Section 5 the results. Section 6 concludes.

2 Electricity and solar energy

2.1 The electricity industry in the Tucson area

We model the Tucson grid area, which includes most of southeastern Arizona. Electricity service is provided by Tucson Electric Power (TEP), a vertically integrated, investor-owned regulated utility. TEP's service territory covers 1,155 square miles and includes a population of approximately one million.[9] In 2011, 41 percent of TEP retail electricity consumption was residential, 21 percent commercial, and 38 percent industrial and public. In Tucson, the demand for power, called load, peaks during the summer, due to high usage of air conditioning (for instance, average August temperature is 88 degrees Fahrenheit). In 2011, residential customers faced retail prices of 9.9 cents/kWh; commercial and industrial customers faced retail prices of 9.4 cents/kWh.[10]

As of the end of 2011, TEP owned or leased generation units with total capacity of 2,275 MW. This capacity is virtually all powered by fossil fuel. In Arizona, there are 2,090 MW of planned new fossil fuel generation capacity, all of which are from natural gas generators. 91% of the planned new fossil fuel capacity in Arizona is from combined cycle generators with the

[8]An exception is Lamont [2008], who does include time-varying generation.

[9]Detailed information about TEP customers and operations are found in the 2011 10-K annual report for UniSource Energy Corp., TEP's parent company; [see UniSource, 2012].

[10]UniSource [2012].

remaining from gas turbine generators. Nationally, 88% of planned new natural gas capacity is combined cycle. Combined cycle generators have lower marginal costs and greenhouse gas emissions but higher construction costs.[11]

We now discuss some details of this market that motivate our later modeling assumptions. 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 grid area. At different times, TEP both imports and exports power over the Western Interconnection.

TEP is subject to a state-mandated Renewable Portfolio Standard 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 2011 the standard was 3 percent. TEP satisfies the standard through a combination of its own solar generation, wholesale purchases of renewable energy, distributed solar generation by its customers, and retirement of banked renewable energy credits.[12] Total distributed (i.e., customer-sited) solar capacity in TEP's service territory was 30.4 MW as of the end of 2011; enough to supply 0.3% of annual electricity consumption.[13]

Since storage is very limited for TEP and most electricity systems, the supply of power must equal (almost exactly) load on a real time basis. Moreover, load can vary unpredictably over the course of a day (e.g., due to weather changes) and available supply can vary quickly and unpredictably due to equipment malfunction or breakdown and due to intermittent renewable generation. To ensure matching of supply and load, the electricity grid operator engages in *system operations*. System operations involve control of generators, decisions about rationing power to customers, and control of backup systems. The system operator ensures reliability in part by having generators available on a stand-by basis so that customers can continue to be served in the event that one or more generators fails and/or load exceeds forecast. In the absence of grid operations, each supplier imposes externalities, because it does not bear the full cost of a system outage [see Joskow and Tirole, 2007].

[11]Source: Authors' calculations from http://www.eia.gov/electricity/monthly/current_year/january2014.pdf, Table 6.5.

[12]TEP owns no wind generation. While there is some wind generation in Arizona, wind is not expected by be a major source of renewable generation in the state.

[13]TEP [2012], Table 2.1.

The generation capacity that is scheduled by the system operator over and above the amount required to serve forecasted load is called *operating reserves*. Broadly speaking, utilities carry two types of operating reserves: contingency reserves, used in the event of a generator failure; and balancing reserves, used to smooth out fluctuations in load and renewable output. The North American Electric Reliability Corporation (NERC) requires utilities to have access to contingency reserves sufficient to cover the failure of its largest unit; for TEP this is a 403 MW coal generator, which is 20.9% of mean load. Ela et al. [2011] note that balancing reserves are typically set between 1-1.5% of peak load. Together, these imply that TEP should maintain approximately a 23% ratio of operating reserves to load.[14]

2.2 Solar energy

Solar photovoltaic (PV) systems utilize panels of materials (typically silicon) that convert solar radiation into direct current (DC) electricity, coupled with inverters that convert DC to alternating current (AC) [see NREL, 2011]. Electricity generation from solar PV panels varies with solar insolation, a measure of energy from sunlight. Solar PV technology has been improving rapidly.[15]

Investment in solar PV capacity has grown considerably in recent years, while wind capacity investment has recently declined. Roca [2013] projects that solar PV capacity investment would exceed wind capacity investment worldwide in 2013. The Tucson area, along with the rest of the Southwest of the U.S., comprise the locations in the U.S. with the most solar insolation; see Figure 1.

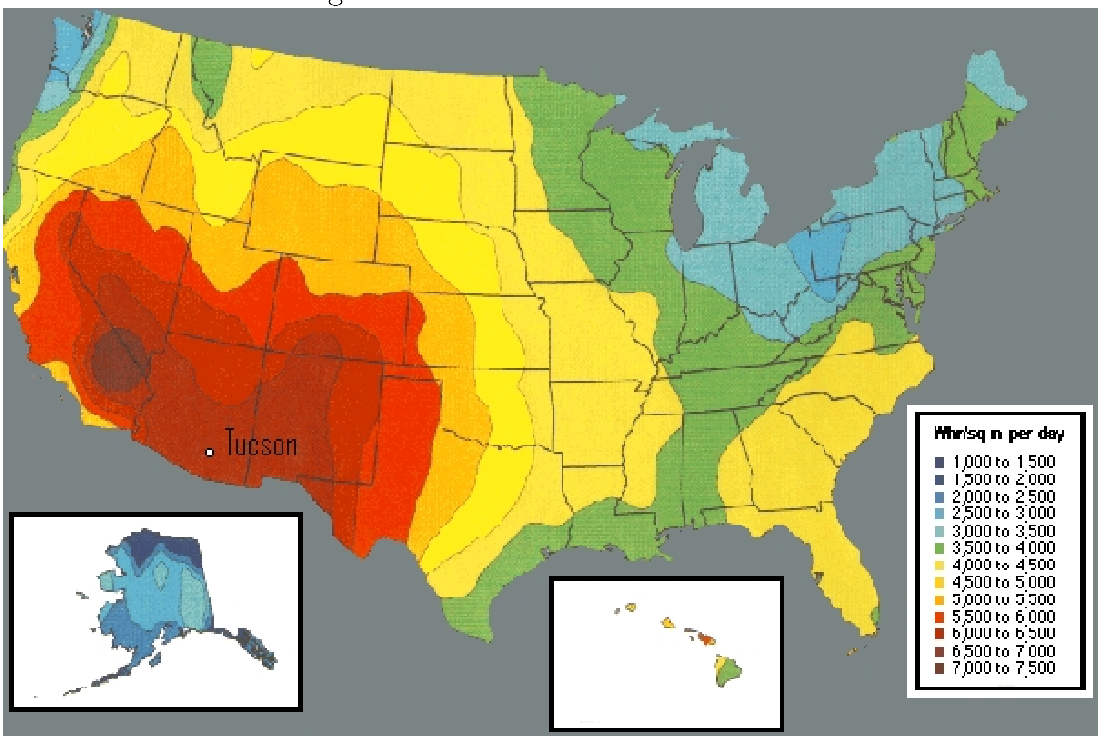
Because open land is plentiful near population centers in the Southwest U.S., it is common to locate large utility-scale solar PV near existing transmission links. In addition, Arizona's Renewable Portfolio Standard mandates that 30% of total renewable energy consist of distributed generation, e.g. solar PV on customers' rooftops.

New plants have to pay and report *interconnection costs*, which cover the extra transmis-

[14]Operating reserves may be lower because TEP is part of the Southwest Reserve Sharing Group. Participation in this group allows TEP to call upon other utilities during emergencies, such as plant outages and system disturbances, and reduce the amount of reserves TEP is required to carry [UniSource, 2012].

[15]Baker et al. [2013] describe the variety of solar PV technologies that are currently available, as well as potential technological improvements that are likely to change PV cost and efficiency in the coming decades.

Figure 1: Photovoltaic solar resource



Source: http://www.nrel.gov.

sion costs necessary to tie them into the grid. In Arizona, the mean interconnection costs for plants built between 2003 and 2011 were $45,454 per MW for natural gas combined cycle plants and $64,802 for solar PV plants.[16] Applying our data on solar production rates, and our assumed discount factor and plant lifespan to these numbers, the extra interconnection costs for solar are very small, about $0.78 per MWh generated.

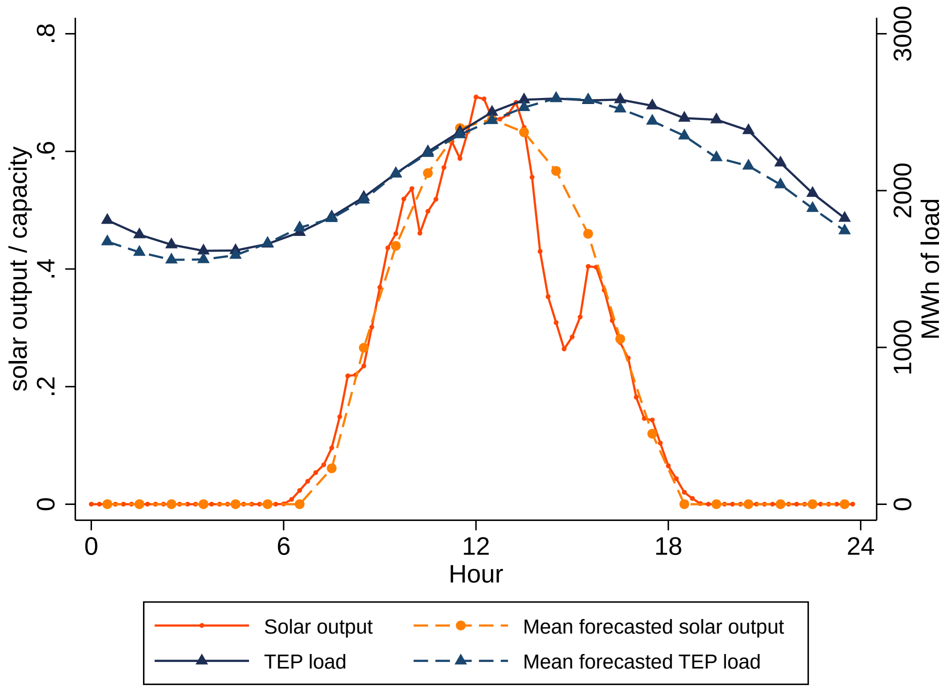
To illustrate the issues of intermittency, Figure 2 shows the Tucson area load and solar PV output in solid lines, for Aug. 15, 2011.[17] With dotted lines, Figure 2 also shows the hourly mean forecasted load and output using day ahead weather forecasts.[18] Load peaks during hot and sunny periods which also have a lot of solar production. Thus, solar output correlates positively with load, although peak load tends to occur later in the day than peak solar output. The positive correlation between solar output and load will increase the value of

[16]EIA [2011b]

[17]Load (in our data and in the figures) is at the hourly level and solar output is at the 15-minute level.

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

Figure 2: Predicted and actual Tucson area load and solar output, Aug. 15, 2011



Source: Authors' calculations from a variety of sources – see Section 4.

solar production because production occurs when load, and hence marginal costs of displaced fossil fuel generation, are high. However, to the extent that many evenings are similar to the figure, with high load but low solar output, the system operator will not be able to reduce fossil fuel capacity much. In addition, particularly during the late-summer monsoon season, solar output has large fluctuations that are not forecasted. The unforecastable intermittency in solar production will also lower its value.

3 Model

3.1 Overview

We now describe our base model of electricity generation, demand, and system operations. Our model takes as given the retail price of electricity and a state-mandated level of solar capacity.

Our model has two stages. In stage one, the system operator decides on capacity investment for new generation units, which are all combined cycle natural gas generators. The new generation units, solar capacity, and existing fossil fuel generators all last a fixed number of years. At this stage, the operator also chooses a price for interruptible power contracts and offers these contracts to customers.

In stage two, the operator is faced with a sequence of generator scheduling and demand-side management decisions for each hour of each year over the lifespan of the generators.[19] Here, she observes the weather forecast (made the previous day) and a list of which fossil fuel generators will be unavailable due to scheduled maintenance. Observing this information, the operator forecasts load and solar output and then chooses which fossil fuel generators to schedule for production and reserves and how much load to curtail from the set of customers who have signed up for interruptible power contracts.[20]

In each hour, following the scheduling and curtailment decisions, solar output and load are realized and some fossil fuel generators may fail. These realizations lead to one of two possible outcomes. First, load could be less than the sum of electricity output and operating reserves, adjusting for line losses. In this case, the operator will adjust the rate of generation for one or more generators to balance output with load. Second, load could be more than output plus reserves, adjusting for line losses. In this (hopefully rare) case, a system outage occurs and a random subset of customers involuntarily lose power.

Appendix B includes extensions to this model in the form of multiple solar observations per hour, real-time pricing contracts, and imports and exports outside the market. We now exposit demand, generation, and the operator's solution.

3.2 Demand and consumer welfare

We assume that retail price is a constant $\overline{p}$. We need to specify the demand curve in order to quantify the costs of system outage and the take-up of interruptible power contracts. We let demand be a random function of the weather forecast, denoted $\vec{w}$. We let $\vec{w}$ include cloud

[19]We abstract from linkages between hours, as would occur with ramping or startup costs.

[20]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.

cover, temperature, and similar variables as well as any factors that might affect load or solar output, notably the time of day, day of week, and time since sunrise and sunset.

We choose a parsimonious specification for demand in order to more easily calibrate the model. Specifically, we assume that demand has a constant price elasticity η for prices up to a reservation value, v, but that the scale $\overline{D}$ varies stochastically with $\vec{w}$, $\overline{D} \sim F^D(\cdot|\vec{w})$. Demand is then

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

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

The term *value of lost load* (VOLL) is used in the electricity industry to describe the mean value of electricity per unit for customers [see Cramton and Lien, 2000]. VOLL and the reservation value v have a closed-form and monotonic relation, detailed in Appendix A.

In our base model, the system operator manages demand through voluntary arrangements to curtail demand when necessary, in exchange for a payment at the time of curtailment. These contracts are known as pay-as-you-go interruptible power contracts [see Baldick et al., 2006]. Many utilities offer this type of contract[21] and there exist firms (e.g., EnerNoc.com) that intermediate these contracts between small- to medium-sized commercial customers and utilities.

The specifics of our curtailment contracts are as follows. In the first stage, the operator offers interruptible power contracts at price p_c. Customers who sign up agree to have their power curtailed as necessary and be paid a net per-unit price of $p_c - \overline{p}$ as compensation. In each stage 2 period, the operator chooses a quantity z of demand curtailment and randomly selects customers for curtailment from the set of customers who have signed up and who are known to use power in that period.[22]

The amount by which the system operator can curtail demand in any period is limited by the curtailment price p_c, as users will sign up for the interruptible contract if and only if

[21]See FERC (2011) "Assessment of demand response and advanced metering." Federal Energy Regulatory Commission Staff Report.

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

their valuation is below p_c. Note that there is a tradeoff from increasing p_c. An increase in p_c implies that the operator can curtail more demand, which increases expected welfare as it lowers the probability of a system outage. However, an increase in p_c also implies an increase in the average valuation of the curtailed user, which decreases welfare. Let $WLC(z, p_c)$ denote the *welfare loss from curtailment* in any stage 2 period; Appendix A derives a closed form expression for $WLC(z, p_c)$ and shows that it does not depend on $\overline{D}$.

3.3 Generation, transmission, and reserves

We index existing generators by $j = 1, \ldots, J$. At any given period, each unit has a maintenance status $m_j \in \{0, 1\}$, with $m_j = 1$ implying that the unit is unavailable for production. We assume that maintenance occurs with probability P_j^{maint}, and that its occurrence is *i.i.d.* across periods and generators. At each stage 2 period, the system operator schedules available units for production and reserves; let on_j denote a $0-1$ scheduling indicator. Note that $m_j = 1$ implies that $on_j = 0$.

Each unit uses a particular generation technology, e.g., coal-fired steam turbine, combined cycle natural gas; has marginal costs (MC) c_j; and capacity k_j. Generators are used for both production and operating reserves. For any generator, we assume that the marginal cost of reserves is a fraction c^r of the cost of producing electricity for whatever fraction of capacity of the generator is under reserve.[23]

We specify that each generator has a probability P_j^{fail} of failure in any period. Analogously to P_j^{maint}, we assume that failure occurrences are also *i.i.d.* Maximum output from

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

generation unit j is then given by

$$x_j(on_j) = \begin{cases} k_j, & \text{with prob. } (1 - P_j^{Fail})on_j \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

Let $\vec{x}(\vec{on})$ denote the vector of maximum outputs for all generators.

The system operator chooses the number of new fossil fuel generators, n^{FF}. Each of the new units is an identical combined cycle gas generator, with the same failure and maintenance probabilities as existing gas generators. We assume that each has fixed capacity size k^{FF}, investment cost of FC^{FF} per MW of capacity, and operating costs of c^{FF} per MWh. We label the new fossil fuel units $j = J+1, \ldots, J+n^{FF}$.

The system operator is faced with a fixed solar PV capacity n^{SL}, which is generated from a set number of installations spread over a metropolitan area. In contrast to fossil fuel units, we assume that solar units have zero MC and maintenance and failure probabilities, and are continuously scalable. Solar capacity costs are FC^{SL} per MW of capacity, which include installation and scheduled maintenance costs. Solar units are not dispatchable. Production from solar PV generation takes on a state-contingent distribution $n^{SL}\overline{S}$, where $\overline{S} \sim F^S(\cdot|\vec{w})$. Let $\vec{F}(\cdot|\vec{w})$ denote the joint distribution $(F^D(\cdot|\vec{w}), F^S(\cdot|\vec{w}))$ of forecasted load and solar output. This formulation allows for the possibility of correlation between forecast errors for demand and for solar generation.

We do not model the possibility that solar generators will be sited distant from existing transmission lines. If renewable facilities are sited far from cities and existing transmission lines – as is often the case with wind but not with solar – then a mandate for increased renewables might imply the need to construct such lines.[24]

In contrast, we assume that a fraction d^{SL} of solar capacity is installed in a distributed environment, which lowers transmission costs in two ways. First, we allow distributed solar to lower the fixed costs of transmission lines. Here, we assume that the discounted present value of equipment investment and maintenance costs is a function of the *maximum* expected

[24]For example, the 290 MW Agua Caliente Solar Project, which went into operation in 2014, is located in southwestern Arizona, right next to an existing 500 kV transmission line.

load net of distributed solar:

$$TFC(n^{SL}) = AFC^T \max_{\vec{w}}\{E[\overline{D}(\vec{w})\overline{p}^{-\eta} - d^{SL}n^{SL}\overline{S}(\vec{w})]\} \quad (3)$$

where AFC^T is a parameter. From (3), solar production will lower the fixed costs of transmission to the extent that it is positive in periods with the highest load.

Second, we model distributed solar as lowering line losses. Following Bohn et al. [1984] and Borenstein [2008], we assume that line losses from transmission in any period are equal to a constant α times the square of non-distributed generation. Letting Q be load net of demand curtailment and distributed solar generation, line losses satisfy $LL = \alpha(Q + LL)^2$. We let $LL(Q)$ be the function implicitly defined by this relationship, and obtain:[25]

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

Note that distributed solar production will reduce line losses, especially when it occurs during periods with high load.

3.4 System operator's problem

The operator seeks to maximize expected discounted total surplus, subject to $\overline{p}$ and n^{SL}. Her discount factor is β and the lifespan for her decisions is T, the life of the generators. Although we have developed our model using a single-agent social optimum approach, our results could be replicated by a competitive market-based model. For example, with multi-unit firms, Vickrey auctions for the generation and reserves markets could be used to induce efficient short-run outcomes.[26] Cullen and Reynolds [2013] show that the long-run competitive equilibrium of a generation market will generate the operator's solution; they do not consider reserve operations.

In the first stage, the system operator chooses n^{FF} and p_c. In each second-stage period, the operator makes two decisions conditional on weather forecast $\vec{w}$, maintenance statuses

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

[26] A Vickrey auction specifies that a firm that sells k units is paid the lowest k losing offers submitted by rival firms [see Krishna, 2010]. Vickrey auctions are not used by current wholesale markets such as ERCOT, likely due to their complexity. For instance, one would have to have Vickrey auctions for both generation and reserve operations.

$\vec{m}$, and first-stage decisions: (1) generator scheduling decisions $\vec{on}$ and (2) the amount of demand to be curtailed, z.

We model the choice of operating reserves as a simplified version of how reserves are treated in unit commitment models.[27] Upon learning $(\vec{w}, \vec{m})$, the operator chooses on_j for each unit with $m_j = 0$. Then, the state-specific random variables are realized. Possibly, a system outage occurs. Otherwise, the operator will adjust generation to exactly match load, by using the generators with the lowest marginal costs to satisfy demand, leaving the generators with the highest marginal costs as reserves. Let $PC(D, \vec{x})$ denote the ex-post minimized costs of power generation and reserves, where D denotes demand (net of curtailment) plus line loss minus solar production, and $\vec{x}$ denotes generator output realization vectors.

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

Let *outage* denote a $0-1$ indicator for a system outage. We have:

$$outage(\vec{on}, z, \vec{w}) = \mathbb{1}\{\sum_{j=1}^{J+n^{FF}} x_j(on_j) + n^{SL}\overline{S} < \overline{D}\overline{p}^{-\eta} - z + LL(\overline{D}\overline{p}^{-\eta} - z - d^{SL} n^{SL} \overline{S})\}. \quad (5)$$

In words, (5) indicates that an outage occurs when the maximum realized fossil fuel and solar output is less than demand and line losses net of curtailed demand. Note that the probability of a system outage is $E[outage(\vec{on}, z, \vec{w})]$. Let d^{outage} denote the expectation of the fraction of customers who lose power times the number of periods for which they lose power in the event of a system outage.

Using these definitions, we now define the system operator's problem for a second stage

[27] A unit commitment model specifies the cost of generation (including startup costs) as well as costs for several types of reserves: 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 [see Bouffard et al., 2005]. Our model is simplified in that we have a single type of operating reserve and no startup costs but, in contrast, our reserve policies are endogenously determined.

period as:

$$W(\vec{w},\vec{m}|n^{FF},p_c) = \max_{\vec{on},z} E\left[(1-d^{outage}outage(\vec{on},z,\vec{w}))\left(\overline{D}\overline{p}^{-\eta}VOLL - WLC(z,p_c)\right)\right.$$
$$\left. -PC\left(\overline{D}\overline{p}^{-\eta} - z - n^{SL}\overline{S} + LL(\cdot), \vec{x}(\vec{on})\right) \mid \vec{w},\vec{m}\right]$$
$$\text{such that } m_j = 1 \implies on_j = 0, \tag{6}$$

where the expectation is taken over the random variables $F(\cdot,\vec{w})$ and $\vec{x}(\vec{on})$ and the argument of $LL(\cdot)$ is the same as in (5). In (6), the first line indicates the gross consumer benefit and the second line indicates the production costs. Generators can only be operated if they are not undergoing scheduled maintenance, which constitutes the third line.

We now analyze the stage one decisions of p_c and n^{FF}. The system operator rolls up the second-stage payoffs by taking the expected value of W in (6) over all the hours in one year. Let H be the number of hours in a year. Given the stock of generators n^{FF}, the system operator chooses the price for curtailment contracts p_c to maximize expected welfare each year:[28]

$$V(n^{FF}) = \max_{p_c} E[H \cdot W(\vec{w},\vec{m} \mid n^{FF}, p_c)], \tag{7}$$

where the random variables are now $\vec{w}$ and $\vec{m}$. Finally, the generation buildout (or, investment) decision for the system operator may be expressed as:

$$V^* = \max_{n^{FF}}\{V(n^{FF}) - n^{SL}FC^{SL} - n^{FF}FC^{FF} - TFC(n^{SL})\}. \tag{8}$$

The function V^* provides the maximized expected annual total surplus. It will vary based on the solar capacity n^{SL} as chosen by the regulator.

4 Data, estimation, and computation

4.1 Data

In order to estimate and calibrate the parameters of our model, we use public data from a variety of sources, notably different federal agencies and grid operators. Our data pertain mostly to a one-year period from May, 2011 until April, 2012.

[28] We could equivalently express the problem as one of simultaneous choice of p_c and n^{FF} in stage one.

We obtain hourly load data for the Tucson service area from a Federal Energy Regulatory Commission Form 714 filing by TEP. Summary statistics on load data are provided in Figure 3. March was the month with the lowest load and August with the highest. Electricity load in the Tucson area has grown by roughly one percent per year.

We draw our data on generation units serving Tucson in 2011 from several sources. The Energy Information Administration (EIA) maintains a database on all existing generation units in the U.S. This database includes information about capacity, fuel source, and location. We take from this database the generators owned by TEP, as reported in UniSource [2012].[29] We also use EIA information on the average retail electricity price and capacity investment cost for new generation units.

We obtain information on heat rates from the Environmental Protection Agency (EPA)'s eGRID2010 database. This database provides heat rates at the plant level, where a plant may have multiple generation units. We assume that each generation unit at a plant has the same heat rate.

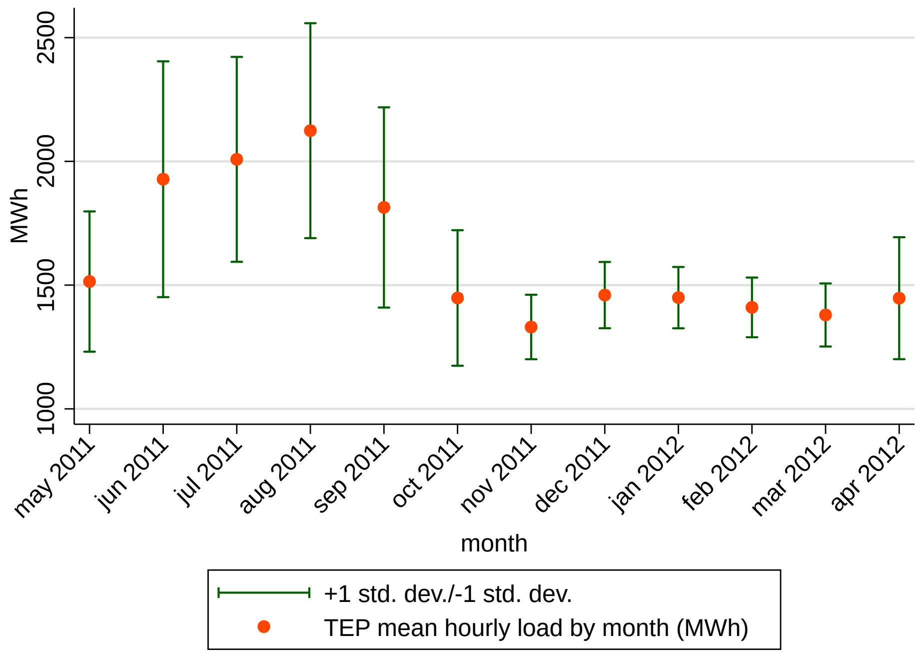
Since our analysis is forward-looking, we use information about projected future fuel costs. EIA Form 423 contains information about the terms of multi-year fuel contracts for each of the coal-fired generators. For natural gas we use NYMEX futures prices at Henry Hub in Louisiana [see CME, 2012]. We collect the last settlement price for each month for futures contracts in September 2012 for delivery from October 2012 through September 2017. Our natural gas price is the average of these prices.

The eGRID2010 database has mean annual emission rates for CO_2, SO_2, and NO_x at the plant level. We apply the same emission rates for each generation unit at a plant. SO_2 permit fees are from the EPA's annual advance auctions for years 2011 – 2017. TEP units are not subject to NO_x permit fees. EPA's NO_x Budget Trading program, a cap and trade program for NO_x, applies to 20 eastern states, but does not apply to Arizona [see EPA, 2011].

We use solar generation data from 58 geographically dispersed installations located within 40 kilometers of the center of Tucson, Arizona. The data were collected by the University

[29]TEP has partial ownership stake in four plants, Luna Energy, Navajo, Four Corners, and San Juan, each with multiple generators. For these plants, we assume that TEP owns one generator each with capacity equal to total plant capacity times TEP's ownership share.

Figure 3: Summary statistics for TEP hourly load (MWh), May 2011 - Apr. 2012

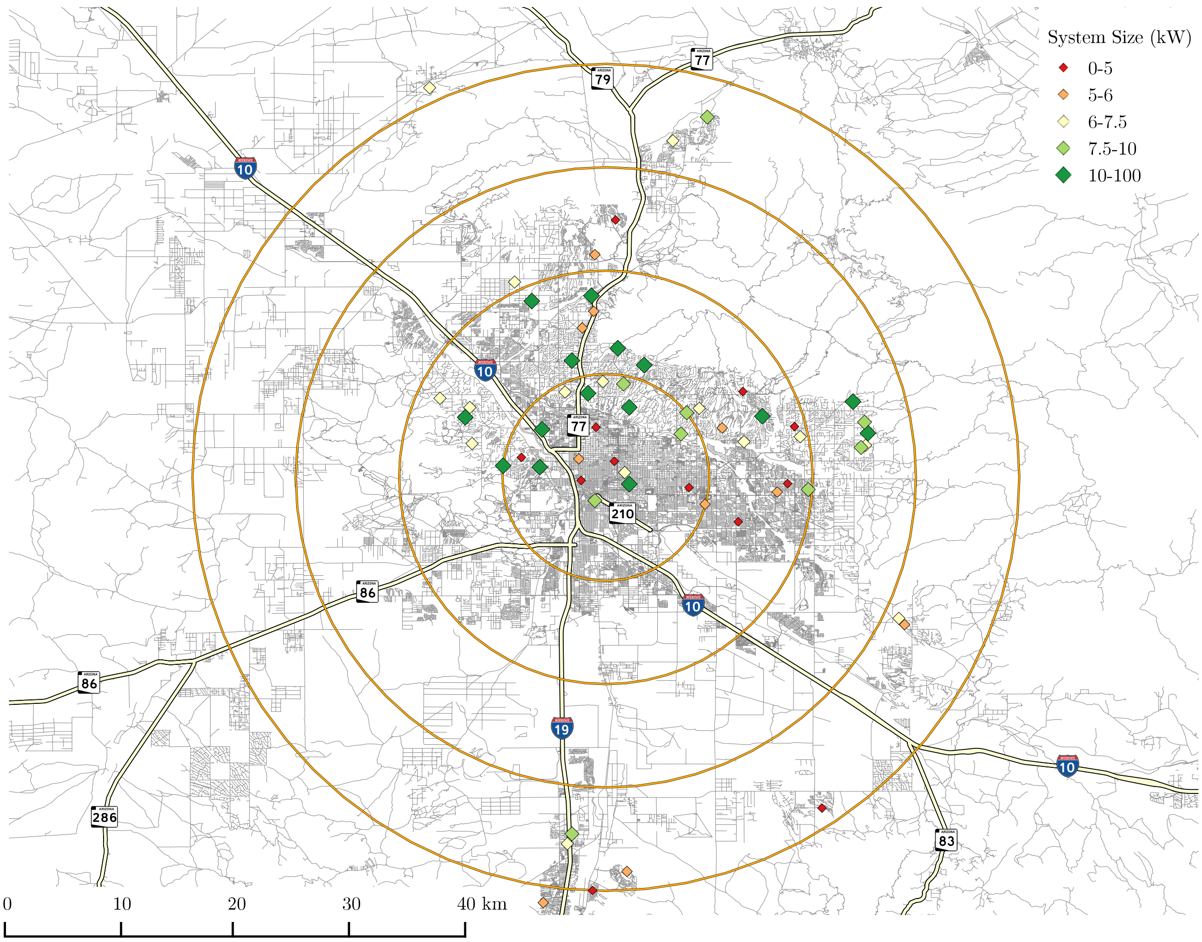


Source: FERC Form 714.

of Arizona Photovoltaics Research Lab. They record the power generated (in kilowatts of alternating current) by each site at 15-minute intervals from May 1, 2011 to April 30, 2012. All installations have fixed tilt PV panels, mainly south-facing. The average system size in the sample is 9.0 (all sizes are in kW of rated DC power), minimum size is 2.3, and maximum size is 84.2. Total capacity for the 58 sites is 517 kW. The age of installations (as of start of data collection period) ranges from one month to 3.5 years, with a mean of 9.5 months. See Figure 4 for a map of the installations. The circles represent distances of 10, 20, 30, and 40 kilometers from the center of Tucson. Figure 5 shows the mean and standard deviation of solar output by each 15-minute interval.

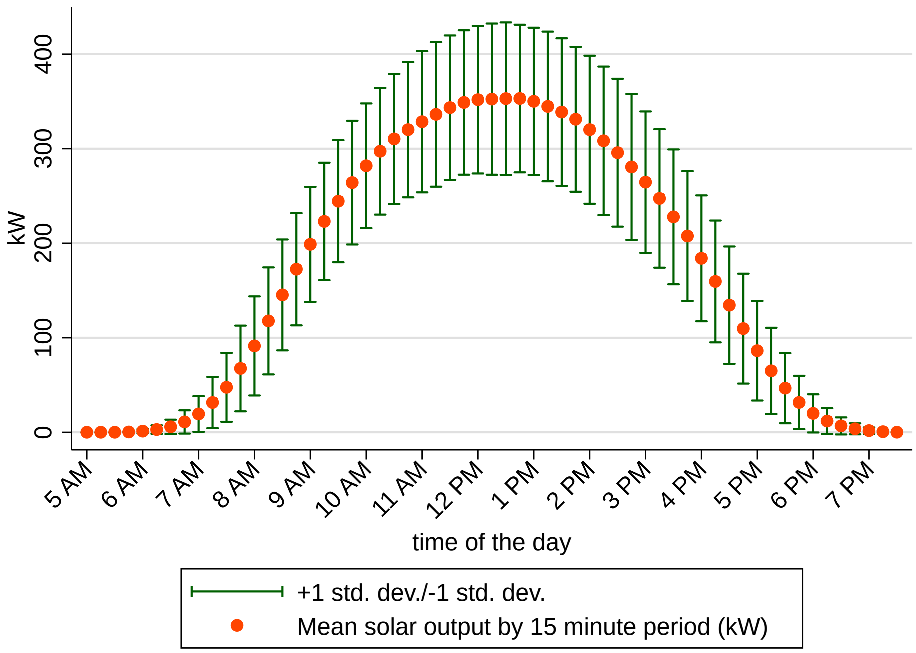
A novel aspect of this project is collection and use of weather forecast data which are used to determine the day-ahead forecasts of load and solar generation. We collect weather forecast data from the National Oceanic and Atmospheric Administration [see NOAA, 2014]. The forecasts are generally made at 2 AM, and forecast weather for the next day at 3 hour intervals. We interpolate to convert to hourly forecasts. Information includes cloud cover,

Figure 4: Map of Tucson with solar sites in our data



Source: University of Arizona Photovoltaics Research Lab.

Figure 5: Summary statistics for Tucson solar output, 2011 - 2012



Source: University of Arizona Photovoltaics Research Lab.

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 weather station, located at the Tucson International Airport. Table 1 provides information on the variables used in the weather forecast. We supplement the weather information with data on sunrise and sunset times at the daily level [see Sunrise, 2011].

We measure net import quantity as follows. We collect hourly data on generation for all TEP units from the EPA Continuous Emissions Monitoring System database [see EIA, 2011a].[30] Since some of these units are only partially owned by TEP, we multiply each unit's hourly output by its respective share of TEP ownership. We then calculated the difference

[30]The generation information for four small natural gas gas turbine generators with a combined capacity of 108 MW was missing.

Table 1: Summary statistics for information used in weather forecasts, 2011 – 2012

Forecast variable	Mean	Standard deviation
Cloud cover (%)	26.8	19.8
Temperature (°F)	69.7	17.6
Dew point (°F)	34.1	15.2
Relative humidity (%)	32.6	19.1
Wind speed (mph)	8.4	4.2
Number of observations	8,784	

Source: National Oceanic and Atmospheric Administration.

between hourly total generation output for TEP and the hourly load to get our measure of net imports. We also obtain a "system λ" for each hour from TEP's FERC Form 714 filing. As this value is meant to reflect the shadow price of additional generation, we use it to proxy for the import price of electricity. Finally, we use data on mean daily temperatures at each weather station in the states of Arizona, California, Nevada, Utah, and New Mexico.

We obtain startup costs from a National Renewable Energy Laboratory report [see Kumar et al., 2012] that calculates median capital and maintenance costs and associated fuel usage per cold, warm, and hot startup by generator type and size category. Data for generation unit outages come from the Generating Availability Data System of NERC. These data include outages due to maintenance and forced outages. We use these data for 2007–2011 for all U.S. generation units. We obtain data from the U.S. on system outage durations and number of affected customers from "Major Disturbances and Unusual Occurrences" reports [EIA, 2010]. Unlike TEP, the ERCOT grid in Texas operates a deregulated electricity market with prices. We use price data from ERCOT ancillary service auctions to define the costs of operating reserves ERCOT [2011]. Finally, we obtain total line losses from TEP [see UniSource, 2012].

4.2 Estimation and calibration of parameters

Table 2 lists the demand parameters of our base model. 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

Table 2: Demand parameters

Parameter	Interpretation	Value	Source
η	Demand elasticity	0.1	Espey and Espey [2004]
$\overline{p}$	Retail price per MWh	\$98.1	EIA
g	Demand growth factor	1.2	Based on historical rate of demand growth
VOLL	Value of lost load	\$8,000/MWh	Cramton and Lien [2000]
$F \equiv (F^D, F^S)$	Forecastable distribution of demand and solar output		Estimated

for Arizona in 2011. To be consistent with the loads projected during the middle part of the life of new generation units, we scale demand quantities by $g = 1.2$, based on historical rates of population and electricity consumption growth in Arizona. Importantly, the 20% growth yields non-zero investment in new fossil fuel generators for all counterfactuals.

Using mostly customer surveys, Cramton and Lien [2000] report estimates of VOLL that range from \$1,500/MWh to \$20,000/MWh. We choose a fairly conservative estimate of VOLL=\$8,000/MWh (with results from a higher value in Table 8).

To recover F^D and F^S, we estimate a seemingly unrelated regression (SUR) with load and solar output as the dependent variables. The unit of observation is a 15-minute interval for all daytime hours over our sample period. As solar output is zero at night, we estimate a separate nighttime load regression. We include linear splines and interaction terms of forecast information as regressors. To predict the joint density of solar output and demand, we simulate from the estimated residuals. We use our load estimates to recover the demand constant $\overline{D}$ by inverting the demand equation (1). We trim the solar output at 0 and at the rated maximum capacity.

We follow the approach outlined in Wolfram [1999] and Borenstein et al. [2002] to construct the marginal cost of operation for generation units. We compute the marginal cost of a fossil fuel generation unit as the sum of fuel cost per unit plus SO_2 emissions cost per unit. We report summary statistics for existing TEP generators in Table 3. Apart from two small solar PV facilities, all of TEP's generation units are fossil-fuel based. We treat the two solar units as though they were dispatchable units producing constantly at full capacity.

Table 3 also lists characteristics of potential new generators. We use a combined cycle natural gas generator with capacity size of $k^{FF} = 191$ MW. This number is the mean of the capacities of ten planned new units of this type in Arizona throughout 2023 as reported in EIA [2014]. Thus this number likely reflects the optimal generator size for a relatively small market such as the Tucson area. We assume that the other characteristics of new generators are the ones given in the EIA forecasts of emissions of new plants reported in EIA [2013]. For one of the extensions of the model, we also allow for potential new natural gas gas turbine generators. The mean size of these planned units is 91 MW for Arizona.

Table 3: Summary statistics for TEP generators, 2011 - 2012

Unit type	# Units	Mean size	Mean MC	Mean NO_x	Mean SO_2	Mean CO_2
Solar PV	2	6.5 (0.5)	0 (0)	0 (0)	0 (0)	0 (0)
Coal	6	263 (133)	23 (10)	3.0 (1.7)	1.6 (1.3)	1.0 (.06)
Natural gas combined cycle	1	185 (0)	35 (0)	.09 (0)	.01 (0)	.4 (0)
Natural gas steam turbine	3	89 (13)	54 (0)	1.5 (0)	.03 (0)	.5 (0)
Natural gas gas turbine	6	39 (20)	71 (13)	3.5 (2.0)	.05 (.01)	.8 (.2)
Potential new natural gas combined cycle	–	191	32.6	0.05	0.01	0.4
Potential new natural gas gas turbine	–	91	47.6	0.31	0.01	0.5

Note: standard deviations in parentheses. MC figures include emissions permits. Size is measured in MW, MC in dollars per MWh, NO_x and SO_2 emissions in pounds per MWh, and CO_2 emissions in tons per MWh. Source: authors' calculations based on a variety of data sources.

We calculate startup costs by multiplying the extra reported fuel used for startups by our fuel prices and adding the reported capital and maintenance costs to obtain our startup cost measure. Table C.1 (in the Appendix) provides details on our calculated startup costs for generators in our sample. These match closely with the startup costs for Spanish electricity generators estimated by Reguant [2014].[31]

[31] Another potential cost is ramping costs, which are the costs of changing output. Reguant [2014] finds that the costs of ramping a substantial amount over one period, 100 MW, are vastly lower than the costs of

Table 4: Remaining supply parameters

Param.	Interpretation	Value	Source
d^{outage}	system outage hours times percent of affected customers	0.98	EIA
d^{SL}	Fraction of solar generation that is distributed	0.3	Arizona Renewable Portfolio Standard
FC^{FF}	New combined cycle gas generator capital cost per MW	$1,095,458	EIA
	New gas turbine gas generator capital cost per MW	$921,927	EIA
FC^{solar}	Solar capital cost per MW DC	$4,410,000	Barbose et al. [2013], Baker et al. [2013]
c^r	Ratio of MC for operating reserves to production MC	0.41	Calculated from ERCOT data
α	Line loss constant	0.000035	Calculated fromTEP Form 10K
AFC^T	Average transmission fixed cost per MW	$1,259,000	Borenstein and Holland [2005], Baughman and Bottaro [1976], & TEP line loss cost
β	Discount factor	0.94	
T	Lifetime of generators in years	25	

A key input into our analysis is the installed cost per unit of solar PV capacity. Based on results for 2012 from Barbose et al. [2013] we find that the installed cost is $3.93 per watt of DC power.[32] We add the net present value of the cost of replacing inverter equipment, to arrive at a capacity cost of $4.41 per watt of DC power.[33]

Table 4 lists solar PV capacity cost and the remaining supply parameters. We compute the ratio of the hourly reserve marginal cost to the hourly generation marginal cost, c^r, using data from the deregulated ERCOT market on the 2008 prices in the up-regulation and responsive reserve markets, which pay firms in exchange for giving ERCOT the option to force them to operate with short notice.[34] If they operate, they receive the price on the

starting up a unit. Hence we do not account for ramping costs.

[32]Using data for 2012, Barbose et al. [2013] report a U.S. average installed cost of $3.30 per watt DC for utility-scale projects ($>$ 2 MW) and an average installed cost of $5.40 per watt DC for distributed PV projects in Arizona (based on their Table B-3). These installed costs are averaged using the Arizona standard that a fraction $d^{SL} = 0.3$ of solar PV capacity must be distributed generation.

[33]Following Baker et al. [2013] we assume inverter equipment is priced at $0.60/W, real inverter prices fall by two percent per year, and inverters are replaced twice over the life of the modules.

[34]The up regulation market gives firms 3 to 5 seconds to adjust production while the responsive reserve

balancing market. The average price is \$65.41/MWh in the balancing (production) market; \$27.05 in the responsive reserve market; and \$22.71 in the up-regulation market. The average of the ratio of the responsive reserve market to balancing market prices over all hours is 0.42, while the average of the ratio of the up regulation to balancing market prices over all hours is 0.40. Our estimate of the reserve costs is the average of these two numbers.

We estimate the transmission cost parameters as follows. For line losses, we find the value of α that matches TEP's reported line losses of 6.6% of 2011 load, using $LL(Q)$ defined in equation (4) to calculate line loss as a function of α in any period and then summing across every hour in 2011 to get total line losses as a function of α. To calculate AFC^T, we approximate average generating costs by \$70/MWh, so that the 6.6 percent line loss represents \$4.62/MWh. We difference the line loss from the total transmission and distribution (T&D) cost of \$40/MWh from Borenstein and Holland [2005] to obtain an average T&D fixed cost of \$35.38/MWh. Using information in Baughman and Bottaro [1976] we calculate that 55.3% of T&D fixed cost can be attributed to transmission. We obtain AFC^T in (3) by multiplying 55.3% of \$35.38 by the discounted sum of expected load and then dividing by the maximum expected load.

To estimate d^{outage} we use major disturbances reported by EIA [2010] whose causes were due to equipment failure (not, for example, due to storms) that impacted more than 50,000 customers. For 2008 there were 21 such disturbances for which we could find both the total number of customers and the number of affected customers, from which we calculate the percent of customers affected. For each of the 21, we multiplied the duration by the percent of customers affected. We estimate d^{outage} as the mean of this product.[35]

We calculate generator maintenance probabilities as the ratio of number of occurrences of outages due to maintenance to total available hours; for failure probabilities we use forced outages. We compute maintenance outage and forced outage probabilities separately for coal units and for natural gas units. We report the probabilities in Table 5. Note that coal generators report a higher P^{fail} than do gas generators.

market gives 10 minutes.

[35] Our estimate here is an approximation because some of these system outages may be due to failure of equipment other than generators.

Table 5: Mean hourly maintenance and failure probabilities

	Failure probability, P^{fail}	Maintenance probability, P^{maint}	Mean number of units over period
Natural gas generator	0.0492%	0.0382%	342
	(0.01%)	(0.008%)	
Coal generator	0.099%	0.047%	859
	(0.027%)	(0.010%)	

Note: time period covered is 2007-2011. Standard errors in parentheses. Source: Generating Availability Data System of NERC.

4.3 Computation of system operator's problem

We compute the system operator's problem conditional on a solar capacity level n^{SL} by finding the values of the long-run choice variables, n^{FF} and p_c from (8) and (7) respectively, that maximize social welfare. Since n^{FF} is a discrete variable, we perform a grid search for this variable, and a search using the simplex method for p_c.

The social welfare of any long-run choice depends on the discounted sum of the welfare levels at each hour in the T-year lifespan of the generators, net of the fixed costs. The observable data (to the economist) for an hour is the weather forecast, $\vec{w}$. We assume that the weather forecasts for the year-long sample period are repeated for each of the T years. The operator also observes maintenance status $\vec{m}$, and we integrate over this variable with simulation as it is not observed.

For a given hour, observing $(n^{FF}, p_c, \vec{w}, \vec{m})$ the operator chooses $\vec{on}$ and z. For each value of $(n^{FF}, p_c, \vec{m}, \vec{w}, \vec{on}, z)$, we simulate generator failures $\vec{x}(\vec{on})$ and demand and solar output $F(\cdot, \vec{w})$, and solve for the social welfare at that hour. We then perform a grid search in $\vec{on}$ and z to maximize social welfare at that hour. For each n^{FF} and p_c, we sum the maximum welfare levels across $(\vec{w}, \vec{m})$ to obtain the total welfare that feeds into our simplex search.

We make a couple of simplifications to our short-run optimization algorithm, in the interest of computational tractability. First, we assume that the operator schedules generators in ascending order of MC when computing optimal generation for a second-stage period.[36]

[36] This is called "sequential optimization" in the electricity literature.

Although this point is intuitively reasonable, because of differences in sizes and failure probabilities across generators, it is possible that a system operator would want to schedule a higher MC generator and not a lower MC one. Second, we assume that the operator curtails demand only if all available generators for which MC is below the marginal cost of curtailment, $dWLC(z)/dz$, are scheduled. Again, this point is intuitively reasonable but may not hold exactly because generators come in discrete chunks.

Finally, we provide some details on the simulation procedures. We use 20 draws to integrate over the joint distribution of load and solar generation conditional on a forecast. Note that the operator's problem also involves simulation of generator failures. Failure probabilities for individual generation units are small, and probabilities of multiple failures – which might cause a system outage – are very small, but the adverse consequences of a system outage are very large. Thus, our computation is challenging because integration using a direct simulation method would be very inefficient. Instead, for each type of generator, we integrate over the probability of a given number of failures given a total number of generators operating, and then simulate the identity of failed generators conditional on the number of failures. We integrate over maintenance probabilities using the same method.

5 Results

We first discuss the results of our forecast estimation. We then analyze the baseline predictions of our model; breakdown the sources of social costs for solar energy; and evaluate the social costs of solar energy under different complementary environments.

5.1 Forecast estimation results

The estimated relationship from the SUR model of daytime load and solar output on weather forecasts is reported in Appendix Table C.2. The nighttime impact of weather forecast on load is reported in Appendix Table C.3.

We find a U-shaped relation between forecasted temperature and load in both regressions, which occurs because electricity is used for both heating and cooling. For solar output, forecasted cloud cover variables are negative and are increasing in magnitude with more cloud cover, as expected. The R^2s are 0.959, 0.945, and 0.878 for daytime load, nighttime load, and solar output, respectively. The correlation in the residuals between load and solar output is 0.093 and statistically significant ($\chi^2(1) = 152, P = 0.00$).

Given that we assume that the system operator can use our regression output for generator scheduling and demand-side management, it is useful to understand how the precisions of our forecasts compare to those used by electricity firms and in the systems engineering literature. While we could not obtain TEP forecast information, HydroOne, the largest electricity provider in Ontario, reports R^2 values between 0.84 and 0.92 for their load predictions [Hydro One, 2012]. Many engineering studies report the mean absolute percentage error (MAPE), defined as the mean of the absolute value of load minus predicted load all divided by load, rather than R^2. Our daytime and nighttime MAPEs are both 0.03. A variety of studies on forecasting report MAPEs for load that range from 0.017 to 0.027 [see Fay and Ringwood, 2010, Weron and Misiorek, 2004].

For solar, MAPE will be infinite, because output can be zero. It is common to report the relative root mean square error (rRMSE) here, which is the square root of mean squared forecast error, divided by mean output. We find a daytime rRMSE of 0.23. This compares to daytime rRMSE values of 0.17 to 0.46 reported from a variety of solar studies using several different forecasting models [see Larson, 2013]. Thus, we believe that our forecast precision – for both load and solar output – is well within the range of forecasts that are currently used by utilities.

5.2 Social costs of large-scale solar energy

Table 6 reports our computational results for the social costs of large-scale solar generation, with more details in Appendix Table C.4. The first column reports the baseline with no solar PV investment. The remaining columns progressively add more solar capacity up to 20% of generation being produced by solar. Overall, we find that the social cost of large-scale solar energy – not accounting for CO_2 benefits – ranges from \$126.7 to \$138.4 per MWh.

Comparisons between solar PV and conventional generation are often based on levelized

Table 6: The social costs of large-scale solar (base specification)

Fraction of generation from solar	0%	10%	15%	20%
Foregone new gas generators (#)	0	2	2	3
Mean system outage probability	4.76e-5	5.82e-5	5.81e-5	8.4e-5
Reserves (as % of energy consumed)	30.5	32.1	33.6	35.2
Curtailment quantity (as % of total load)	0.11	0.19	0.14	0.24
Curtailment price p_c ($/MWh)	661	469	431	804
Production costs	437.2	380.0	355.2	332.2
Reserve costs	78.1	81.5	82.8	84.8
Gas generator investment costs	2,090	1,672	1,672	1,463
Solar capacity investment costs	0	4,148	6,221	8,295
Transmission fixed costs	331.4	319.4	317.4	316.2
Loss in $ surplus per MWh solar produced	–	126.7	133.7	138.4
Loss in $ surplus per ton CO_2 reduced	–	293.1	283.5	279.1

Note: cost variables are measured in millions of dollars per year. The second to last row (Loss in $ surplus per MWh solar produced) does not account for environmental benefits except for SO_2 permits.

cost, which is the average cost over the life of the unit. The realized solar output and our assumption about the cost of solar panels together yield a levelized cost of $181.2 per MWh for solar PV generation. The levelized cost for a new combined cycle generation unit is $66.3 per MWh [see EIA, 2014]. Thus, on the basis of a simple average cost comparison, solar PV imposes an additional per unit cost of $114.9 per MWh. Our analysis, which endogenizes choices in response to a solar mandate, finds that the social costs of 20% solar penetration are higher than the levelized cost difference by $23.5 per MWh.

We now evaluate the predictions of our model in a little more depth. Building solar capacity allows the system operator to forgo natural gas generators. Yet, because mean solar generation is well below capacity, the reduction is relatively small; e.g., for the 20% solar case, the operator forgoes three generators, each with a capacity of 191 MW (or 573 MW in total), in exchange for 1,881 MW of solar capacity.

A system outage occurs with probability 0.0048% each hour in the baseline, rising to 0.0084% with 20% solar. NERC mandates a "one day in ten years" standard for system outage. The interpretation of the NERC standard varies across utilities [see Cramton et al., 2013]. If the standard is interpreted as "one event in ten years" then this corresponds to a system outage probability of 0.0011% for each hour. If it is interpreted as "24 hours of system outage in ten years," then it corresponds to an expected failure probability of 0.027% in any hour. Our model uses a mean system outage duration of 4 hours implying an analogous figure of 0.019% for our model. Thus, depending on the interpretation, our model predicts a baseline optimal failure probability that is either 30% lower or 4 times higher than the NERC outage standard. Moreover, the optimal outage probability rises given large-scale solar, but not dramatically.

On average over all hours, operating reserves are 30.5% percent of load with 0% solar, going up to 35.2% for the 20% solar case. So our predicted reserves for 0% solar are similar, though somewhat higher, than the 23% ratio necessary to meet NERC reserve standards (noted in Section 2.1). The extra reserves may be due to the fact that we approximate the scheduling problem with sequential optimization, whereby the operator turns on generators solely in increasing order of marginal cost.

The optimal quantity of demand curtailment increases in the level of solar output, going from 0.11% of load with no solar to 0.24% with 20% solar generation. The optimal price for curtailment contracts is $661 per MWh in the baseline and is not monotonic in the level of solar. Here, the multiple tradeoffs between solar generation, backup fossil fuel generation, system outage probability, and curtailment do not lead to an easily quantifiable trend. Baldick et al. [2006] report that prices for these contracts range from $150 to $600 per MWh. The comparable price for our baseline is $661 minus the retail price of electricity, or $563 per MWh. Thus, our baseline number is within the range that exists in the real world.

We find that the probability of having at least one gas generator scheduled goes down from 97% to 90% (Appendix Table C.4) as we go from 0% and 10% solar to 20% solar and that the social costs of solar increase over the 10-20% range. Borenstein [2008], Lamont [2008], and Joskow [2011] point out that valuation of a renewable generation source should be based on the marginal cost of the generation it displaces. Since coal generators have lower marginal costs, the fact that solar substitutes more from coal at higher capacity levels contributes to the increase in the social costs of solar as solar capacity increases.

We also examine the social costs of solar including startup costs of fossil fuel generators in Appendix Table C.5. We find that adding large-scale solar does add to the total costs of startups. For instance, with 20% solar penetration, the total costs per startup accounts for an increase in \$5.3 per MWh of solar generated, raising the social cost of solar from \$138.4 per MWh to \$143.7 per MWh. Note that this likely overstates the extra startup costs generated by large-scale solar as, in our model, the system operator does not account for these costs in making her decisions.

Finally, Table 6 compares the loss in social surplus from each solar policy relative to the baseline to the offset CO_2 emissions. We find that the dollar cost per ton CO_2 reduced ranges from \$293.1 with 10% solar penetration to \$279.1 with 20% solar penetration. The U.S. government values the social cost of carbon at \$36 per ton. Thus, at present, one could not justify large-scale investment in solar on purely a social cost basis unless one believed the social cost of offset carbon to be almost an order of magnitude greater than the value specified by the U.S. government. Not reported in the table, solar capacity costs would have to drop to \$1.48 per W for 20% solar to be welfare neutral at the U.S. government figure.

Interestingly, the loss in social surplus per offset CO_2 emissions decreases in solar penetration. The reason for this is that the substitution from coal plants, though offsetting cheaper generation, is also offsetting much more carbon-intensive generation. In this case, the carbon-intensity of the coal generation is the dominant effect.[37]

5.3 Components of social costs for solar

Table 7 decomposes the sources of the social costs of 20% solar. The first row repeats the base 20% solar case from Table 6. The next four rows examine the impact of different components of intermittency on the social costs of solar. First, we find that eliminating the unforecastable

[37]The displacement of fossil fuel generation displaces other pollutants beside CO_2. Our analysis accounts for part of this benefit by including the EPA permit price for SO_2 emissions. However, the recent SO_2 permit price we use is close to zero and likely understates the marginal damage of SO_2 emissions [Schmalensee and Stavins, 2013]. Muller and Mendelsohn [2009] estimate marginal social damages from SO_2, NO_x and other pollutants. Using their values, we nonetheless find very small benefits to solar from offset emissions of these pollutants.

component of solar output lowers social cost by $6.1 per MWh relative to the base case. This drop is small compared to the overall additional social cost of solar generation. System engineering studies that account for backup capacity and operating reserve changes often find higher costs of unforecastable intermittency,[38] while studies that employ a narrower set of cost factors find somewhat lower costs.[39]

We next consider the total intermittency costs, which we define to be the difference in social costs between large-scale solar and a perfectly dispatchable energy source that produced the same total output.[40] We find that perfect dispatchability would lower the social cost of 20% solar by a substantial $45.9 per MWh. While many factors together result in the relative unimportance of unforecastable intermittency, it is worth noting that the R^2 of 0.878 for solar output is very similar to the ratio of 0.867 of unforecastable to total intermittency costs.

A hypothetical solar facility that always produced at its mean output level would be better than the base case by $4.6 per MWh. Distributed generation is also costly, raising the cost of solar by $19.5 per MWh. Here, the transmission cost savings from distributed generation are small relative to the extra fixed costs of installation.

All of these savings are small compared to the fixed cost of solar. Our fixed costs of $4.41 per W reflects average installed costs in 2012 for a mix of utility-scale and distributed (residential and commercial) sites, but the lowest cost utility-scale projects in the U.S. were approximately $3.00/W DC power (including costs of inverter replacements) [see Barbose et al., 2013]. If the cost of solar were to drop to $2 per W as many industry observers believe will occur, then the social costs of large-scale solar would drop by $99 per MWh.

The final two rows of Table 7 compare the social cost of solar under alternative policy

---

[38] For instance, Skea et al. [2008] find a cost of roughly 15% of the cost of wind generation for 20% wind power penetration in the U.K. This figure is based on the difference between the *capacity credit* for wind and average wind production. The capacity credit of a generation resource is meant to capture its contribution to system reliability and hence is similar to our notion of unforecastable intermittency.

[39] For 9-17% solar PV penetration in the Arizona Public Service (Phoenix, Arizona) territory, Mills et al. [2013] estimate that eliminating PV forecast errors would reduce costs by $2-4 per MWh of solar generation. Their analysis does not adjust backup generation, for instance.

[40] The systems engineering literature has no single accepted definition for total intermittency costs, or as it often called in this literature, integration costs. A common approach is to define integration costs as the difference between total system cost with renewables and total system cost for a reference case without renewables, although such a definition depends on how the reference case is defined [Milligan et al., 2011].

Table 7: Decomposition of social costs of 20% solar

Experiment	Loss in $ surplus per MWh solar
Base case – feasible solar	138.4
No unforecastable intermittency	132.3
Fully dispatchable	92.5
Equal generation profile	133.8
Eliminate distributed generation: $d^{SL} = 0$	118.9
Fixed costs FC^{SL} drop from \$4.41/W to \$2/W	39.4
Same policies as without solar	281.5
Rule-of-thumb policy	194.5

Note: "no unforecastable variance" produces at the forecastable mean. "Fully dispatchable" solar can be dispatched based on the demand forecast. "Equal generation profile" produces equally at every hour with the same total capacity as the baseline. "Rule of thumb" reduces fossil fuel capacity by the capacity factor of installed solar capacity and increases reserves by three standard deviations of the change in solar from one period to the next.

environments. We first examine a policy where the system operator uses the same n^{FF} and p_c as in the base case, and, in each stage 2 period, employs the same quantity of operating reserves and curtailed demand.[41] We find that this policy results in a much higher social cost of large-scale solar: $281.5 per MWh instead of $138.4 in the base case. Part of the reason is the extra generator cost. Not reported in Table 7, much of the loss comes from a much higher system outage probability than in the base case, of 0.3% in each period. We also examine the social costs given rules-of-thumb that have been suggested by the systems engineering literature in valuing solar energy. First, we reduce fossil fuel capacity by the solar capacity times the "capacity factor" (mean production level per unit of capacity) of solar in our data.[42] Second, following Mills and Wiser [2010], we increase both operating reserves and fossil fuel capacity by three standard deviations of the change in solar from

[41]Since reserve operations come in discrete quantities, we cannot implement exactly the same reserve operations. We use the smallest level of reserve operations that is weakly greater than in the base case.

[42]The systems engineering literature has generally suggested reducing fossil fuel capacity by a "capacity credit" [Mills and Wiser, 2012]. A capacity credit is similar to a capacity factor but takes into account other attributes of solar production in heuristic ways.

one 15-minute period to the next. The social cost of this rule of thumb policy is lower than without any adjustment of decision rules, but still high, at $194.5 per MWh with a still-high system outage probability of 0.09%.[43]

Thus, our finding is that following rules-of-thumb policies from the literature implies huge social costs of intermittency. Utilities often express concern over the high cost of intermittency with large-scale solar. We believe that this may be because they do not know how policies should change in the case of large-scale solar. The envelope theorem implies that it is not necessary to reoptimize policies with only a small amount of solar on the grid. But, the envelope theorem no longer holds with 20% of electricity generation from solar. Thus, a very useful complement to large-scale renewable energy is policy reoptimization with the goal of minimizing social costs.

5.4 Robustness to environment

Table 8: Social costs of 20% solar given different environments

Environment	Loss in \$ surplus per MWh solar	Δ surplus with 20% solar from baseline to new environment, in \$ per MWh solar
Base environment	138.4	—
No interruptible power contracts	137.8	-7.5
Imports and exports allowed	139.2	$+7.5$
Investment in additional generator type	138.7	$+0.3$
2PM (instead of 2AM) forecasts	139.3	-2.0
Forecasts with 24-hour lagged demand	138.5	$+1.0$
VOLL increased to $12,000	138.9	$+19,999$

Note: "investment in additional generator type" allows the system operator to choose NGCC and NGGT plants to be built.

We now examine how the social cost of solar changes across different environments, with results in Table 8 and Figure 6. The first row of Table 8 repeats the base environment 20%

[43]In contrast, using similar rules of thumb, Mills et al. [2013] finds that the *accounting costs* of intermittency for large-scale solar are relatively low, e.g., $3.77 per MWh for 17%.

solar. The second row presents the case without curtailment contracts. In this case, the value of 20% solar goes down by $7.5 per MWh solar produced. But, the value of no solar goes down by almost as much. The net effect is that the social cost of solar is virtually unchanged from the base environment.

A similar pattern is repeated across all rows in Table 8. The third row examines the value of adding imports and exports, which we obtain by first estimating a net import function and then by solving the operator's problem with the estimated net import function.[44] With imports and exports allowed, the welfare of the market increases by $7.5 per MWh solar produced, but the social cost of solar is very similar to the base environment, although net exports increase substantially in the presence of 20% solar production. The assumption of no imports or exports has been used in other studies of electricity in the Western U.S. [for instance Borenstein and Holland, 2005, Mills et al., 2013]; our results show that this assumption does not have a large impact on solar valuation. An important caveat is that, by holding constant the current import elasticity into the future, we are essentially assuming that TEP expands solar while other utilities do not.

We next examine the impact of allowing the system operator to choose two types of new gas generators: combined cycle and gas turbine. The values from this change in environment, both with and without solar, are very modest. However, not reported in the table, the system operator does choose some gas turbine generators. For instance, with 20% solar, the system operator goes from building 7 new combined cycle generators under the base environment to building 6 combined cycle and 3 (smaller) gas turbines when allowed.

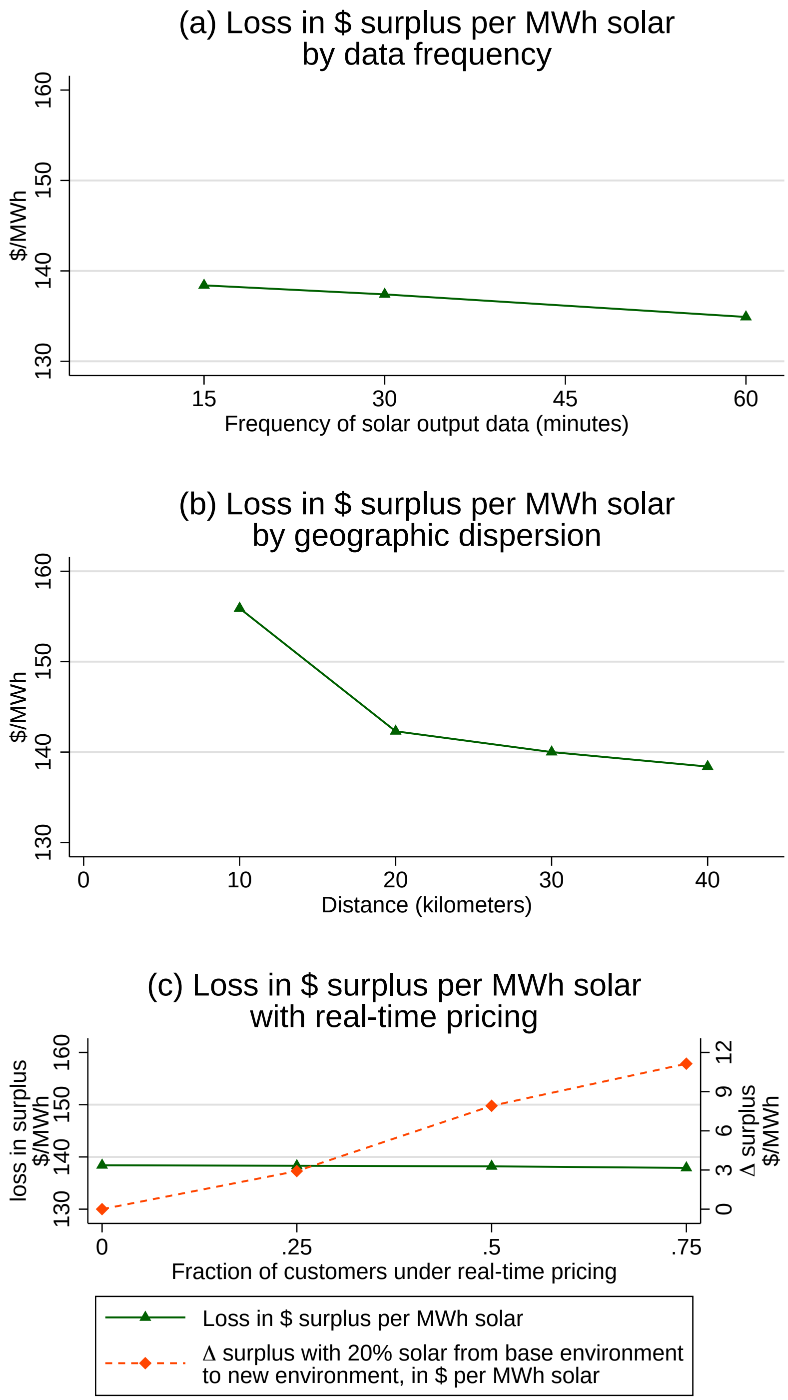
We next consider the value of better forecasts, and again find little change. Finally, Table 8 examines the impact of increasing the value of lost load from $8,000 to $12,000. The higher VOLL causes a huge increase in the aggregate consumer surplus – but only changes the social cost of solar by a tiny amount, of 50 cents per MWh.

Figure 6 shows the impact of three more environments. Part (a) shows the impact of the intermittency costs as the time scale of solar output observations increases.[45] We find

[44]We follow the estimation method of Bushnell et al. [2008] here. See Appendix B for details on the estimation and model.

[45]Appendix B provides modeling details for the model with multiple solar observations per hour.

Figure 6: Costs from unforecastable intermittency, geographic dispersion, and different real-time pricing regimes



a \$3.5 decrease in the social cost of solar per MWh with 60 minute data. Our results on the impact of the solar generation time scale are roughly consistent with results in Mills and Wiser [2010]. Thus, the high frequency nature of solar intermittency adds costs.

Our baseline analysis uses data from 58 solar sites within 40 kilometers of the center of Tucson. Part (b) shows the impact of having fewer, more geographically concentrated sites, on the social cost of high solar penetration. Here, we find that the results are quantitatively important: having only sites within 10 kilometers of the center results in social costs that are \$17.5 per MWh higher than the baseline. Relatedly, the R^2 of the solar forecast increases monotonically from 0.866 with the 10 kilometer radius to 0.876 with the 40 kilometer radius. By including all sites within 20 kilometers of the center, the social costs are very similar to the 40 kilometer costs. As a caveat, we note that we are comparing sites that may use slightly different solar technologies and have different environmental characteristics, such as the prevalence of shade.

Finally, Part (c) shows the impact of real-time pricing contracts on the social cost of solar.[46] Adding customers on real-time pricing adds to the social value. For instance, conditioning on 20% solar, adding 75% of customers to real-time pricing contracts adds \$11.1 in social surplus per MWh of solar produced (the dashed line). Yet, a very similar total benefit of real-time pricing is present even in the absence of solar. Thus, overall, real-time pricing has a negligible effect in changing the social costs of large-scale solar.

6 Conclusions

A variety of current and potential policies have stimulated investment in renewable energy generation. Intermittency of renewable generation may have a significant impact on electric grid reliability, system operations, and requirements for backup generation capacity. Because a grid operator must make different long- and short-run decisions in response to intermittent renewable output, we believe that the social costs of intermittency can best be understood in the context of a model that maximizes welfare.

[46] In our model, the system operator chooses the prices for these contracts at the same time as she schedules generators, one day in advance. Appendix B provides modeling details for this model also.

We develop an empirical approach to estimate the social costs of renewable energy accounting for their intermittent nature. Our approach has three parts: (1) a theoretical model that is based on the work of Joskow and Tirole [2007]; (2) a process to estimate and calibrate the parameters of this model using publicly-available data; and (3) a computational approach to evaluate the impact of counterfactual policies and environments.

The biggest limitations of our approach are that we do not optimize over dynamic linkages from period to period[47] and that we do not consider market power, evaluating only the social optimum. Moreover, our assumed operating reserve costs are at best an approximation of reality.

We examined the impact of large-scale solar on the Tucson, Arizona grid area. We find that the social cost of 20 percent solar generation are $138.4 per MWh of which unforecastable intermittency only accounts for $6.1 per MWh. Other environments, such as allowing imports and exports outside of the grid area or allowing a greater fraction of consumers on real-time pricing contracts have very small impacts in terms of changing the social costs of large-scale solar.

Our study has a number of broader implications beyond the results for solar generation in Arizona. First, our finding that the costs of intermittency for solar energy are lower than many industry observers believe may be important. In particular, costs associated with intermittency are a relatively small component of the overall welfare cost of large-scale solar; the bulk of the welfare cost is simply the high installation cost of solar. Our results on intermittency stem from the assumption that utilities implement substantial changes in demand-side management policies, reserve operations, and investment in backup fossil fuel generators in response to substantial solar capacity. 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, our study has implications about the benefits of different renewable energy poli-

[47]Cullen [2010a] estimates a dynamic model of startup costs for generators. A similar model would hugely complicate our analysis.

cies. While we find that an immediate investment in large-scale solar with current technology would reduce welfare, if solar capacity costs drop to \$1.48, large-scale solar PV generation becomes welfare increasing when accounting for the benefit from offset CO_2 emissions.

Finally, our approach can be used to analyze a variety of other energy policies many of which might require changes in complementary policies to minimize social costs. These policies include understanding the social costs of reducing emissions from a renewable energy standard versus a carbon tax; and how technologies such as battery storage and electric cars which change the effective time pattern of demand can change the value of renewable mandates.

References

Baker, E., Fowlie, M., Lemoine, D., and Reynolds, S. (2013). The economics of solar electricity. *Annual Reviews of Resource Economics*, 5:387–426.

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.

Barbose, G., Darghouth, N., Weaver, S., and Wiser, R. (2013). Tracking the Sun VI: An historical summary of the installed price of photovoltaics in the United States from 1998 to 2012. Technical report, Ernest Orlando Lawrence Berkeley National Laboratory.

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

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

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

Borenstein, S. (2012). Effective and equitable adoption of opt-in residential dynamic electricity pricing. Energy Institute at Haas working paper 229.

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

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

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., Mansur, E., and Saravia, C. (2008). Vertical arrangements, market structure, and competition: An analysis of restructured U.S. electricity markets. *American Economic Review*, 98:237–266.

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

CME (2012). Henry hub natural gas futures. `http://www.cmegroup.com/trading/energy/natural-gas/natural-gas.html`. Last accessed September 15, 2012.

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

Cramton, P., Ockenfels, A., and Stoft, S. (2013). Capacity market fundamentals. *Economics of Energy & Environmental Policy*, 2(2).

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.

Cullen, J. and Reynolds, S. (2013). Notes on dynamic competition in wholesale electricity markets. Working Paper.

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, electric disturbance events, tables b.1 and b.2. http://tonto.eia.doe.gov/ftproot/electricity/epm/02261003.pdf. Last accessed September 21, 2012.

EIA (2011a). Continuous Emissions Monitoring System. http://ampd.epa.gov/ampd. last accessed July 15, 2014.

EIA (2011b). EIA Form 860 Detailed data. http://www.eia.gov/electricity/data/eia860/. last accessed July 23, 2014.

EIA (2012). Renewable capacity in the Annual Energy Outlook 2012. http://www.eia.gov/forecasts/aeo/MT_renewable.cfm#wind. Last accessed October 5, 2012.

EIA (2013). Updated Capital Cost Estimates for Utility Scale Electricity Generating Plants. http://eia.gov/forecasts/capitalcost/pdf/updated_capcost.pdf. last accessed August 20, 2014.

EIA (2014). EIA Monthly Report January 2014. http://www.eia.gov/electricity/monthly/current_year/january2014.pdf. last accessed August 20, 2014.

EIA (2014). Levelized cost of new generation resources in the annual energy outlook 2014. http://www.eia.gov/forecasts/aeo/pdf/electricity_generation.pdf. Last accessed August 21, 2014.

Ela, E., Milligan, M., and Kirby, B. (2011). Operating reserves and variable generation. Technical Report TP-5500-51978, National Renewable Energy Laboratory.

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.

EPA and DOE (2013). Technical update of the social cost of carbon for regulatory impact analysis. `http://www.whitehouse.gov/sites/default/files/omb/assets/inforeg/technical-update-social-cost-of-carbon-for-regulator-impact-analysis.pdf`. Last accessed August 9, 2014.

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

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

Fabbri, A., Romn, T., Abbad, J., and Quezada, V. (2005). Assessment of the cost associated with wind generation prediction errors in a liberalized electricity market. *IEEE Transactions on Power Systems*, 20:1440–146.

Fay, D. and Ringwood, J. (2010). On the influence of weather forecast error in short-term load forecasting models. *IEEE transactions on power systems*, 25.

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

Hydro One (2012). Business load forecast and methodology. `http://www.hydroone.com/RegulatoryAffairs/Documents/EB-2012-0031/Exhibit\%20A/A-15-02.pdf`. last accessed July 23, 2014.

Joskow, P. (2011). Comparing the costs of intermittent and dispatchable electricity generating technologies. *American Economic Review: Papers and Proceedings*, 100(3):238–241.

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

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

Kumar, N., Besuner, P., Lefton, S., and Agan, D. (2012). Power Plant Cycling Costs. `http://www.nrel.gov/docs/fy12osti/55433.pdf`. last accessed June 13, 2014.

Lamont, A. (2008). Assessing the long-term system value of intermittent electric generation technologies. *Energy Economics*, 30:1208–1231.

Larson, V. (2013). Forecasting solar irradiance with numerical weather prediction models. In Kleissl, J., editor, *Solar Energy Forecasting and Resource Assessment*, pages 299–318. Academic Press.

Lueken, C., Cohen, G., and Apt, J. (2012). Costs of soilar and wind power variability for reducing CO2 emissions. *Environmental Science & Technology*, 46.

Madaeni, S. and Sioshansi, R. (2013). The impacts of stochastic programming and demand response on wind integration. *Energy Systems*, 4:109–124.

Milligan, M., Ela, E., Hodge, B.-M., Kirby, B., Lew, D., Clark, C., Cesaro, J. D., and Lynn, K. (2011). Integration of variable generation, cost-causation, and integration costs. *The Electricity Journal*, 24(9).

Mills, A., Botterud, A., Wu, J., Zhou, Z., Hodge, B., and Heaney, M. (2013). Integrating solar PV in utility system operations. Technical report, Argonne National Laboratory. Working paper no. ANL/DIS-13/18.

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

Mills, A. and Wiser, R. (2012). An evaluation of solar valuation methods used in utility planning and procurement processes. Technical report, Ernest Orlando Lawrence Berkeley National Laboratory. Working paper no. LBNL-5933E.

Muller, N. and Mendelsohn, R. (2009). Efficient pollution regulation: Getting the prices right. *American Economic Review*, 95(5):1714–1739.

NOAA (2014). NNDC climate data online. `http://has.ncdc.noaa.gov/pls/plhas/HAS.StationYearSelect?datasetname=9957ANX&subqueryby=STATION&applname=`

SRRSBTNSEL&outdest=APPS&dtypesort=dtypeord&stationsort=id. Last accessed July 15, 2014.

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

Reguant, M. (2014). Complementary bidding mechanisms and startup costs in electricity markets. *Review of Economic Studies*.

Roca, M. (2013). Annual solar installs to beat wind for first time. *Bloomberg New Energy Finance*. September 26, 2013 issue.

Schmalensee, R. and Stavins, R. (2013). The SO2 allowance trading system: The ironic history of a grand policy experiment. *Journal of Economic Perspectives*, 27(1):103–122.

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 (2012). Tucson electric power 2011 renewable energy standard and tariff compliance report. http://www.azcc.gov/Divisions/Utilities/2011\%20TEP\%20REST.pdf. Last accessed August 16, 2014.

Tuohy, A., Meibom, P., Denny, E., and OMalley, M. (2009). Unit commitment for systems with significant wind penetration. *IEEE Transactions on Power Systems*, 24:592601.

UniSource (2012). Unisource energy corp 10-K report for 2011. http://http://www.sec.gov/Archives/edgar/data/100122/000119312512082165/d285669d10k.htm. Last accessed August 19, 2014.

Weron, R. and Misiorek, A. (2004). Modeling and forecasting electricity loads: a comparison. *Eurpoean Electricity Market EEM-04 Conference Proceedings*.

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

Wolfram, C. (2005). The efficiency of electricity generation in the U.S. after restructuring. In Griffin, J. and Puller, S., editors, *Electricity Deregulation: Choices and Challenges*. University of Chicago Press.

Appendix A: Derivations of demand-side properties

Lemma 1. *With demand specified in (1) and retail price fixed at $\overline{p}$, VOLL is constant across states. The relation between VOLL and v is monotonic and satisfies*

$$v = \left((1-\eta)\overline{p}^{-\eta}\text{VOLL} + \eta\overline{p}^{1-\eta}\right)^{\frac{1}{1-\eta}}. \tag{9}$$

Proof

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

$$= \frac{\overline{D}(\frac{1}{1-\eta})(v^{(1-\eta)} - \eta\overline{p}^{(1-\eta)})}{\overline{D}\overline{p}^{-\eta}} = \frac{v^{1-\eta}\,\overline{p}^{\eta} - \eta\overline{p}}{1-\eta}$$

Solving for v in terms of VOLL and the other parameters, we obtain the expression in the statement of the lemma.

Lemma 2. *Assume that the demand that can be curtailed in any period is limited by the minimum demand scale, $\overline{D}^{min}(\vec{w})$. If $p_c < v$, then curtailment quantity is bounded above by $\overline{D}^{min}(\vec{w})\,[\overline{p}^{-\eta} - p_c^{-\eta}]$. The welfare loss associated with curtailment quantity z is:*

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

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

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

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

Appendix B: Extensions to the model

Multiple solar observations per hour

Most of our specifications, including our base specification, include solar data measured at the 15-minute level even though decision-making occurs at the one-hour level. In the interest of ease of exposition, the model described in Section 3 considered the case where the decision-making, solar output, and load all occur at the hourly level. We now consider the case where there are $Y \geq 1$ solar observations per hour period.

Allowing for multiple solar observations per hour results in $\overline{S}$ being vector-valued, $\overline{S} \equiv (\overline{S}^1, \ldots, \overline{S}^Y)$. Let $\overline{S}^{mean}$ denote the mean over the Y values, and $\overline{S}^{min}$ the minimum over the Y values.

With this new model, (3), the transmission fixed costs will then include $\overline{S}^{mean}$ instead of $\overline{S}$. Equation (5), the outage event, substitutes $\overline{S}^{min}$ for $\overline{S}$. Finally, for the system operator[s second-stage maximization problem W in (6), the PC term becomes:

$$\sum_{y=1}^{Y} PC\left(\overline{D}\overline{p}^{-\eta} - z - n^{SL}\overline{S}^y + LL(\overline{D}\overline{p}^{-\eta} - z - d^{SL}n^{SL}\overline{S}^y), \vec{x}(\vec{on})\right).$$

The model is then computable using this modified version of the operator's problem.

Real-time pricing

For this extension, we specify that an exogenous fraction γ of customers are on real-time pricing (RTP) contracts, with the remaining customers paying a fixed retail price. We assume that the system operator sets the real-time prices at the same time as scheduling generators, which is one day in advance. This RTP specification is similar to Borenstein and Holland [2005]. In this extension, we eliminate demand-side management through curtailment contracts that are in the base model.

As before, $\overline{p}$ is the fixed retail price for customers not on RTP. Let $p^{RTP}(\vec{m}, \vec{w})$ be the real time price, where we make explicit that this price is a function of maintenance status and the weather forecast. The demand for electricity is then:

$$Q^D(p^{RTP}(\vec{m}, \vec{w}), \overline{p}, \overline{D}) = \begin{cases} \overline{D}(1-\gamma)\overline{p}^{-\eta}, & p^{RTP}(\vec{m}, \vec{w}) > v \\ \overline{D}(\gamma p^{RTP}(\vec{m}, \vec{w})^{-\eta} + (1-\gamma)\overline{p}^{-\eta}), & p^{RTP}(\vec{m}, \vec{w}) \leq v. \end{cases} \quad (10)$$

Our computation of the RTP solution solves for the optimal $p^{RTP}(\vec{m}, \vec{w})$ via a grid search. We let p^{RTP} be bounded between \$0 and \$5000 per MWh.

Finally, note that there are many other potential demand-side management tools besides the two that we consider. For example, Borenstein [2012] explores the use of critical peak pricing (CPP) for residential customers. Customers on a CPP plan pay relatively low retail electricity prices except during a limited number of days of exceptionally high demand. These types of contracts could also be considered within our framework.

Imports and exports

Finally, we add imports and exports to our model by introducing a net import supply function for the home (TEP, for our application) grid area. Denote the net import quantity as q^I; a positive value indicates imports into the home area and a negative value indicates exports from the home area. Correspondingly, denote the price of imports p^I. We assume that the system operator is faced with a net supply curve for imports of:

$$q_t^I = \alpha_t^I + \beta^I \log p_t^I, \tag{11}$$

where α_t^I, the import supply level, varies across hourly time periods t, and β^I is the net import supply semi-elasticity.

Our model with imports and exports is as follows. Net import decisions are made one day ahead, at the same time as scheduling generators. For each hour-long period, the operator observes $\vec{w}$, $\vec{m}$ and the net import supply curve for that time, and then commits (except in the case of system outage) to a point on the net import supply curve.

The operator's objective function includes the cost (benefit) or imports (exports), but is otherwise the same as in the base model – she does not internalize the fact that if there is a system outage, she may not export the amount that she promised, which may cause system outages in other localities. In this extension, we do not allow demand-side management. The reason is that the system operator can obtain electricity at lower social cost by importing it than by curtailing demand. Our computation of the import/export solution solves for the optimal import quantity, as a function of $(t, \vec{w}, \vec{m})$, via a grid search.

We estimate the parameters of net import supply using the methods of Bushnell et al. [2008]. Specifically, we parametrize α_t^I to include the cooling and heating needs of the surrounding states of California, Nevada, Utah, and New Mexico. For each of these states at each hour, we include for α_t^I the mean cooling and heating degree days and their squares.[48] We calculate the mean by averaging over all weather stations within a state. In addition to these 16 variables, we also include indicators for month, day of the week, and hour of the day in α_t^I. We also obtain p^I and q^I as described in Section 4.1.

As in Bushnell et al. [2008], we estimate (11) with an instrumental variables regression and instrument for p^I with the local (TEP) load, as local load is correlated with p^I but not with the operator's chosen import and export quantity. The unit of observation for our analysis is every hour during our sample period. We estimate $\beta^I = 8,862.2$ (with a standard error of 963.0). Our mean estimated α_t^I is $-30,616.4$. At the mean α_t^I, the system operator would pay \$32.7 per MWh to import 300 MW in an hour and would receive \$30.6 per MWh to export 300 MW in an hour.[49]

Appendix C: Additional tables

Table C.1: Summary statistics for startup costs of TEP generators, 2011 – 2012

Unit type	# units	Additional fuel costs per startup	Capital and maintenance costs per startup	Total costs per startup
Solar PV	2	0	0	0
Coal (< 300 MW)	3	20.3	132.7	152.9
Coal (> 300 MW)	3	29.0	76.3	105.3
Natural gas combined cycle	1	1.0	56.3	57.3
Natural gas steam turbine	3	1.0	56.3	57.3
Natural gas gas turbine	6	1.9	87.0	88.9
Potential new natural gas combined cycle	–	1.0	56.3	57.3
Potential new natural gas gas turbine	–	1.9	87.0	88.9

Note: All costs in \$/MW. Source: Kumar et al. [2012] .

[48] HDD is defined as 0 if the day's mean temperature is greater than 65 and it is equal to (65 - day's mean temperature) if the mean is less than 65. CDD is 0 if the days mean temperature is less than 65, and as (day's mean temperature - 65) if the mean is greater than 65.

[49] Our numbers cannot directly be compared to Bushnell et al. [2008], since their equivalent of q^I included local production by a fringe.

Table C.2: Estimation of daytime load and solar output forecasts

	Load (MWh)			Solar output (kWh)		
	Coefficient on spline for			Coefficient on spline for		
	1st decile	5th decile	10th decile	1st decile	5th decile	10th decile
Temperature	–25.7** (1.4)	–2.3 (1.3)	39.0** (1.0)	4.3** (0.8)	1.5* (0.7)	1.9** (0.6)
Dew point	0.6 (1.3)	7.7** (1.9)	22.3** (1.5)	–4.1** (0.7)	–3.4** (1.0)	–1.3 (0.8)
Relative humidity	8.1* (3.9)	–10.8** (1.4)	–3.0** (0.6)	10.2** (2.1)	2.8** (0.8)	1.6** (0.3)
Wind	5.9* (2.7)	–1.0 (3.3)	–2.1** (0.5)	–3.7* (1.5)	6.4** (1.8)	1.2** (0.3)
	2–15%	38–60%	78–94%	2–15%	38–60%	78–94%
Cloud cover	568** (52)	296** (51)	502** (102)	–108** (28)	–238** (27)	–332** (55)
	Coefficient on hour			Coefficient on hour		
	1	4	6	1	4	6
Hours since sunrise, AM	–146.2** (7.7)	–55.5** (6.3)	0.7 (7.6)	43.3** (4.2)	35.0** (3.4)	36.4** (4.1)
Hours till sunset, PM	306.8 (22.5)	49.0** (5.2)	8.6 (6.1)	63** (12.2)	52.4** (2.8)	35.1** (3.3)
Temp × cloud	–7.5**	(0.7)		1.5**	(0.4)	
RH × cloud	–3.4**	(0.7)		1.0**	(0.4)	
Wind × cloud	–1.6*	(0.7)		–4.6**	(0.4)	
Dew × cloud	5.8**	(0.8)		–0.2	(0.5)	
6AM dummy	132**	(8.3)		–5.3	(4.5)	
…						
12PM dummy	667**	(31.1)		124**	(16.8)	
…						
6PM dummy	118**	(20.3)		–93**	(11.0)	
R–squared	0.959			0.878		
MAPE	3.4%					
rRMSE				0.234		
Correlation of residuals			0.093**			

Note: Model estimated with a SUR specification. Unit of observation is 15 minute interval in May 2011 – April 2012. Standard errors are clustered at hour level. Number of observations is 17,825. We include as regressors day-of-week and month-of-year indicators and full sets of spline coefficients. MAPE is the mean absolute percentage error. rRMSE is the relative root mean squared error.
** Statistically significant at 1% level.
* Statistically significant at 5% level.

Nighttime Load Forecast

Table C.3: Estimation of nighttime load forecast

	Load (MWh)		
	Coefficient on spline for		
	1st decile	5th decile	10th decile
Temperature	–22.0**	2.2	36.7**
	(2.8)	(2.5)	(5.3)
Dew point	3.9	8.2*	12.7**
	(3.0)	(3.2)	(3.3)
Relative humidity	–0.3	–6.3*	–3.1**
	(9.1)	(3.0)	(1.0)
Wind	–5.0	–3.2	–5.4**
	(4.1)	(4.0)	(1.2)
	2–15%	38–60%	78–94%
Cloud cover	257**	83	31
	(74)	(69)	(130)
Temperature × cloud cover	–5.5**	(1.2)	
Relative humidity × cloud	–1.6*	(0.8)	
Wind × cloud cover	4.1**	(1.4)	
Dew point × cloud cover	5.1**	(1.5)	
9PM dummy	263**	(5.5)	
…			
3AM dummy	–67**	(4.1)	
R–squared	0.945		
MAPE	3.2%		

Note: Model estimated with OLS. Unit of observation is 15 minute interval in May 2011 – April 2012. Standard errors are clustered at hour level. Number of observations is 17,311. We include as regressors day-of-week and month-of-year indicators and full sets of spline coefficients. MAPE is the mean absolute percentage error.
** Statistically significant at 1% level.
* Statistically significant at 5% level.

Table C.4: Detailed results for base specification in Table 6

Fraction of generation from solar	0%	10%	15%	20%
Solar PV capacity (MW)	0	941	1,411	1,881
Solar production (1000 MWh/year)	0	1,694	2,541	3,387
Load (1000 MWh / year)	16,937	16,937	16,937	16,937
New 191MW natural gas generators (#)	10	8	8	7
Foregone new gas generators (#)	0	2	2	3
Scheduled non-solar prod. + res. (1000 MWh/year)	24,021	22,466	21,852	21,216
Realized non-solar prod. + reserves (1000 MWh/year)	24,003	22,449	21,835	21,199
Reserves as % of energy consumed	30.5	32.1	33.6	35.2
Average system outage prob.	4.76e-5	5.82e-5	5.81e-5	8.4e-5
Curtailment price p_c ($/MWh)	661	469	431	804
Total curtailment (as % of total load)	0.11	0.19	0.14	0.24
Prob. of some curtailment Jul. 12PM	0.033	0.032	7.8e-5	2.6e-7
Prob. of some curtailment Jul. 6PM	0.097	0.37	0.33	0.58
Production costs	437.2	380.0	355.2	332.2
Reserve costs	78.1	81.5	82.8	84.8
Gas generator investment costs	2,090	1,672	1,672	1,463
Solar capacity investment costs	0	4,148	6,221	8,295
Transmission FC	331.4	319.4	317.4	316.2
Transmission line losses (1000MWh/yr)	1,480	1,378	1,335	1,287
Prob. natural gas generator scheduled	0.97	0.97	0.95	0.90
Loss in surplus relative to baseline (million $/year)	–	214.5	339.7	468.9
Loss in surplus per unit solar production ($/MWh)	–	126.7	133.7	138.4
Loss in $ surplus per ton CO_2 reduced	–	293.1	283.5	279.1
NO_x emissions (1000 metric tons / year)	12.8	12.6	12.2	11.8
SO_2 emissions (1000 metric tons / year)	6.3	6.3	6.1	6.0
CO_2 emissions (mill. metric tons / year)	15.1	14.4	13.9	13.4

Note: Cost variables are measured in millions of dollars per year.

Table C.5: The social costs of large-scale solar including startup costs

Fraction of generation from solar	0%	10%	15%	20%
Startup costs	20.9	24.0	29.8	38.7
Loss in surplus per unit solar prod. ($/MWh)	–	128.5	137.2	143.7

Note: Cost variables are measured in millions of dollars per year. Last two rows include the startup costs.