

参赛论文

参赛学生姓名: Tsz Sen Zhu

中学: BASIS INTERNATIONAL SCHOOL GUANGZHOU

省份: Guangdong District

国家/地区: China

指导老师姓名: Lingshu Zhong

指导老师单位: School of Systems Science and Engineering, Sun Yat-Sen University

论文题目: Information, Action, and Incidence: The Heterogeneous Economic Value of Forecasts across Virtual Power Plants

Information, Action, and Incidence: The Heterogeneous Economic Value of Forecasts across Virtual Power Plants

Tsz Sen Zhu

BASIS INTERNATIONAL SCHOOL GUANGZHOU

Abstract

The intermittency and heterogeneity of renewable energy sources create a difficulty in the coordination of energy source distribution. Virtual power plants (VPPs) coordinate many small energy resources and operate them as a single market participant. When a VPP is in an electricity market, system operators predict demand one day in advance so that enough generation and backup capacity can be scheduled. This paper examines the economic value and impacts of forecasts, especially if they are wrong. A modeled market is constructed from public California Independent System Operator data with a network based on the open PGLib-OPF five-bus system. Open-Meteo weather records and European Central Bank exchange rates support the construction and calibration of inputs. In each experiment, the VPPs and actual operating conditions remain unchanged. Only the grid operator's day-ahead forecast changes. In the scarcity experiment, 23 dates produce complete comparisons. Forecast error increases average system cost by USD 7.959 million and unserved electricity demand by 774.54 MWh per day, while modeled VPP profit increases by USD 117.21 thousand per VPP per day. The cost of unserved electricity makes up 97.32% of the system-cost increase. The second experiment considers sufficient reserves, imports, and capability to eliminate shortages. Across 14 dates, the VPP profit effect falls from USD 126,655.72 to USD 112.70 per VPP per day, a 99.911% reduction. A further settlement sensitivity assigns additional balancing quantities while leaving the physical operation unchanged. Under the specification, the remaining VPP profit effect becomes USD -743.77 per VPP per day. The results show that a large VPP gain occurs in extreme scarcity, and when sufficient balancing resources are available, the remaining effect is smaller and depends on how market rules allocate balancing costs.

Keywords: load forecast, day-ahead clearing, virtual power plant, two-settlement market, scarcity pricing, reserve performance, settlement responsibility, load shedding

Contents

List of Symbols and Abbreviations	3
1 Introduction	4
1.1 Research question and unified claim	4
2 Literature Review	4
2.1 Virtual power plants as distributed energy aggregators	4
2.2 Forecast value in electricity markets	4
2.3 Constraints, scarcity, and nonlinear price responses	5
2.4 Private VPP returns and social resource cost	5
2.5 Household incidence, flexibility, and fairness	5
2.6 Dispatch and contractual coverage	5
2.7 Adjacent literature	5
3 Market Model	5
3.1 Experimental system	5
3.2 Day-ahead Co-optimization	5
3.3 Real-time Recourse	6
3.4 VPP Settlement and Modeled Profit	6
3.5 Three non-interchangeable accounts	6
3.6 Fixed-dispatch contractual-coverage sensitivity	7
4 Mechanism	7
4.1 From forecast to incidence	7
4.2 Continuous quantity effects and regime switches	7
4.3 Separation of private and social directions	9
4.4 Negative controls and boundaries	9
4.5 Testable Propositions	9
5 Experimental Design	9
5.1 Evidence structure	9
5.2 Data sources	9
5.3 Two-bus mechanism	9
5.4 High-fidelity cases	10
5.5 Unified thirty-day scarcity experiment	10
5.6 Zero-loss adequacy and settlement sensitivity	10
5.7 Outcome definitions and acceptance	11
5.8 Pair construction, missing arms, and estimands	11
6 Results	11
6.1 Mechanism map of loss, zero, interaction, and ambiguity regions	11
6.2 High-fidelity dates results	11
6.3 Inactive slack network forecasting and unidentified binding effect	12
6.4 Thirty-date extension identifies a scarcity regime	12
6.5 Adequacy removes shortage, substantially reducing the apparent profit effect	13
6.6 Contractual cash allocation can reverse the residual VPP-account effect	13
6.7 Four-channel decomposition	13
6.8 Symmetric forecast directions reject a one-sided explanation	13
6.9 The unified result is stable to reconciliation	13
7 Robustness and Identification	14
7.1 Duplicate-record reconciliation	14
7.2 Missing-outcome sensitivity	15
7.3 Two negative controls	15
7.4 Binding congestion and structural non-identification	15
7.5 Missingness and numerical failures	15
7.6 Zero-loss gate and common adequacy	15
7.7 Contract-coverage grid and decomposition checks	16
7.8 Established results of robustness checks	16

8 Analysis and Distributional Consequences	16
8.1 VPP gains are private settlement gains	16
8.2 Joint interpretation across two regimes	16
9 Discussion, Policy Implications, and Limitations	16
9.1 Forecast quality should be valued by state	16
9.2 Settlement design must distinguish dispatch and VPP fault	16
9.3 A practical market-monitoring rule	17
9.4 Evidence interpretation and limitations	17
10 Conclusion	17
References	18
Acknowledgements	19
List of Symbols and Abbreviations	

Symbol	Definition
VPP DA	Virtual power plant Day-ahead
RT	Real-time
LMP	Locational marginal price
d, t, n, g, i	Date, hour, bus, conventional generator, and VPP indices
z	Forecast arm: accurate source forecast, symmetric 5% underforecast, symmetric 5% overforecast
$q_{itd}^{DA}, q_{itd}^{RT}$	Day-ahead net energy injection, dispatched real-time net energy injection
r_{itd}^{DA}	Day-ahead reserve capacity assigned to VPP i
$\lambda^{DA}, \lambda^{RT}$	Day-ahead and real-time nodal energy prices
μ^{DA}	Day-ahead reserve-capacity price
Π_{id}	VPP profit
H_d	Aggregate household net payment
C_d^{sys}	System resource cost
Q_d^{shed}	Real-time load shedding
P_d	Indicator for whether the forecast changes accepted DA energy or reserve.
S_d	Indicator that the active constraint or marginal resource changes.
$r_{itd}^{call}, d_{itd}^{call}$	Called reserve, dispatch credited against call
B_{id}, K_{id}	Balancing quantity outside call credit, called quantity not covered by dispatch adjustment
m_B, m_K	Hypothetical contract-coverage multipliers for B and K
c_d^{imp}	Date-specific marginal cost of added resource
MRE	Minimum relevant effect for VPP profit

1 Introduction

Renewable generation, electric vehicles, and distributed storage are moving power-system flexibility away from a small number of centralized assets and toward millions of heterogeneous devices, yet these systems cannot enter the electricity market as a whole. A virtual power plant (VPP) aggregates such devices and other small resources so that they can be coordinated and scheduled as a market participant in the electricity market [1,2]. In a day-ahead electricity market, the grid operator uses a system level forecast of nonflexible load and available transmission capacity to clear day-ahead energy and upward reserves before electricity is delivered. When actual conditions arrive, the real-time market must correct any imbalance. A forecast error comes into play in these conditions. A small forecast error might require only a small adjustment, but an error near a reserve or scarcity boundary could change how resources are dispatched, prices are formed, and whether demand can be served. This paper investigates how forecasts and errors present in the forecasts create heterogeneous economic impacts by moving through clearing, dispatch, scarcity, and settlement. We design the electricity market with VPP inputs, realized load, and network capability held common.

Forecasting performance is often summarized statistically in scores such as mean absolute error or root mean square error. These measures indicate how far a forecast is from realization, but they do not show whether the error changes a market decision. However, in an electricity market, two forecast errors might lead to significantly different market consequences if an error alters a VPP choice or crosses an economic boundary. The economic effect of a forecast is therefore jointly determined by the magnitude of the information change, the local response of the market solution, and any regime switch [3–8].

This paper focuses on the distributional and institutional complication of the question. Better system information might not increase an individual VPP's income, and a VPP's settlement gain might not reduce physical resource cost. In a two-settlement market, an undercommitted resource may receive a large real-time deviation payment in scarcity, yet scarcity alone cannot identify the source of the private gain. Household payments may move in another direction as they are part of a transfer ledger. Thus, statements such as “a forecast error is costly” or “forecast errors benefit VPPs” are incomplete as they do not consider the market state and the heterogeneous economic impacts of the erroneous forecasts.

This distinction changes the empirical burden, as a large VPP gain in a load-shedding run identifies a scarcity settlement outcome. An adequacy enhanced inference requires comparisons of common dates after involuntary shortage is eliminated. Any post-dispatch overlay must distinguish between real-time quantities from metered autonomous delivery, and we need to identify the contractual basis, payer, and receiver.

1.1 Research question and unified claim

The main research question can be stated as: when the grid operator changes the system load forecast or available transmission forecast used in day-ahead clearing, how do quantity adjustments and cross-region switching affect VPP settlement, reliability costs, and the contractual allocation of bal-

ancing quantities?

Overall, the claim separates two response margins, which is an active constraint region and regime switch. In an active constraint region, a forecast can change schedules and settle quantities continuously. On the other hand, a regime switch can amplify or reverse the regional effect. Net VPP settlement also depends on whether a contract assigns cash adjustment to balancing quantities in real time. These layers determine the market's physical outcomes and incidence resulting from forecast errors.

This paper contributes in three primary ways. First, we design a two-stage electricity market with network constraints to add nuance to earlier work about the values of forecast improvement in VPPs. Second, it identifies the contribution of scarcity through an adequacy comparison experiment, separating it from a contract allocation overlay. Third, we trace the economic meaning of real-time energy quantities and we report a date-level experiment with a transparent data-construction pipeline.

2 Literature Review

2.1 Virtual power plants as distributed energy aggregators

VPPs aggregate distributed energy resources, such as batteries, electric vehicles, or flexible loads, into one single participant in an electricity market. Aggregation is economically important because many of the individual assets are either too small or too unstable to satisfy market participation requirements on their own. A VPP can aggregate these resources, combine them and submit an energy schedule, hold reserve headroom, and respond to real-time conditions, such as fluctuations in the market [1,2,9–12]. Recent work also emphasizes that a VPP's credible capacity and locational deliverability determine whether its nominal flexibility can be treated as a reliable market product [24].

While creating controllability, VPP aggregation also creates forecasting exposure. A VPP operator must decide on how much energy to buy or sell and how much flexibility to reserve before demand, renewable energy production, device availability, and network conditions are realized. Current forecasting studies frequently judge VPP performance via statistical loss functions. While this approach is useful for model comparisons, it is incomplete for market valuation because the economic effect of a VPP decision depends on the decision made with the forecast and the price and constraint regime where the decision is settled. Research connecting VPP forecasting to day-ahead income confirms that forecast quality and energy storage scheduling both impact revenue, but it does not establish, by itself, that a reduction in forecast error has the same marginal value [24].

2.2 Forecast value in electricity markets

Electricity markets have a sequential scheduling because generation, demand, and network capability cannot be perfectly known in advance. In an electricity market, day-ahead schedules coordinate expected production and consumption values, while real-time clearing treats deviations in the real market. Therefore, the economic value of information arises from a comparison made between different information sets, not forecast accuracy alone [3–8]. An improvement in the forecast can alter commitment, reserve procurement, charge storage, flexible demand, or network congestion. If none

of these decisions change, the market settlement effect can remain trivial even if the statistical forecast changes.

Prior research about probabilistic forecasting also shows that joint distributions of renewable supply and demand can improve the operating-reserve allocation relative to deterministic forecasts [26]. That result shows a decision-oriented view of forecast quality, namely uncertainty information is valuable when it changes a reliability decision. This paper extends this logic to distributional incidence. It asks whether decision changes benefit the VPP, reduce household payment, and lowers system resource cost simultaneously, or if these accounts move in different directions.

2.3 Constraints, scarcity, and nonlinear price responses

Electricity markets modeled are network-constrained, piecewise systems. For example, in one active set of constraints, a change in a forecast might leave the same generator, line status, and reserve provider in place. However, if the forecast crosses a transmission, ramp, reserve, or feasibility boundary, the solution might be moved to another face of the feasible region. Thus, prices and quantities may change discontinuously, even if the forecast changes linearly [6-8, 13-16].

Scarcity in the market also creates nonlinear responses in the price of energy. In our market, load shedding is registered when the flexibility delivered by our generation cannot cover realized demand. Current research also shows that relatively small changes in supply conditions can create disproportionate impacts in severe shortage events [25], so the physical reliability of the system should also be separately considered.

2.4 Private VPP returns and social resource cost

Market payments and physical resource cost in the electricity market answer different questions. We differentiate market payments and physical resource costs. A VPP might receive a high payment in the real-time market because its injection is valuable at the margin, but the system might experience costs from a large shortage because the injection is insufficient to restore balance. Their signs are thus not required to offset. This distinction is necessary to interpreting a forecast error. A wrong day-ahead forecast can create more deviation quantity and a higher real-time price, so both errors raise the deviation payment received by the VPP. However, if the same error causes load shedding, the private gain and social resource-cost change would both be positive under the paper's wrong-minus-accurate convention. Thus, the result analyzed should be focused on incentive alignment, not accounting contradictions.

2.5 Household incidence, flexibility, and fairness

Households enter VPPs as the owners of devices, which provide flexible load during charging, retail customers, or recipients of reliability service. However, the recorded payment of households is only one dimension of their economic incidence. Current research shows that VPP incentives can reduce peak demand without increasing the financial burden of vulnerable households in a particular program [20]. This finding supports distributional measurements of forecast impacts to households, and also shows why an outcome specific to a program should not be generalized.

Our model records household net payment and household-action shortfall, although not estimating utility, discomfort, or compensating variation. A lower household payment might

therefore occur while the market has greater unserved energy. This paper thus treats payment and reliability separately, and does not infer household welfare from payment alone.

2.6 Dispatch and contractual coverage

A two-settlement market needs to balance incentives and also clearly define what each quantity in the contract represents. Real-time prices reward energy when it is scarce, while reserve-capacity payments reward option value before delivery [6,7,13,17]. In this model, however, q^{RT} is chosen by the system's real-time cost optimization, not a separate action.

The institutional sensitivity therefore is more nuanced. Upward balancing is first mapped against a proxy call from the system forecast shortfall. The residual absolute balancing quantity and the called quantity that are not already covered by the dispatch adjustment become the contract-coverage bases. Both are valued at the marginal cost of the firm-import resource, and do not estimate breach or an authorized fee. The calculations therefore only ask about how profits would change if such cash amounts are assigned by a contract.

2.7 Adjacent literature

Prior work establishes that statistical accuracy does not mean trading profit. For example, forecasting links predictive information to decisions in renewable trading, market rules might alter the forecast value, and forecasting can reduce uncertainty for retailers. The contributions of different sources are defined in Table 1 below.

3 Market Model

3.1 Experimental system

The paper's experimental system is a modeled electricity market with two sequential stages. In the first stage, the day-ahead stage, the system operator receives forecasts of non-flexible demand and available transmission capacity and co-optimizes energy and upward reserve. Each VPP submits the same frozen strategic inputs within a matched forecast-error accurate-error run pair. Only the forecast for system demand, and forecast for capacity in designated mechanism experiments, change. Commitments in the day-ahead market are then fixed.

In the second stage, realized conditions replace the forecasts. The market dispatches available recourse, and prices are settled at real-time nodal prices. The forecast changes the input of the system operator, where the clearing model selects different award amounts and real-time dispatch quantities. Let d represent dates, t represent hours, n represent buses, g represent conventional generators, and $i \in \{1, 2, 3\}$ the VPPs. Also, let $z = 0$ represent the oracle-accurate demand arm and $z = 1$ the source-forecast error arm. In this experimental system, every treatment effect is defined as:

$$\Delta Y_d = Y_d^{\text{wrong}} - Y_d^{\text{accurate}}. \quad (1)$$

3.2 Day-ahead Co-optimization

In the day-ahead market, an selects conventional generators, P_{gtd}^{DA} , VPP net energy positions q_{itd}^{DA} , generator and VPP reserve r_{gtd}^G , and r_{itd}^V , network flows, and reliability slacks. This can be represented by the following formula:

Table 1: Literature table and paper positioning

Study	Forecast object	Market stage	Network and action	Economic outcome	Responsibility allocation	Role in paper
Muñoz et al. (2020)	Renewable forecast used for trading	Market trading	Forecast-to-trade features	Trading performance	Participant exposure	Establishes forecast improvement
Klyve et al. (2023)	PV production forecast used by trader	Day-ahead, intraday	Historical Nordic bidding areas, settlement- rule comparison	VPP income and forecast value	Single- and dual-price imbalance rules	Shows market rules can make perfect forecasts non-profit- maximizing
Wang et al. (2021)	Retail-load uncertainty reduced by data for a retailer	Two-settlement market	Retailer optimization under uncertainty	Profit enhancement and data value	Retailer bears uncertainty exposure	Provides an analytical information-value benchmark
Sarathkumar et al. (2025)	VPP generation forecast used for day-ahead arbitrage	Day-ahead	IEEE 14-bus OPF, storage arbitrage, fixed network assumptions	Forecast error and income frontier	VPP, storage operator	Connects VPP forecasting with arbitrage revenue
Terrén-Serrano et al. (2026)	System renewable supply and demand forecast used for reserve decisions	Day-ahead operation	System-level reserve allocation	Reliability-oriented reserve allocation	Operator decision	Demonstrates the operational value of joint probabilistic forecasts
This paper	Grid-operator system load and forecasts, VPP strategies fixed	Day-ahead energy and reserve and real time	Nodal network, reserve, scarcity, adequate-resource counterfactual	VPP settlement, household cash payment, system resource cost, load loss	Hypothetical coverage of system-dispatched balancing quantities in fixed dispatch	Distinguishes regional and regime-switching effects, identifies scarcity reduction, derives a cash-allocation sign boundary

3.4 VPP Settlement and Modeled Profit

The VPP's modeled daily settlement surplus is

$$\min_{x^{DA}} J_d^{DA}(x; \hat{L}_d, \hat{F}_d) = C^{gen} + C^{deg} + C^{curt} + C^{shed} + C^{res} + C^{hhshort}, \quad (2)$$

At each bus, scheduled VPP injections of energy must cover all scheduled energy withdrawals and exports. This can be shown as:

$$\sum_{g \in n} P_{gtd}^{DA} + R_{ntd}^{DA} + \sum_{i \in n} q_{itd}^{DA} - \hat{L}_{ntd} = \sum_{\ell} A_{nl} F_{ltd}^{DA}. \quad (3)$$

The energy and reserve both have the same headroom. This can be represented by the following formula.

$$P_{gtd}^{DA} + r_{gtd}^G \leq \bar{P}_g. \quad (4)$$

3.3 Real-time Recourse

During real-time clearing, the system considers realized demand, which is represented as L^{RT} , and realized line capacity F^{RT} , while the day-ahead positions remain constant. In the clearing process, the system operator selects recourse generation, feasible VPP injections, flows, reserve deployment, and shortage quantities, which can be modeled as

$$\min_{x^{RT}} J_d^{RT}(x; x_d^{DA}, L_d^{RT}, F_d^{RT}). \quad (5)$$

This model includes the physical reliability costs, real-time load shedding and household-action shortfall, considering the registered penalty of each. q^{RT} is a decision variable in the system optimization, where the solver can redispatch the VPP's full daily trajectory in response to realized conditions. The real-time stage is a full recourse benchmark with perfect information, so this assumption might overstate the flexibility available to storage.

$$\Pi_{vtd} = \sum \left[q_{itd}^{DA} \lambda_{n(i)td}^{DA} + (q_{itd}^{RT} - q_{itd}^{DA}) \lambda_{n(i)td}^{RT} + r_{itd}^{DA} \mu_{td}^{DA} \right] - C_{id}^{deg}, \quad (6)$$

In this equation, λ^{DA} and λ^{RT} represent energy prices at different nodes and μ^{DA} is the reserve-capacity price. The inherited contract for the market includes reserve-capacity revenue but not the separate model activation payment or contractual coverage adjustment. Equation (6) settles the system-dispatched adjustment $q^{RT} - q^{DA}$ once at the real-time nodal price.

For any bilinear payment $q\lambda$, the symmetric decomposition can be represented as

$$\Delta(q\lambda) = \bar{\lambda} \Delta q + \bar{q} \Delta \lambda, \quad \bar{q} = \frac{q_1 + q_0}{2}, \quad \bar{\lambda} = \frac{\lambda_1 + \lambda_0}{2} \quad (7)$$

and the equation decomposes a payment difference into quantity and price contribution, so it produces the price and injection-quantity channels. We use the same identity when calculating real-time deviations and reserve payments.

3.5 Three non-interchangeable accounts

The economic model of this paper separates VPP profit, household payment, and system resource cost as three different accounting perspectives. System resource cost can be represented by the following equation:

$$C_d^{sys} = C_d^{gen} + C_d^{deg} + C_d^{spill} + C_d^{loadloss} + C_d^{hhshort} + C_d^{other}, \quad (8)$$

where total system cost includes the costs of generation, degradation, spill, load-loss, household shortfall, and other

physical costs. Market payments are not considered in this calculation because a high nodal price can also create a high physical production cost. Household net payment H_d then records settlement amounts assigned to a household group. As a cash incidence measure, the calculation does not include household income, utility, discomfort, or compensating variation. Therefore, a decrease in H_d is not equivalent to a direct household-welfare improvement. Cash closure is necessary for verifying the transfers without adding them to total welfare.

3.6 Fixed-dispatch contractual-coverage sensitivity

The VPP profit calculation in Equation (6) is used as the gross benchmark. After the electricity system decides all physical operations, we consider two possible cash-allocation adjustments after dispatch. The system's electricity demand forecast belongs to the grid operator, and the VPP is not responsible for producing it, so if the grid operator forecasts demand incorrectly, the model does not directly charge the VPP for the mistake. Moreover, the model does not separately measure what the VPP independently chooses to deliver, but instead calculates q^{RT} , the VPP's real-time amount of energy injection as part of the system operator's optimal dispatch. Therefore, this paper cannot directly conclude VPP fault, but rather asks a narrower question: if a hypothetical contract placed an additional payment responsibility on parts of the system-dispatched balancing quantity, how much would the VPP's profit change?

Let $u_{it} = q_{it}^{RT} - q_{it}^{DA}$ represent VPP i 's balancing adjustment, r_{it}^{DA} represent reserve-capacity award, a_t the positive system forecast shortfall, and R_t the augmented upward-reserve requirement. The proxy call fraction can be expressed as:

$$\kappa_t = \min \left\{ 1, \frac{a_t}{R_t} \right\}. \quad (9)$$

Then, based on the proxy call fraction, the proxy called reserve and the part of upward dispatch mapped to that call can be expressed as:

$$r_{it}^{call} = \kappa_t r_{it}^{DA}, \quad d_{it}^{call} = \min \{ r_{it}^{call}, \max(0, u_{it}) \}. \quad (10)$$

The cap at r_{it}^{call} prevents upward dispatch above the call from being counted twice as service. Then, we define two accounting bases, the balancing quantity outside the called credit, B_i , and the called quantity not covered by the dispatch adjustment, K_i , as shown in:

$$B_i = \sum_t \max \{ 0, |u_{it}| - d_{it}^{call} \}, \quad K_i = \sum_t \max \{ 0, r_{it}^{call} - d_{it}^{call} \}. \quad (11)$$

With a firm-import marginal cost c_d^{imp} on a given date, the balancing-quantity multiplier m_B , and called-quantity coverage multiplier m_K , the VPP's adjusted profit can be represented by

$$\Pi_{id}^{adj}(m_B, m_K) = \Pi_{id} - c_d^{imp}(m_B B_i + m_K K_i). \quad (12)$$

This design evaluates $m_B \in \{0, 0.5, 1.0, 1.5\}$ and $m_K \in \{0, 1.0, 1.5, 2.0\}$, with the central illustration given by $(m_B, m_K) = (1.0, 1.5)$.

Equation (12) is treated as an additional cash-allocation term. To avoid double recovery, this should only be interpreted as a conditional with explicit rules defining the VPP obligation and cost calculation.

The bases B_i and K_i both measure leftover quantities, but account for different quantities.

B_i accounts for the absolute balancing quantity that is not already credited against the proxy reserve call, while K_i captures the portion of the proxy reserve call that is not covered by the mapped upward dispatch. These bases do not behaviorally interpret the VPP. Table 2 explains the sequence from the original forecast to the contract adjustment, separating the market information, decision, and cash allocation.

4 Mechanism

4.1 From forecast to incidence

The electricity model is a sequential mechanism, where an information change goes through a day-ahead plan, constraint set, recourse, and prices before being settled. A forecast error that leaves economically relevant decisions unchanged has no settlement effect, while an error of a similar statistical size could matter significantly if it changes a schedule, quantity, or status. The mechanism includes five stages. First, the model begins with day-ahead forecasting, then day-ahead energy and reserve planning, real-time deviation and scarcity, realized information of price and quantity, and finally settlement and physical incidence. The conditions and separation of the accounts can be explicitly shown in Figure 1.

As shown in Figure 1, a forecast error matters economically only if it changes an active decision, price, reserve position, or scarcity outcome. It also shows that the resulting consequences of a forecast error are divided in three distinct accounts. In the first stage, the operator-side forecasts of non-flexible demand and transmission enter day-ahead clearing about the next day. These forecasts influence the residual demand scheduled resources must cover and transfers that the network must support, but do not directly enter VPP profit. Their economic effect only begins when the day-ahead optimizer selects a different energy schedule.

At the second stage, a changed day-ahead plan alters the system's starting position in the real-time market. A lower schedule can create more upward room, but less scheduled energy, and a higher schedule can cause the reverse. When reserve award is changed, the capacity revenue and amount of flexibility held are both changed. These changes occur even when no generator changes its status and no network constraint changes.

At the third and fourth stages, market conditions are held fixed and are considered across forecast arms. The real-time optimization shows the results of entering the same real-time state from different day-ahead positions. These responses determine nodal prices and the quantities settled at the prices. The final stage assigns cash flows to the VPP and records transfers in the system's account. The whole chain thus shows a statistical forecast creates value by changing an economically relevant decision which influences pricing and settlement.

4.2 Continuous quantity effects and regime switches

Let $x^{DA}(I)$ represent the day-ahead plan in the information state I , and $\mathcal{A}(I)$ represent the plan's active constraints and

Table 2: Sequence of stages in VPP model

Stage	Event	Party held accountable	Obligation and Cash Flow	Counterparty
System forecast	A VPP grid operator forecasts nonflexible demand and available transmission for day-ahead clearing	Grid operator	Information input with no direct VPP charge	Market clearing process
VPP submission	Submit the frozen energy schedule portfolio and strategy inputs	VPP	Offers and resource constraints	Day-ahead market
Day-ahead award	Energy clearing and realized reserve capacity	Market operator	VPP receives energy and reserve-capacity settlements	Market settlement account
Real-time dispatch	The system operator reoptimizes feasible VPP energy injections in realized conditions	System operator	The system dispatch amount is q^{RT} , and the VPP follows the modeled solution	Real-time market
Proxy-call mapping	System shortfall is converted in the call fraction κ_t	Accounting rule	Map dispatch upwards	Settlement record
Balancing quantity outside call credit	Compute B_i from system-dispatched quantities	Accounting rule	Profit adjustment to $c_d^{imp} m_B B_i$	Unspecified without contract
Call quantity coverage	Compute K_i from system-dispatched quantities	Accounting rule	Profit adjustment to $c_d^{imp} m_K K_i$	Unspecified without contract

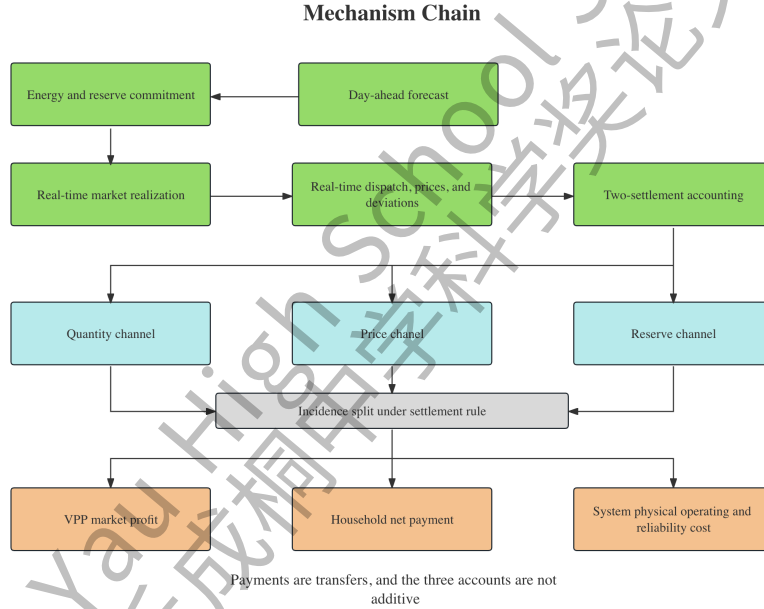


Figure 1: Forecast to incidence mechanism. A forecast error creates a significant economic impact only if it crosses an economically active boundary and meaningfully changes a commitment, price, reserve allocation, or scarcity quantity. The settlement then distinctly divides results in the market among the VPP, household, and system accounts.

marginal resources. Indicators of this plan are

$$P_d = \mathbf{1}\{x_d^{DA}(I^{wrong}) \neq x_d^{DA}(I^{accurate})\}, \quad (13)$$

$$S_d = \mathbf{1}\{A_d(I^{wrong}) \neq A_d(I^{accurate})\}.$$

P_d indicates whether the day-ahead allocation differs between the wrong and accurate forecast arms. The switch indicator, S_d , records whether the set of active constraints or marginal resources change.

Three potential cases come from this equation. First, if $(P_d, S_d) = (0, 0)$, the forecast effect is an economic zero. Second, if $(P_d, S_d) = (1, 0)$, the information change alters quantities in the same region. In two scenarios, let ΔI represent a change in forecast information. The profit response in this change can be represented by the following:

$$\Delta \Pi_i \approx \nabla_x \Pi_i^\top \frac{\partial x^{DA}}{\partial I} \Delta I, \quad S_d = 0. \quad (14)$$

where $\partial x^{DA}/\partial I$ represents how the accepted day-ahead plan responds to changes in the forecast information and $\nabla_x \Pi_i$ represents how the change in the plan affects the VPP's profit. Third, if $S_d = 1$, the forecast crosses an active set, marginal resource, or feasibility boundary. For example, a previously slack transmission limit may bind. The price and recourse might change nonlinearly in these situations.

Two common errors might be assumed. First, no regime switch does not directly imply zero effect because schedules and reserves can move continuously on the same face of the feasible set. Second, a large effect cannot be directly attributed to forecast magnitude alone because a smaller error

can also create a large result if it crosses a large boundary.

This mechanism explains why the experiments need several kinds of evidence. The two-bus surface shows the impact of active-region boundaries. High-fidelity cases show whether the same distinction remains visible when hourly schedules, reserve headroom, and network constraints interact with each other. The multi-date scarcity experiment also evaluates repeated outcomes in a regime where shortage boundaries are crossed, and the adequacy experiment supplies deliverable recourse to move forecast contrasts away from the shortage boundary.

4.3 Separation of private and social directions

In a hypothetical situation, an underforecast causes the day-ahead market to schedule too little supply compared to realized demand. At real time, a VPP providing positive deviation energy might be paid at a scarcity price, and its profit might be positive, as shown in:

$$\Delta \Pi_i > 0. \quad (15)$$

At the same time, the system may shed load or have a shortage penalty, which can be shown as:

$$\Delta C^{sys} > 0, \quad \Delta Q^{shed} > 0. \quad (16)$$

While VPP payment is determined by marginal prices, system cost includes physical reliability loss. This separation is economically meaningful because the market could reward a resource for its marginal value even when its forecast error is socially costly.

This can be illustrated by an underforecast. If too little supply is scheduled, a real-time optimizer might increase a VPP's injection. When marginal price rises due to scarcity, the VPP would earn more than under an accurate forecast. This payment thus rewards the VPP's marginal flexibility, but the system would simultaneously incur losses. The accounting distinction thus prevents a larger VPP payment from being directly added through system resource cost and positive VPP profit offsetting the value of unserved energy.

This coexistence is the paper's main investigative question. Paying a VPP to balance energy can be marginally efficient, but there is a mismatch between the system forecast error and the shortage cost assigned based. Diagnosing this mismatch requires evidence on the forecast origin, physical adequacy, and contract governing the reserve obligations.

4.4 Negative controls and boundaries

We use two negative controls in this paper. First, a capacity forecast error should be inactive when both capacity assumptions leave the same line constraint slack. Second, a demand error can be an economic zero if the induced schedule and price changes are below the minimum relevant effect.

A separate boundary comes from the multiplicity of optimization. If several allocations attain the same primary objective but imply different settlements, the forecast effect is not point identified. Selecting the most convenient allocation would change the solver choice into an economic conclusion, so we report the high-fidelity congestion case as unidentified until an allocation rule preserves the primary optimum, or if the settlement is the same in the optimal set.

4.5 Testable Propositions

The experiments in the next section are based on these propositions:

Proposition 1. Regional Quantity Response: When holding realized conditions fixed, a forecast change can influence hourly energy or reserve awards and therefore alter VPP profit. If it does not change the accepted plan or any economically relevant prices, the effect is negligible.

Proposition 2. Regime-switch Amplification: When a forecast change causes a regime switch, such as switching an active constraint, marginal resource, or feasibility state, the magnitude and direction of VPP profit would change nonlinearly. In a shortage situation, both the revenue deviation and system reliability costs might rise simultaneously.

Proposition 1 does not predict that a response must be profitable. Instead, it predicts that forecast information can affect settlement through continuous quantity movements, while leaving the realized conditions constant. Proposition also predicts the possibility of a nonlinear amplification or signal reversal.

The two-bus surface separates regional quantity responses from regime switches. High-fidelity dates can trace commitment and settlement channels, while the unified 30-date experiment tests are repeated. The adequacy experiment then holds the forecast comparisons closely aligned while removing the shortage mechanism. The comparison between these regimes identifies if the large effect is associated with scarcity and if smaller gross settlement differences exist under zero load loss. Reconciliation for duplicate records and the responsibility overlay are evidence checks, not additional propositions.

5 Experimental Design

5.1 Evidence structure

This section identifies interventions, samples, estimands, and acceptance gates before results are presented. The empirical experiments are organized in four layers by role, which can be shown in Table 3.

5.2 Data sources

The experiment uses publicly available data from different sources. Network parameters are taken from PGLib-OPF v23.07 [18]. Market forecast, demand, and price data come from CAISO OASIS [19]. Weather variables come from Open-Meteo [29].

The accurate-forecast treatment replaces day-ahead demand with realized demand, while the underforecast and overforecast treatments use $(0.95L_{d,t}^{RT})$ and $(1.05L_{d,t}^{RT})$.

The sample contains all 30 eligible dates remaining after 19 dates are used for model development or analysis excluded from a 49-date candidate set. Eligible dates must have complete 24-hour records for CAISO demand, prices, required weather variables. Dates are ordered in predefined strata using a deterministic SHA-256 key. The procedure produces an evaluation from a sample covering September 16th to November 30th 2024.

5.3 Two-bus mechanism

The artificial fixture two bus mechanism crosses actual line-capacity multipliers 0.09, 0.10, and 0.11, which limit

Table 3: Experimental layers and evidence roles

Layer	Design	Purpose	Evidence grade	Permitted inference
Mechanism phase map	Two buses, two hours, actual line-capacity multipliers crossed with positive capacity-forecast bias, four demand-by-capacity arms per cell	Locate loss, economic zero, interaction, and ambiguity regions	Mechanism development and frozen model interpolation	Explains why and where a constraint crossing can matter
High-fidelity cases	Five buses, 24 hours, three VPPs, matched forecast arms on selected dates	Test schedule, reserve, price, and settlement transmission in the full model	Development evidence and negative controls	Shows state dependence in identified cases without estimating a market average
Unified 30-date scarcity experiment	Same 30 frozen input dates, one reconciled effect per date	Test repeated scarcity-state direction while preventing duplicate counting	Paired-date evidence with 23 complete and 7 incomplete dates	Describes the identified dates in the model but does not provide a population-representative causality
Adequacy and contract-coverage sensitivity	Same 30 dates, four forecast arms, common reserve, locational, firm import, and ramp capability, with a 16-point allocation grid	Separate extreme scarcity, forecast direction, and contractual coverage	Paired zero-loss sensitivity with common-date and whole-date resampling checks	Tests whether gross or adjusted VPP-profit effects survive without load loss without establishing breach

the transmission capacity, with positive day-ahead capacity biases 0.010, 0.025, and 0.050, where overestimation is modeled. Each cell contains demand-accurate/wrong and capacity-accurate/wrong arms. VPP 1 effects are classified with a 1 USD/day minimum relevant threshold when accounting for economic gains or losses, while VPP 2 is located where the forecast error should have low economic effect as an inactive comparison. A separate four-cell block for parameter combinations in the study is also generated to check whether the local regime pattern is stable within the artificial model.

5.4 High-fidelity cases

Three matched dates isolate their state dependence in the five-bus market. In each pair, the realized demand, actual line capacity, resources, VPP portfolios, strategies, and settlement rules are kept fixed. Then, the grid operator uses a system-level forecast of nonflexible load and available transmission in day-ahead clearing. The source forecast defines the wrong arm, while the realized demand is directly substituted into the operator's day-ahead input in the accurate arm. The profit decomposition retains day-ahead energy revenue, real-time deviation settlement, reserve-capacity revenue, and degradation cost.

A 2×2 diagnostic crosses demand and forecast separately. At the actual line-capacity multiplier 0.90, both the accurate 0.90 forecast and wrong 0.95 forecast remain nonbinding in day-ahead clearing. The binding diagnostic test reduces line capacity enough that transmission becomes economically restrictive. The binding is retained to test identification, not to force an effect estimate. Alternative primary-optimal commitments are then compared before any settlement magnitude is admitted.

5.5 Unified thirty-day scarcity experiment

The extension freezes 30 dates and two demand-forecast arms per date. The date is the only treatment effect unit. Each date contributes at most one wrong-accurate contrast, irrespective to the number of certified numerical records that exist for the

date. If two certified records exist, then their metric-specific mean is formed within a date. On the other hand, if only one certified record exists, that record is directly retained. A date without a certified outcome for both arms remains incomplete and is never assigned zero.

This rule produces 23 complete and 7 incomplete date pairs. Fifteen complete dates have duplicated certified records, and eight have one certified record. In the fifteen duplicate-record dates, the mean relative numerical difference is 0.072% in system resource cost and 0.855% in total VPP profit, with the largest differences being 0.191% for system resource cost and 2.202% for total VPP profit. Reconciliation removes the duplicate counting while preserving a numerical consistency check without creating any new market variation. The 30 source forecasts contain 29 underforecast dates and 1 overforecast date, and all complete pairs used for the headline extension are underforecasts.

5.6 Zero-loss adequacy and settlement sensitivity

The fourth evidence level uses four operator-forecast arms for every frozen date. ACCURATE substitutes realized demand into clearing, SOURCE_WRONG retained the frozen forecast, UNDER_MATCHED and OVER_MATCHED set every realized nodal nonflexible-demand component to 0.95 and 1.05 times its realized value respectively. The two matched arms impose the same proportional error magnitude with opposite signs to test whether a result is specific to the forecast sample consisting mostly of underforecasts. Other VPP resources, inputs, and realized conditions remain fixed across the arms.

Adequacy is sized before any arm is solved and is identical across arms within a date. In each hour, the upward reserve becomes the larger of the inherited requirement and 1.10 times the maximum positive system forecast error across the four arms. Every bus receives a separately registered firm-import resource with capacity equal to the larger value of 150 MW and 1.25 times the largest local positive forecast error across all hours and arms. Its marginal cost is 1.50 times the date's highest inherited generator's marginal cost

and remains below the inherited value of lost load. The added import can increase in the hour up to its local capacity, while original generators and VPPs retain the inherited link between upward recourse and day-ahead reserve award. These rules buy enough operational capability to prevent the forecast contrast from being identified through involuntary shortage.

The experiment plans 30 dates by 4 arms, creating 120 paths. A contrast is only eligible if its forecast arm and AC-CURATE pass the numerical gate and both have an absolute load loss and reserve shortage less than 10^{-6} MWh, while failed paths are not converted to zero. The effect unit is the market date with all three VPPs retained together. All headline whole-date bootstrap intervals use 5,000 resamples and the same registered seed 3,409,606.

Three sample rules prevent comparisons of unlike dates. First, the scarcity adequacy comparison uses the 14 source dates observed in both regimes. Second, comparisons among SOURCE_WRONG, UNDER_MATCHED, and OVER_MATCHED use a separate 14-date set eligible for all arms. Third, arm-specific estimates use every eligible pair. The analysis also reports medians, ranges with one date left out, a circular three-date block bootstrap, failure categories, preprocessing covariates for completed versus missing source pairs, and missing outcome tipping points.

The 16-point contract-coverage grid combines the four balancing-quantity multipliers and the four called-quantity coverage multipliers. Adjustments are evaluated after dispatch, so that every grid cell has the same physical outcome and only differs in cash allocation. The central illustration is (1.0, 1.5), and the inherited-settlement benchmark is (0, 0). No cell is labeled a breach or observed penalty.

5.7 Outcome definitions and acceptance

Every accepted pair reports the VPP's modeled profit, settlement components, household payments, full system resource cost, load shedding, shortage, and physical status. The adequacy experiment additionally calculates and reports the proxy call, dispatch quantity mapped to the call, the two contract-coverage bases, all 16 of the adjusted profit cells, four channel contributions, and residuals. A path that fails its numerical or uniqueness gate is retained as a failure and excluded from pair identification, and no failed path is replaced with a favorable date.

Uncertainty intervals resample complete dates in whole blocks. For the unified scarcity experiment, they summarize the heterogeneity across 23 identified dates and do not repair any missing outcomes for the remaining 7 dates. For the zero-loss sensitivity, they summarize 17, 19, or 18 arm-specific eligible dates, 14 common-four-arm dates, or 14 common-regime dates. The numerical reconciliation differences are computed only on the 15 dates with duplicate certified records and never enter the sample.

5.8 Pair construction, missing arms, and estimands

The date is the treatment-effect unit. For each frozen date, the accurate and wrong arms share the same realized demand, network capacity, resources, costs and nonforecast inputs. The contrast isolates the consequence of using a source forecast rather than the oracle-accurate demand in the computational market. A date contributes to a paired

estimate only when both arms reach an accepted terminal state.

This rule defines a complete-case estimand for 23 identified dates without repairing missing arms or assuming that failed paths are exchangeable. The rest of the incomplete dates contribute to coverage and tipping-point diagnostics, but not the complete-case mean. The primary descriptive quantities are whole-date means and medians. For VPP outcomes, each date averages the three VPP effects, and bootstrap intervals then resample dates so that the three VPPs are not treated independently. The scarcity experiment and zero-loss sensitivity estimate different conditional quantities and are never pooled. Their percentage reduction is computed only on the 14 dates available to both regimes. Duplicate-record percentage differences are diagnostics of numerical consistency, not additional treatment effects.

After the primary objective is fixed, a strictly convex secondary problem selects a reproducible schedule and representative. A settlement effect is identified only when it is invariant across the primary set, and economic prices are only accepted when the solution satisfies primal and dual feasibility, stationarity, and duality. If prices or allocations remain non-unique, the path is excluded from the estimation or is reported unidentified.

The accepted paths must satisfy checks, such as generator, network, and cash closure checks. The core equality tolerance is 10^{-7} , and the physical zero-loss threshold is 10^{-6} MWh for both load-shedding and reserve shortage. These checks are applied before treatment effects are calculated.

6 Results

6.1 Mechanism map of loss, zero, interaction, and ambiguity regions

In Figure 2, an artificial mechanism surface is described where 0 means the difference is within USD 1 per day. Hatching marks an interval that crosses a sign or a boundary. Based on the phase map, the statement that more error always raises or lowers VPP profit is rejected. Demand error is a loss in 6 of nine cells and an economic zero in 3. Capacity error is a loss in 5, zero in 3, and ambiguous in 1.

The positive interaction is not a direct positive outcome. In several cells, both main effects are losses, but the interaction offsets part of their sum. The ambiguous cell near the actual capacity 0.09 and bias 0.025 is preserved as an interval for a relevance boundary. The four midpoint cells reproduce the frozen criteria, where all arms retain feasibility and cash closure. VPP2 remains economically neutral, while loss, zero, and positive-or-ambiguous interaction patterns persist.

6.2 High-fidelity dates results

On September 16th 2024, demand error lowered the combined VPP profit by USD 276.68/day. All three VPPs lost profit, and the largest decline occurred in VPP2. Of the combined effect, USD -221.84 is reserve-capacity revenue and USD -54.84 is the day-ahead energy revenue, while the real-time deviation and degradation are negligible. The error in forecast thus changed hourly schedules even though the net day-ahead position was essentially constant, so this result is influenced by reserve-headroom and timing.

On November 24th 2024, the combined loss was USD -324.16/day. Day-ahead energy contributed USD -147.53, re-

Economic Value of Forecasts across VPPs

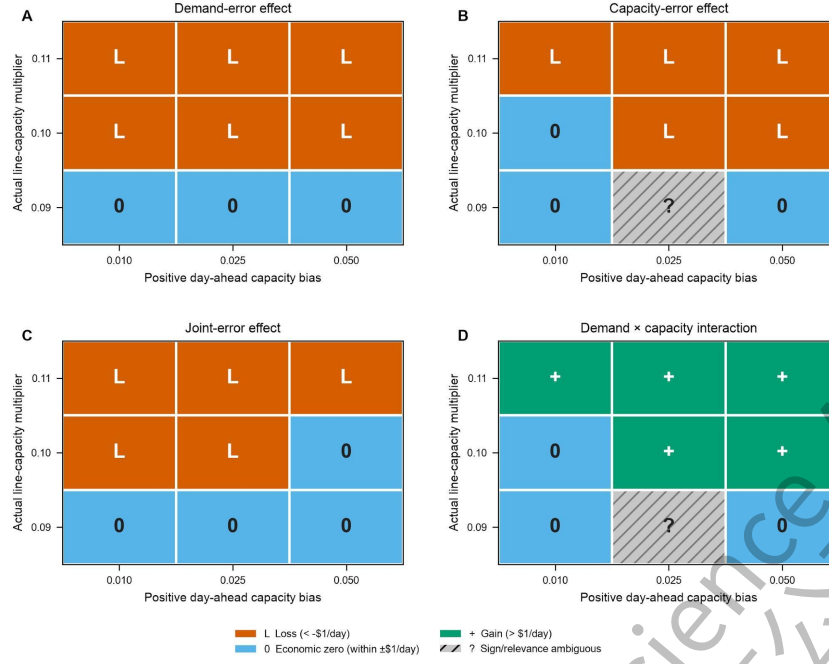


Figure 2: Two-bus mechanism phase map. Each panel classifies VPP1's wrong-minus-accurate effect over nine combinations of actual capacity and positive forecast bias.

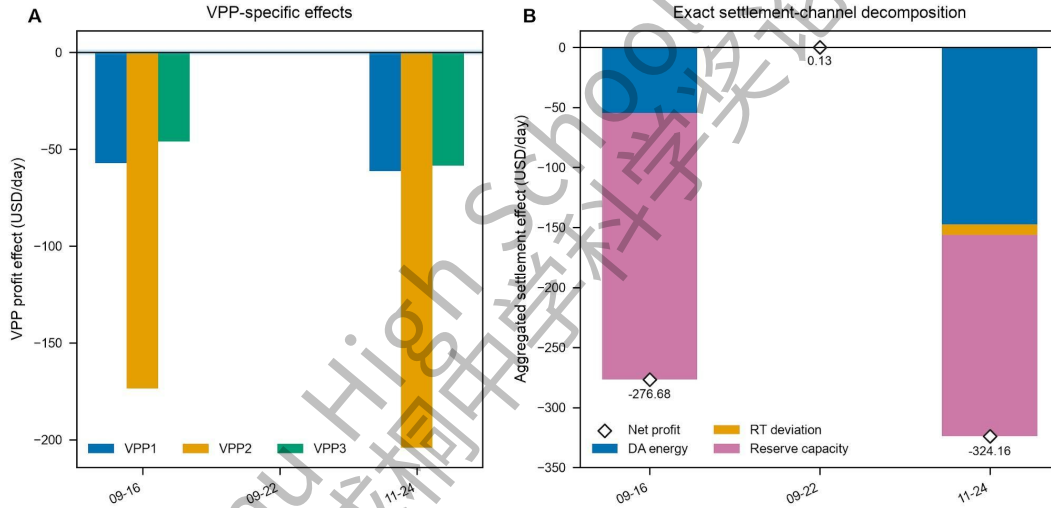


Figure 3: High-fidelity profit effects and settlement channels. Panel A reports VPP-specific wrong-minus-accurate profit effects, while Panel B sums the verified settlement components over all three VPPs. Diamonds show the total profit, and the September 22 result is visible as an economic zero.

serve capacity contributed USD -167.88, and the real-time deviation settlement contributed USD -8.95. VPP2 also had the largest private loss, with the maximum hourly schedule change reaching 1.541 MWh. This demonstrated that daily net energy can conceal different hourly schedule movements.

On September 22nd, the combined effect is only USD +0.13/day, and every VPP-specific effect was below its minimum relevant threshold. This is the first negative control. While the forecasting intervention is real, the market state does not amplify it to a settlement outcome.

6.3 Inactive slack network forecasting and unidentified binding effect

In the representative network diagnostic, both the accurate and wrong line capacities remained outside the binding region. The capacity-forecast error changed the system by only USD +0.70, household payment by USD -0.20, and any single VPP profit by no more than USD 0.14. This result shows that a day-ahead capacity forecast can have little value if it

does not change the day-ahead feasible or optimal face because the real-time transmission lines can still be congested.

The binding high-fidelity state supplies a different result, that multiple primary-optimal day-ahead allocations carry materially different VPP commitments and profits. Because the registered secondary allocation does not establish a unique optimum, the binding congestion-forecast magnitude is not identified. Thus, we do not report a zero nor preferred point estimate. The two-bus surface demonstrates mechanism possibility, but its artificial dollar magnitudes cannot fill the high-fidelity identification gap.

6.4 Thirty-date extension identifies a scarcity regime

In the unified experiment, the 30-date design produces 23 numerically certified resource pairs with seven incomplete dates. In the complete pairs, mean system resource cost increases by USD 7.959 million/day, with a whole-date bootstrap 95% interval of USD 6.661 to 9.313 million/day. Real time load-shedding increases by 774.54 MWh/day, with an

interval of 643.41 to 911.69 MWh/day. Mean VPP profit rises by USD 117.21 thousand per VPP-day, with an interval of USD 103.86 to 130.59 thousand. Aggregating the three VPPs, the average gain is USD 351.62 thousand/day. Household net payment also falls by USD 353.60 thousand/day, with an interval from USD -396.13 to -312.02 thousand/day.

The direction of all panels is unanimous in the 23 identified dates. All dates have a higher system cost and VPP profit, while having a lower household net payment with more load shedding in the wrong demand-forecast arm. This is the extended scarcity state previously mentioned in the mechanism, where the source underforecast changes the day-ahead system and the real-time system enters a costly shortage.

The cost decomposition shows why the dollar magnitude is significant. The mean load-loss penalty increases by USD 7.745 million/day, which is equal to a 97.32% total system cost increase, and the household shortfall component adds USD 255.86 thousand/day. Then, conventional generation cost decreases by USD 41.57 thousand/day and degradation cost decreases by USD 1.14 thousand/day, but these costs are relatively small compared to the load-loss penalties. Thus, the load-shedding cost makes up the majority of total system cost. Across the 69 reconciled observations, mean real-time deviation settlement contributes USD 116.77 thousand per day. Day-ahead energy and reserve-capacity changes contribute only USD 4.09 and 51.35 per day, while the reduction in degradation cost adds about USD 379.06. The wrong forecast therefore impacted VPP profit by changes in real-time deviation settlement. This thus closes the central mechanism chain from forecast error to changed commitment, real-time deviation, settlement gain, and reliability cost.

6.5 Adequacy removes shortage, substantially reducing the apparent profit effect

In the 120 simulation paths, all reached a terminal state, with 97 passes and 23 remaining in numerical or price-uniqueness failures. The 23 failures consist of 10 nonunique qualification results, 9 scientific path rejections, 1 nonunique singleton global-equality pivot, and 3 singleton global-equality recheck failures. On the 14 dates available for the scarcity and adequate resource specifications, mean VPP settlement surplus falls from USD 126,655.72 to USD 112.70 per VPP per day. This responds to a 99.911% reduction. This result can be shown in the following Figure 5.

6.6 Contractual cash allocation can reverse the residual VPP-account effect

Under the central cash-allocation illustration, the SOURCE_WRONG effect becomes USD -743.77 per VPP-day, with a median of -744.45 and a 95% interval from -865.56 to -625.51. The grid is monotone in the called-quantity multiplier, but increasing the balance-quantity multiplier does not always reduce the profit difference because both the inaccurate and accurate forecast cases receive call credits. For SOURCE_WRONG, a balancing-quantity multiplier of 0.5 with no called-quantity adjustment leaves a small positive mean of +37.50 and a multiplier of 1.0 yields -45.31. A called-quantity multiplier of 1.0 with no balancing-quantity adjustment produces a change of -345.32.

Because the fixed-dispatch overlay is linear, the adjusted profit 17 SOURCE_WRONG dates can be represented as:

$$\overline{\Delta\Pi}^{adj} = 120.317 - 165.629m_B - 465.637m_K. \quad (17)$$

where 120.317 is the mean gross profit across the 17 dates, 165.629 is the average sensitivity to m_B , and 465.637 is the average sensitivity to m_K . With $m_K = 0$, the mean reverses at $m_B = 0.726$, while with $m_B = 0$, it reverses at $m_K = 0.258$. These sign boundaries show that the residual positive wedge depends on whether the coverage bases receive an additional cash allocation wedge.

6.7 Four-channel decomposition

For SOURCE_WRONG under the central illustration, the mean price channel is USD +33.95 per day, the injection quantity is USD 31.98, and the reserve-capacity-income channel is USD +54.38. These channels sum up to USD +120.32 before negligible degradation and audit terms. The contract allocation channel is USD -864.08, with the adjusted profit USD -743.77. The 95% intervals are +18.83 to +48.95 for price, +24.34 to +40.23 for quantity, +37.11 to +72.37 for reserve income, and -997.74 to -737.21 for contractual allocation.

This result adds on to the mechanism because in the zero-loss state, no reliability loss is needed to explain the gross gain, and the VPP exposed to small changes in price and injections retains additional reserve-capacity income. Under the illustrative rules, the negative contractual-allocation channel is about 7.2 times the sum of the positive gross channels. This result shows how an additional layer would reassign cash.

6.8 Symmetric forecast directions reject a one-sided explanation

The four-arm comparison holds 14 dates fixed, where gross means are +111.63 for SOURCE_WRONG, +58.55 for the matched 5% underforecast, and +35.34 for matched 5% overforecast. Their medians are +90.61, +58.24, and +31.97, and the corresponding 95% intervals are +85.21 to +140.04, +39.85 to +77.55, and +21.99 to +49.08. This means that the small positive gross wedge is not unique to the underforecast source series.

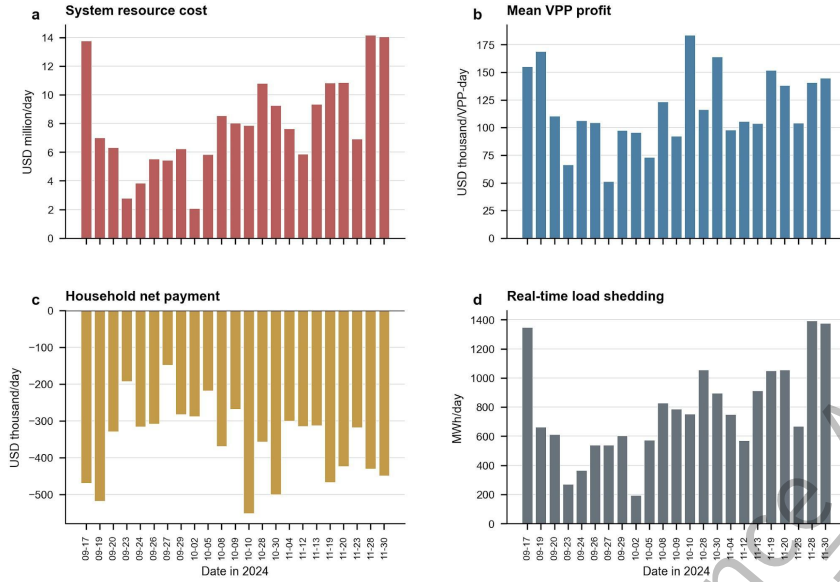
Under the central illustration in the experimentation of the 14 dates, the means become -759.33, -1,567.41, and -522.78, with intervals of -906.45 to -619.56, -1,700.68 to -1,443.70, and -591.39 to -463.37. The different magnitudes show dispatch adjustments and proxy calls in the coverage formula. Their common sign supports a conditional conclusion that with zero loss, the direction of VPP profit depends on the contract-coverage rule.

6.9 The unified result is stable to reconciliation

The 15 dates with duplicate certified records provide an internal numerical-consistency check before reconciliation. Mean relative differences are 0.072% for system resource cost, 0.855% for total VPP profit, 0.070% for load-loss cost, 0.070% for load shedding, and 0.843% for household payment. Maximum relative differences remain 0.191%, 2.202%, 0.307%, 0.307%, and 2.141% respectively.

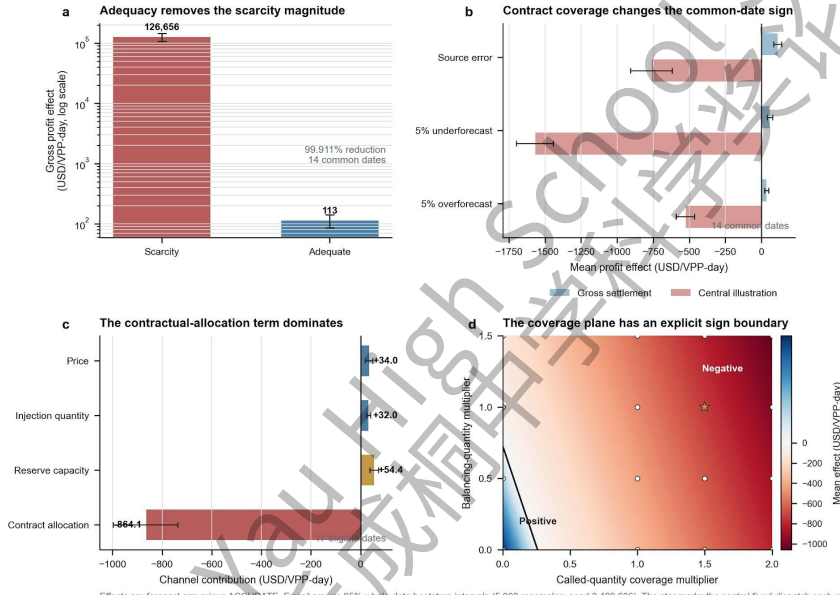
Deleting any one of the unified dates would leave mean VPP profit between USD 114.19 and 120.21 thousand per day, and mean system resource cost between USD 7.677 and

Economic Value of Forecasts across VPPs



Wrong-minus-accurate effects. Each of the 23 completed dates contributes one reconciled date-level observation; seven dates remain incomplete.

Figure 4: Joint outcomes in the unified 30-date experiment. The panels report the wrong-minus-accurate effects for the 23 completed dates, with each date contributing one observation. The panels show system resource cost, mean VPP profit, household net payment, and real-time load shedding in chronological order. Seven dates are incomplete and are not imputed



Effects are forecast arm minus ACCURATE. Error bars are 95% whole-date bootstrap intervals (5,000 resamples; seed 3,409,606). The star marks the central fixed-dispatch cash-allocation illustration (mB=1.0, mK=1.5); it is not a breach finding.

Figure 5: Adequacy and fixed-dispatch cash-allocation sensitivity. Panel A compares gross VPP-settlement surplus effects on the 14 dates available to both scarcity and adequate regimes. Panel B-D reports adequate regime gross effects, adjusted effects, channel decomposition, and the sample's sign boundary. The contractual adjustment is hypothetical.

8.226 million per day. To make the full-30 VPP-profit mean equal zero, the seven incomplete dates would need an average effect of USD -385.11 thousand per day, while to make the mean system cost equal zero, they would need to have an average of USD -26.149 million per day.

Table 4 puts all evidence layers together while retaining their different samples, preventing the different experiments from being read as one pooled estimate.

7 Robustness and Identification

Central findings of this paper depend on pairing, qualification, and a distinction between a missing outcome and an economic zero.

7.1 Duplicate-record reconciliation

We assign an observation value to each date. Fifteen dates have two certified numerical records, which are compared before their mean is taken. All duplicate records have the same signs for system cost, VPP profit, household payment, and load shedding. Mean relative differences remain below 1% for these headline quantities, and the maximum headline difference is 2.202% of the total VPP profit. This check supports the use of date values while preventing duplicate values from inflating the scientific sample. A stricter sensitivity restricted to the 10 dates where records retain their original PASS status preserves all four headline directions, showing that directional conclusions do not depend on the broader compatibility classifications.

Table 4: Core numerical results under the wrong-minus-accurate convention.

Evidence block	Identified dates	VPP profit effect (in USD/day)	System resource-cost effect (in USD/day)	Household-payment effect (in USD/day)	RT Load-shedding effect
High fidelity Sep. 16, 2024	1 development date	-276.68, three VPP total	+7.47	-5,393.53	Approximately zero
High fidelity Sep. 22, 2024	1 development date	+0.13, economic zero	+0.31	-10.97	Approximately zero
High-fidelity Nov. 24 2024	1 development date	-324.16, three VPP total	+22.35	-1,519.16	Approximately zero
Unified 30-date scarcity experiment	23 complete pairs	117.21 thousand [103.86, 130.59]	7.959 million	-353.60 thousand [-396.13, -312.02]	+774.54 MWh/day [643.41, 911.69]
Zero-loss source error with inherited settlement	17 eligible pairs	+120.32 [95.95, 146.61]	Not the primary estimand	Not the primary estimand	Absolute value at most 3.08×10^{-8} MWh
Zero-loss source error central illustration	17 eligible pairs	-743.77 [-865.56, -625.51]	Physical dispatch unchanged	Physical dispatch unchanged	Absolute value at most 3.08×10^{-8} MWh
Zero-loss common-four-arm gross	14 dates	Source +111.63, Under +58.55, Over + 35.34	Not the primary estimand	Not the primary estimand	Below 10^{-6} MWh gate
Zero-loss-common-four arm central	14 dates	Source -759.33, Under -1,567.41, Over -522.78	Physical dispatch unchanged	Physical dispatch unchanged	Below 10^{-6} MWh gate

7.2 Missing-outcome sensitivity

In the two primary scarcity directions, VPP profit is between USD 114.19 and 120.21 thousand per day, while system resource cost is between USD 7.677 and 8.226 million per day. The seven incomplete dates are not inputted, and their average effects would need to be USD -385.11 thousand per day and USD -26.149 million per day to bring the full system cost to zero.

7.3 Two negative controls

Two negative controls constrain the interpretation. First, September 22nd of 2024 is an economic zero date, where the aggregate profit change is only USD 0.13/day, and each VPP remains below the threshold of minimum relevance effect. The mechanism chain is active, but economically trivial. Second, the high-fidelity transmission capacity diagnostic holds both day-ahead forecasts outside the binding region. The maximum absolute VPP-profit effect is 0.14 USD, and the system-cost change is only about +0.70 USD, and the household-payment change is about -0.20 USD, showing that real time congestion alone does not establish the value of day-ahead forecasts.

7.4 Binding congestion and structural non-identification

The binding high-fidelity congestion experiment admits multiple optimal day-ahead allocations with different VPP commitments and settlements. Choosing the most favorable or least favorable allocation or an arbitrary solver return would create a point estimate. Therefore, we report the effect unidentified under the current settlement rule. This boundary is substantive because a unique optimum would not need to determine a unique distribution of participation payments.

7.5 Missingness and numerical failures

In the unified scarcity experiment, a date is complete when at least one certified wrong arm record and accurate arm record exist, and the experiment identifies 23 dates complete, 7 incomplete. The complete case averages are selected toward

states easier to certify because numerical difficulty is more common near binding constraints and scarcity. The reconciliation rule improves coverage without claiming that the seven missing dates are exchangeable with the observed dates.

The zero-loss registry is more diagnostic as all four arms were in one frozen protocol. Its 23 failures are classified as 10 nonunique qualification results, 9 scientific-path rejections, 1 nonunique singleton equality pivot, and 3 singleton equality recheck failures. For SOURCE_WRONG pairs, completed dates have a 558.93 MWh higher mean realized load and 16.93 MWh higher mean reserve requirement than missing dates, and the standardized mean differences are 0.408 and 0.171. Absolute source-forecast bias is 265.28 MWh higher in completed dates, with a standardized difference of 0.606. Two bounded checks qualify the complete-case result. Leaving out any one of the SOURCE_WRONG dates keeps the gross mean positive, between +114.80 and +125.70, and the central mean negative, between -768.32 and -710.41. A circular three-date block bootstrap gives intervals of +87.22 to +158.53 for gross profit and -894.95 to -584.03 for the central illustration. For all the gross mean to equal zero, the 13 missing dates would need an average effect of USD -157.34 per day, and for the central mean to equal zero, they would need +972.62.

7.6 Zero-loss gate and common adequacy

For the zero-loss conclusion, contrast date is included only when both arms pass and load loss and reserve shortage are below 10^{-6} MWh. The observed maxima are orders of magnitude below this threshold, so eligibility is not created by rounding a shortage to zero. Adequacy resources are sized jointly across all arms before solution, preventing any arm from receiving favorable ex-post capacity. Imports are installed at every bus, removing the possibility that aggregate capacity cannot reach the local imbalance due to transmission. The added adequacy package creates a counterfactual where a lack of real-time capability cannot explain the settlement effect. The remaining gross-profit difference results

are thus conditional on ample balancing access.

7.7 Contract-coverage grid and decomposition checks.

The contract-coverage grid holds dispatch fixed and uses the same replacement-cost benchmark. The boundary is represented in Equation (17). A called-quantity multiplier of 1.0 with no balancing quantity adjustment can produce a negative SOURCE_WRONG interval, while a balancing-quantity multiplier of 1.0 alone makes the mean negative, but the interval still includes zero. The four channels reconcile to an adjusted profit with a maximum residual of $\text{USD } 2.73 \times 10^{-12}$

The entire experiment is conditional on inputs and the common bootstrap seed, with cases passing through Equations (9) to (11). The added case requires a departure from the day-ahead plan, and assuming a full compliance with instructions, records zero participant breach. Repeating the summary produces the same identity.

7.8 Established results of robustness checks

Together, the robustness checks establish five points. First, reconciliation prevents duplicate counting while preserving the identified scarcity-state directions. Second, these directions survive every leave-one-date-out deletion. Third, economic zero and structural nonidentification are different from numerical failure. Fourth, on 14 common-regime dates, the scarcity-profit magnitude collapses as the adequate experiment eliminates shortage. Finally, the illustrated contract-coverage rates reverse the remaining gross gain and the channel closes numerically.

8 Analysis and Distributional Consequences

8.1 VPP gains are private settlement gains

In the 30-date scarcity regime, the wrong forecast arm raises modeled VPP settlement surplus primarily because the real-time deviation account becomes more valuable. The result shows that under the modeled gross settlement, a resource delivering a deviation in a scarcity state receives high marginal revenue. The private account rewards the service at the margin, while the whole system account records the reliability condition that made the service scarce.

The three VPPs do not receive identical gains. In the unified complete pairs, the mean profits are USD 114.86 thousand per day for VPP1, USD 182.54 thousand per day for VPP2, and USD 54.22 thousand per day for VPP3.

The zero-loss experiment also adds a distinction between inherited and contract-adjusted returns. Gross profit still rises slightly, but adjusted profit falls under the central cash-allocation rule.

Looking at household payment, the net payment falls on all 23 identified dates. This direction might seem favorable, but because the payment is a modeled transfer without considering income, utility, interruption damage, or compensating variation, the model cannot directly prove that household welfare improved.

The dominant effect is load-loss, which strengthens the conclusion that wrong forecasts are socially costly to scarcity dates, as it leads to more physical burden. Also, the reduction in conventional generation cost does not contradict the increase in total resource cost. If the system fails to serve load, it may produce less electricity and thus spend less on

generation. Reporting generation cost alone would label the shortage as saving. Therefore, full resource accounting must retain the penalties.

8.2 Joint interpretation across two regimes

The 30-date regime shows the following pattern:

$$\Delta \Pi^{VPP} > 0, \quad \Delta H < 0, \quad \Delta C^{sys} > 0, \quad \Delta Q^{shed} > 0. \quad (18)$$

This equation describes who receives money, who pays, and what physically happens in the market, showing modeled VPP profit increases, recorded household net payment falls, system resource cost rises, and load shedding rises. Preserving all four entries is necessary for a defensible policy conclusion.

The zero-loss sensitivity defines a second regime as shown:

$$\Delta Q^{shed} \approx 0, \quad \Delta Q^{res, short} \approx 0, \quad \Delta \Pi^{gross} > 0, \quad \Delta \Pi^{adj} < 0. \quad (19)$$

This equation isolates an institutional wedge, showing load shedding and reserve shortage approximately zero, gross profit slightly positive, and adjusted profit negative. Ample physical recourse removes the emergency of reliability, while the gross settlement still produces a smaller positive difference. Adding the contract-coverage terms, the pattern is reversed under the illustrated rates. The two regimes thus show that physical feasibility changes whether shortage occurs while cash allocation changes who retains the settlement from balancing quantities and the two regimes should be diagnosed separately.

9 Discussion, Policy Implications, and Limitations

The policy question is not whether forecasting should improve, but where the forecast improvement alters operational decisions, how recourse limits the cost of the errors, and how settlement rules assign cash. This section translates computational findings into monitoring and design principles that could be used to observe trends in larger markets or alternative regimes.

9.1 Forecast quality should be valued by state

The results suggest replacing a uniform measure of value per unit of forecast accuracy with a state-contingent measure. Useful predictors include the expected magnitude of error, proximity to a commitment boundary, reserve tightness, deliverable flexibility, the price spread, and consequences of shortage. Forecast improvement is the most valuable when an error is likely to change a binding decision or prevent the entry into a reliability-penalty region. The prioritization rule is

$$\begin{aligned} \text{Forecast priority}(s) = & p_{cross}(s) \left[\text{deliverable adjustment}(s) \right. \\ & \times \text{marginal avoided cost}(s) + \text{avoided reliability loss}(s) \left. \right] \\ & - \text{information cost}(s). \end{aligned} \quad (20)$$

9.2 Settlement design must distinguish dispatch and VPP fault

The 30-date scarcity experiment showed a difference between VPP settlement returns and social resource costs. A VPP can

earn more when a wrong forecast creates real-time scarcity, even though its injection is valuable. The zero-loss grid also supplies a sensitivity showing how an additional debit would change the VPP account. A balancing-quantity multiplier of 1.0 reverses the mean of the erroneous arm. A defensible rule needs to separate instruction, reserve call logic, settled credits, and the payer and receiver. A defensible rule would need to separate instruction and delivery records, reserve-call logic, credits that are already settled, and a named payer and receiver.

Possible policy instruments include reserve rules based on performance, and any instrument should be tested about whether it preserves VPP participation, maintains physical reliability, and avoids double recovery.

9.3 A practical market-monitoring rule

Operators can assess four layers when evaluating the market value of a forecast.

1. Plan response: Did the forecast change energy or reserve commitments?
2. Constraint response: Did a line, reserve, ramp, or short-age constraint change status?
3. Settlement response: Did price and deviation exposure change exceed an economic threshold?
4. Incidence response: Did VPP profit, household payment, system cost and reliability move together?

If neither commitments or status changes, improved statistical accuracy may have little market value. Where a market response occurs, its full incidence vector should be reported.

9.4 Evidence interpretation and limitations

The 15 duplicate-record dates support the numerical consistency of the principal directions. Records are reconciled within dates and never are counted as separate observations. Agreement strengthens confidence in the numerical implementation but does not add any new variation. External validity still requires new frozen date blocks, additional network topologies, and alternative designs. This study identifies how forecast errors affect VPP settlements under scarcity and adequate operating conditions. The scarcity experiment identifies 23 complete pairs from 30 frozen dates, while the zero-loss experiment produces 17-19 eligible pairs depending on the contrast. While the main directions are stable, reported means describe an experimental sample, and a population-wide effect cannot be completely proved. The accurate treatment uses realized demand and should be seen as an upper-information benchmark, not a directly deployable forecasting method. Similarly, the two-bus and five-bus models identify mechanisms and regime boundaries, but can not directly represent situations in real markets. The model measures VPP profit, costs, shortage, and household payments, but household payments cannot completely measure welfare. Income differences, interruption costs, and behavioral responses need further evidence to prove.

Furthermore, the numerical magnitudes depend on the five-bus network, three modeled VPP portfolios, and two-settlement rules. Larger networks might produce different specific magnitudes, but the central mechanism applies

generally. The large profit effect should be interpreted as an incentive mismatch in extreme scarcity. When sufficient balancing capability eliminates loss, the effect becomes much smaller and its sign depends on the contract. Future work can test these mechanisms in larger systems, broader date samples, and observed contract data.

10 Conclusion

In conclusion, this study shows that the heterogeneous economic value of forecast accuracy in virtual power plants depends on an operational and settlement response. For an error to effect settlement, a forecast error must change an economically relevant event, such as a schedule, resource, or constraint. The settlement rule then allocates the financial consequences of that adjustment. These linked responses explain gains, losses, economic zeros, and non-unique settlement outcomes.

Across complete modeled date pairs, forecast errors increase the average modeled VPP profit by USD 117.21 thousand per VPP per day, while increasing system cost by USD 7.959 million per day and load shedding cost by 774.54 MWh/day. All completed dates share these directions, which survive reconciliation and leave-one-date-out checks. In the extreme-scarcity states, private VPP gains coexist with deteriorating system reliability.

The zero-loss comparison provides the paper's most significant extension. After sufficient reserve, imports, and ramping eliminate shortage, the VPP's profit effect falls by 99.911%, from USD 126,655.72 to USD 112.70 per VPP per day. Small gross settlement differences remain across source, underforecast, and overforecast treatments, but its sign changes when balancing responsibility is considered. Under the contractual-coverage specification, the source-error effect becomes USD -743.77 per VPP per day. Price, injection, reserve, and the contract allocation decomposition help explain this reversal.

The main contribution is thus a distinction among forecast quality, physical balancing adequacy, and settlement responsibility. The common-date comparison links large private gain to a scarcity regime, and ample recourse leaves a smaller difference, whose allocation depends on the hypothetical contract. This finding motivates joint attention to forecast information, balancing capability, and transparent participant obligations. While the conclusions are based on the modeled network and identified dates, the pattern and analytical framework can be tested in and applied to larger systems and alternative market designs.

References

- [1] Morstyn, T., Farrell, N., Darby, S. J. et al. Using peer-to-peer energy-trading platforms to incentivize prosumers to form federated power plants. *Nature Energy* 3, 94–101 (2018). <https://doi.org/10.1038/s41560-017-0075-y>
- [2] Naval, N. & Yusta, J. M. Virtual power plant models and electricity markets: A review. *Renewable and Sustainable Energy Reviews* 149, 111393 (2021). <https://doi.org/10.1016/j.rser.2021.111393>
- [3] Muñoz, M. Á., Morales, J. M. & Pineda, S. Feature-driven improvement of renewable energy forecasting and trading. *IEEE Transactions on Power Systems* 35, 3753–3763 (2020). <https://doi.org/10.1109/TPWRS.2020.2975246>
- [4] Klyve, Ø. S., Nygård, M. M., Riise, H. N., Fagerström, J. & Marstein, E. S. The value of forecasts for PV power plants operating in Scandinavian energy markets. *Solar Energy* 255, 208–221 (2023). <https://doi.org/10.1016/j.solener.2023.03.044>
- [5] Wang, B., Guo, Q., Yang, T., Xu, L. & Sun, H. Data valuation for decision-making with uncertainty in energy transactions: A case of the two-settlement market system. *Applied Energy* 288, 116643 (2021). <https://doi.org/10.1016/j.apenergy.2021.116643>
- [6] Zavala, V. M., Kim, K., Anitescu, M. & Birge, J. A stochastic electricity market clearing formulation with consistent pricing properties. *Operations Research* 65, 557–576 (2017). <https://doi.org/10.1287/opre.2016.1576>
- [7] Morales, J. M., Conejo, A. J., Madsen, H., Pinson, P. & Zugno, M. *Integrating Renewables in Electricity Markets: Operational Problems*. Springer (2014). <https://doi.org/10.1007/978-1-4614-9411-9>
- [8] Conejo, A. J., Carrión, M. & Morales, J. M. *Decision Making Under Uncertainty in Electricity Markets*. Springer (2010). <https://doi.org/10.1007/978-1-4419-7421-1>
- [9] Nguyen, H. T., Le, L. B. & Wang, Z. A bidding strategy for virtual power plants with intraday demand response exchange market using stochastic programming. *IEEE Transactions on Industry Applications* 54, 3044–3055 (2018). <https://doi.org/10.1109/TIA.2018.2828379>
- [10] Shafiekhani, M., Badri, A., Shafie-khah, M. & Catalão, J. P. S. Strategic bidding of virtual power plant in energy markets: A bi-level multi-objective approach. *International Journal of Electrical Power & Energy Systems* 113, 208–219 (2019). <https://doi.org/10.1016/j.ijepes.2019.05.023>
- [11] Naughton, J., Wang, H., Riaz, S., Cantoni, M. & Mancarella, P. Optimization of multi-energy virtual power plants for providing multiple market and local network services. *Electric Power Systems Research* 189, 106775 (2020). <https://doi.org/10.1016/j.epsr.2020.106775>
- [12] Naughton, J., Wang, H., Cantoni, M. & Mancarella, P. Co-optimizing virtual power plant services under uncertainty: A robust scheduling and receding horizon dispatch approach. *IEEE Transactions on Power Systems* 36, 3960–3972 (2021). <https://doi.org/10.1109/TPWRS.2021.3062582>
- [13] Caramanis, M. C., Ntakou, E., Hogan, W. W., Chakraborty, A. & Schoene, J. Co-optimization of power and reserves in dynamic transmission and distribution power markets. *Proceedings of the IEEE* 104, 807–836 (2016). <https://doi.org/10.1109/JPROC.2016.2520758>
- [14] Bohn, R. E., Caramanis, M. C. & Schweppe, F. C. Optimal pricing in electrical networks over space and time. *RAND Journal of Economics* 15, 360–376 (1984). <https://doi.org/10.2307/2555444>
- [15] Borenstein, S., Bushnell, J. B. & Wolak, F. A. Measuring market inefficiencies in California's restructured wholesale electricity market. *American Economic Review* 92, 1376–1405 (2002). <https://doi.org/10.1257/000282802762024557>
- [16] Wilson, R. Supply function equilibrium in a constrained transmission system. *Operations Research* 56, 369–382 (2008). <https://doi.org/10.1287/opre.1070.0421>
- [17] Kirschen, D. S. & Strbac, G. *Fundamentals of Power System Economics*, 2nd ed. Wiley (2018). ISBN 978-1-119-21325-3.
- [18] Power Grid Lib. PGLib-OPF v23.07. <https://github.com/power-grid-lib/pglib-opf/releases/tag/v23.07>
- [19] California ISO. Open Access Same-Time Information System. <https://www.caiso.com/systems-applications/portals-applications/open-access-same-time-information-system-0asis>
- [20] Wang, Z., Lu, B., Wang, B. et al. Incentive-based emergency demand response effectively reduces peak load during heatwave without harm to vulnerable groups. *Nature Communications* 14, 6202 (2023). <https://doi.org/10.1038/s41467-023-41970-8>
- [21] Calver, P. & Simcock, N. Demand response and energy justice: A critical overview of ethical risks and opportunities. *Energy Policy* 151, 112198 (2021). <https://doi.org/10.1016/j.enpol.2021.112198>
- [22] de Souza Dutra, M. D. & Alguacil, N. Fairness of prosumers' incentives in residential demand response. *International Journal of Electrical Power & Energy Systems* 148, 109015 (2023). <https://doi.org/10.1016/j.ijepes.2023.109015>
- [23] Huang, B., Meijssen, A. G., Annema, J. A. & Lukszo, Z. Are electric vehicle drivers willing to participate in vehicle-to-grid contracts? *Energy Policy* 156, 112410 (2021). <https://doi.org/10.1016/j.enpol.2021.112410>
- [24] Sarathkumar, T. V., Goswami, A. K., Khan, B. et al. Forecasting of virtual power plant generating and energy arbitrage economics in the electricity market using machine learning approach. *Scientific Reports* 15, 3812 (2025). <https://doi.org/10.1038/s41598-025-87697-y>
- [25] Zheng, D., Tong, D., Davis, S. J. et al. Climate change impacts on the extreme power shortage events of wind-solar supply systems worldwide during 1980–2022. *Nature Communications* 15, 5225 (2024). <https://doi.org/10.1038/s41467-024-48966-y>
- [26] Terrén-Serrano, G., Deshmukh, R. & Martínez-Ramón, M. Probabilistic day-ahead forecasting of system-level renewable energy and electricity demand. *Nature Communications* 17, 3307 (2026). <https://doi.org/10.1038/s41467-026-69015-w>

Acknowledgements

My research began with an interest in Virtual Power Plants and electricity markets, eventually becoming an academic endeavor spanning several months, testing my skills and resilience. My initial ideas for this project stemmed from a curiosity about how virtual power plants operate, specifically in how day-ahead forecasts influence the actions and results of virtual power plants. I was motivated by the complexity observed in previous research about virtual power plants, about when forecasts are used to shape VPP decisions, and what are heterogeneous economic impacts of such forecasts in the electricity market.

Researching and reading current sources, I discovered that the existing literature on VPPs that discussed VPP dispatch of energy sources and economic values of forecasting tended to rely on statistical measures of forecast quality, but tended to overlook a complexity of different electricity market mechanisms that influence the profits and costs of different participants. From this gap arose my central research ambition, to show that higher forecast accuracy of electricity demand or network conditions do not necessarily produce a higher profit in a VPP by creating an economic model and experimenting with different possible mechanisms of forecast impact.

I would like to express my deepest gratitude to Professor Lingshu Zhong for his selflessness and supportive guidance throughout my research. He devoted significant amounts of time and effort and guided me, encouraged me when I faced challenges, and consistently provided meaningful directions when I was stuck with a problem. In the process of research, I was responsible for the main research activities, including proposing and refining the topic, searching for and organizing literature, obtaining publicly available data, conducting data cleaning, and writing the main research paper. In the research process, I also was mainly responsible for designing the research framework for the energy market and revenue accounting, designing the main experiments modeling forecast errors, as well as interpreting the evidence. Professor Zhong supported me significantly, helping me review methodological logic and evaluate my experimental design, supervising the rigor of the scientific results, and providing revision suggestions when I faced obstacles in design and experimentation.

In the research process, I encountered many difficulties. The first major difficulty I encountered was having inconsistencies in my data sources. Different datasets involved different units and definitions. Professor Zhong guided me in aligning observations by hour and conducting checks. Another major difficulty was market optimization. Some cases had convergence problems and incomplete solutions. Facing this difficulty, Professor Zhong encouraged me to retain failure records instead of removing unsuccessful runs, introducing new checks for feasibility.

Artificial intelligence was used to assist with locating code, explaining and debugging errors, drafting Python code, checking data-processing procedures and mathematical expressions, organizing literary sources, and checking for errors in my paper. The model is the OpenAI GPT-5.6 Sol. Artificial intelligence was not used to fabricate or supplement experimental data or observations, choose favorable results, or determine conclusions. Numerical values in the research are obtained from computational outputs, and the code, results, figures are thoroughly reviewed.

I sincerely thank Professor Zhong for guidance in my research, modeling what research truly means with persistence. This project has given me more than just experimental results. I learned what rigorous research, when confronting obstacles and uncertainty, is like. In the future, I would carry forward this experience that taught me how research, pursued with persistence and integrity, can answer impactful questions we observe.