The Economics of Negative Price Phenomenon in Renewable-Integrated Electricity Markets

Nima Rafizadeh[†]

Abstract

Understanding how renewable energy integration affects electricity market efficiency and price formation is an important challenge in energy economics and environmental policy. Negative electricity prices, where generators pay to produce power, now occur with increasing frequency across wholesale markets, yet their economic drivers require better understanding. This paper addresses this gap by developing a theoretical framework linking generator behavior to market outcomes, then testing it empirically using over 14 million observations from New York's wholesale markets (2010-2022). The theoretical analysis demonstrates that negative prices can achieve welfare-maximizing allocations under operational constraints and production subsidies. The empirical analysis, using binary response and count data models with high-frequency data, identifies a clear hierarchy of drivers: renewable energy integration emerges as primary, with solar energy reducing negative price occurrences while wind energy increases them. Weather conditions rank second in importance, while grid constraints show limited influence, contrary to policy focus on transmission expansion. These findings can inform policy discussions by suggesting that rather than suppressing negative prices through regulatory constraints, policymakers should preserve these efficient price signals while prioritizing technology-specific renewable policies and weather-responsive mechanisms over transmission expansion to enhance investment signals and market stability.

Keywords: Negative Electricity Prices, Renewable Energy Integration, Wholesale Electricity Markets, Price Formation, Market Design
JEL Classification Codes: Q41, Q42, D47

[†]Department of Resource Economics, University of Massachusetts Amherst, USA. Email: nrafizadeh@umass.edu

1. Introduction

The global electricity sector is experiencing a transformation driven by renewable energy deployment. Renewable capacity increased from approximately 1,566 GW in 2010 to 4,448 GW by 2024, with renewables representing 92.5% of new capacity additions in 2024 (IRENA, 2025). This structural shift has generated an economic phenomenon that creates challenges for market design: wholesale electricity prices that turn negative, where generators effectively pay to produce power. Once rare, negative electricity prices now occur with increasing frequency across wholesale systems, challenging conventional market theory and raising questions about optimal resource allocation during the renewable energy transition.

This trend has accelerated worldwide. In 2024, Europe experienced 7,841 hours of negative prices in just eight months—a 160% increase from 2023—while California's negative pricing frequency increased from 4% to 15% of all operating hours (3Degrees, 2024). Finland recorded over 650 hours of negative prices in 2024 compared to just 11 hours in 2020, while Spain witnessed 769 hours within its first year of allowing negative pricing (gridX, 2024). These developments represent more than technical market anomalies; they indicate a restructuring of electricity market economics that requires theoretical and empirical analysis.

Negative electricity prices occur when markets signal that additional electricity generation imposes net costs on the system, creating incentives for producers to pay rather than receive compensation for their output (Seel et al., 2021). This price inversion now characterizes portions of trading intervals across global wholesale markets, with some regions experiencing negative pricing in over 25% of operating hours during peak renewable generation periods. The growing frequency demonstrates a relationship between clean energy integration and market price dynamics that challenges traditional assumptions about electricity market operation.

The economic consequences extend beyond immediate pricing effects. Recent empirical analysis by Newbery (2023) demonstrates that marginal curtailment rates for wind will be at least three times average curtailment rates, with the last megawatt of wind capacity curtailed 3+ times more hours than the average. When curtailment becomes excessive, this can distort investment signals for both renewable and flexible generation resources. Negative prices also affect operational decisions by market participants, with Ketterer (2014) documenting impacts on grid stability and system coordination during negative pricing events. Price volatility increases during negative pricing periods, complicating risk management and financial planning for electricity market participants (Woo et al., 2011). Regional electricity trade patterns shift in

response to negative price differentials, modifying power flows and congestion patterns across interconnected systems (Genoese et al., 2010).

These market dynamics have policy implications, particularly with ongoing energy transition investments. Global investment in energy transition technologies reached $2.1 trillion in 2024, yet this represents only 37% of the $5.6 trillion required annually through 2030 to achieve net-zero pathways (BloombergNEF, 2025). Governments worldwide have committed resources to clean energy deployment, with the United States mobilizing over $11 billion through the Infrastructure Investment and Jobs Act and Inflation Reduction Act for grid modernization and renewable integration. The European Union's electricity market design reform, agreed in December 2023, represents a policy response to negative pricing challenges (European Commission, 2023). This creates a policy challenge: while negative prices can achieve welfare-maximizing allocations under operational constraints and renewable energy characteristics, excessive frequency may undermine investment signals and market stability. Policymakers should therefore determine optimal frequency and design policies that preserve efficiency benefits while managing problematic drivers (Cramton, 2004, Farrell, 2021).

Current research provides fragmented insights into negative pricing mechanisms and lacks integrated analysis of their relative importance and policy implications. Existing studies examine renewable energy effects in isolation (Nicolosi, 2010, Genoese et al., 2010), analyze weather impacts separately (Valor et al., 2001), or focus on transmission constraints (Joskow, 2006) without establishing their comparative importance. Recent advances in theoretical modeling, including multi-agent reinforcement learning frameworks that can simulate markets under 100% renewable penetration (Zhu et al., 2025), demonstrate that traditional merit-order concepts require revision when zero-marginal-cost resources dominate supply. However, no previous analysis integrates these theoretical insights with empirical evidence to establish the relative contributions of different drivers to negative price formation. This fragmentation creates gaps in policy guidance when effective action is needed.

Recent industry experience supports these policy challenges. Industry analysis by smartEn and LCP Delta (2024) demonstrates that flexibility integration through demand-side participation in wholesale markets could reduce energy costs while supporting renewable integration, yet debate continues over whether negative prices represent market failures requiring intervention or efficient signals that should be preserved. California's experience illustrates the complexity: despite storage capacity tripling from 4 GW to 11 GW between December 2022 and

June 2024, the state still curtailed 3.4 million MWh in 2024—a 29% increase from 2023 (Utility Dive, 2025). This persistence of curtailment alongside large storage deployment suggests that understanding the economic drivers of negative pricing has become relevant for designing effective renewable integration policies.

This study addresses two core research questions. First, what are the primary economic drivers of negative electricity prices in renewable-integrated markets? Second, what is the relative importance of these drivers in negative price formation? To answer these questions, this paper develops a formal theoretical framework that links generator-level optimization under operational constraints and production subsidies to market-level negative pricing outcomes, deriving formal conditions determining when each mechanism dominates price formation and generating empirical predictions about the relative importance of different drivers that emerge from microeconomic optimization principles. These theoretical predictions are then tested empirically by examining renewable energy integration, weather conditions, and transmission constraints as candidate drivers, using over 14 million observations from New York's wholesale electricity markets spanning 2010 to 2022, exploiting high-frequency data collected at 5-minute intervals across all market zones to establish a hierarchy of negative pricing drivers through econometric analysis.

The empirical approach addresses key identification challenges that have limited previous research. High-frequency data collected at 5-minute intervals across all market zones are exploited to capture rapid price dynamics and intra-hour variations invisible in traditional hourly analysis. Fixed effects specifications control for unobserved heterogeneity across zones and time periods, while weather-based proxies for renewable generation potential provide exogenous variation needed for causal identification, avoiding simultaneity bias that would arise from using actual generation data. Both the probability and frequency of negative price events are analyzed using econometric techniques designed for rare event analysis.

The results establish an empirical hierarchy of effects with policy implications that challenge conventional assumptions about negative pricing drivers. Renewable energy integration emerges as the primary driver, but with different effects across technologies that support technology-specific rather than technology-neutral policy approaches. Solar generation reduces negative price occurrences by aligning production with daily demand patterns, while wind generation increases negative pricing likelihood through production subsidy structures and temporal demand mismatches. Weather conditions rank second in explanatory importance,

with temperature and humidity effects showing opposing impacts on price formation, while transmission constraints demonstrate limited influence despite their prominence in policy discussions.

These findings offer implications for policy design during the renewable energy transition. They suggest that preserving negative price signals while addressing their underlying drivers offers a more efficient approach than suppressing these prices through regulatory constraints. The results indicate that effective negative price management may benefit from technology-specific policies recognizing distinct characteristics of different renewable sources, weather-responsive market mechanisms that adapt to meteorological conditions affecting renewable output, and demand-side interventions that may be more effective than supply-side approaches for managing negative pricing. They reveal that transmission expansion, while relevant for renewable access and system reliability, may not reduce negative pricing frequency, suggesting infrastructure investments should be justified on broader system benefits rather than negative price reduction.

This paper makes several contributions to energy economics and environmental policy. First, a theoretical framework is developed that formally connects generator-level optimization under operational constraints and production subsidies to market-level negative pricing outcomes, suggesting that negative prices can achieve efficient allocations in renewable-integrated markets. Second, empirical predictions from this framework are tested by establishing a hierarchy of negative pricing drivers, showing that renewable energy integration is the primary factor, weather conditions are secondary, and grid constraints have limited impact. Third, important heterogeneity within this hierarchy is revealed by identifying differential effects across renewable technologies, with solar generation reducing and wind generation increasing negative price occurrences. Fourth, these empirical findings inform evidence-based policy guidance, suggesting that effective negative price management should prioritize technology-specific renewable integration policies and weather-responsive mechanisms rather than focusing primarily on transmission expansion.

The remainder of this paper proceeds as follows. Section 2 develops the theoretical framework and derives empirical predictions. Section 3 presents the empirical methodology. Section 4 describes the data and provides descriptive evidence. Section 5 presents the estimation results. Section 6 discusses policy implications. Section 7 concludes.

2. Economic Theory and Market Context

This section establishes the theoretical and institutional foundations for analyzing negative electricity prices in renewable-integrated markets. The discussion begins by describing New York's electricity market structure and its renewable energy transformation. A formal theoretical model is then developed characterizing the economic mechanisms underlying negative price formation. Finally, theory is bridged to empirical implementation and empirical predictions about the key drivers of this phenomenon are derived.

2.1. Market Structure: New York's Renewable-Integrated Electricity System

New York State represents one of the earliest adopters of market deregulation among competitive electricity markets in North America and a leader in renewable energy integration. Its diverse energy landscape and ambitious clean energy goals make it a suitable case study for examining the evolving economic dynamics of renewable-integrated wholesale electricity markets. Since December 1, 1999, the New York Independent System Operator (NYISO) has managed the state's wholesale electricity markets (NYISO, 2011). The state is segmented into 11 interconnected load zones, as illustrated in Figure 1, each operating within a two-settlement system.

Figure 1: NYISO Control Area Load Zones

Source: https://www.eia.gov/electricity/gridmonitor/about

The NYISO administers both day-ahead and real-time markets through a well-developed pricing mechanism. The day-ahead market accepts hourly bids and schedules, incorporating

reliability commitments to optimize production costs. This market establishes financially binding schedules for energy and ancillary services one day before the operating day. The real-time market then calculates locational marginal prices (LMP) at 5-minute intervals, based on real-time demand, regional system reliability, and transmission constraints. This market balances the differences between day-ahead schedules and real-time operations. This dual-market structure facilitates efficient resource allocation and dynamic pricing that reflects real-time conditions, while providing market participants with mechanisms to mitigate price risks (IRC, 2017). The system's design allows for the integration of variable renewable energy sources and demand response programs, contributing to grid stability and supporting New York State's evolving energy landscape (NYISO, 2022a).

NYISO's market structure has become increasingly important as New York State pursues ambitious renewable energy goals. The state aims to derive 70% of its electricity from renewable sources by 2030 and achieve 100% clean energy by 2040. As of 2022, renewable sources account for approximately 30% of the state's electricity production, with hydropower as the primary contributor, followed by wind and solar energy (NYISO, 2022a). This shift toward renewables has been marked by significant growth in certain sectors. Solar capacity expanded from less than 50 megawatts (MW) in 2010 to over 2,000 MW by 2022, while wind capacity increased from approximately 1,300 MW to exceeding 2,000 MW over the same period. Hydropower capacity, however, has remained relatively stable at approximately 5,000 MW, reflecting its established role in the state's energy mix and the complex operational characteristics that distinguish hydropower from variable renewable sources through state-dependent water values under inflow uncertainty (NYISO, 2022b, Pearce et al., 2025). These changes in the renewable-integrated energy landscape have important implications for market dynamics, including the economic phenomenon of negative electricity prices examined in this study.

2.2. Economic Theory of Negative Price Formation

A theoretical framework for understanding negative electricity prices in renewable-integrated markets is developed here. The approach proceeds in three stages: first establishing the core economic mechanisms through a formal optimization model, then translating these theoretical insights into empirically testable relationships, and finally deriving specific predictions about the key drivers of negative price formation.

2.2.1. Market Economics and Price Dynamics

The emergence of negative electricity prices stems from the economic interplay of limited storage capabilities, unit commitment constraints, and highly inelastic demand. To formalize these dynamics, a wholesale electricity market operating over discrete time intervals $t \in \mathcal{T} = \{1, \ldots, 24\}$ representing hours in a day is considered.[1] Let $\Omega = \{1, \ldots, N\}$ denote the set of zones, with each zone $i \in \Omega$ containing a set of baseload generators Ω_B^i and renewable generators Ω_R^i.

The theoretical framework emphasizes unit commitment variables and their economic implications, establishing the foundations for understanding why generators might rationally offer negative prices. While the model incorporates detailed commitment variables for theoretical completeness, the subsequent empirical analysis focuses on observable market characteristics available through public data sources, providing a practical bridge between theory and empirical testing.

Conventional generators face turn-on and turn-off constraints, and to accommodate these, system operators clear the market in the day-ahead timeframe. The day-ahead market establishes financially binding schedules for energy and ancillary services one day before the operating day, allowing baseload generators to optimize their commitment decisions. The real-time market then balances differences between day-ahead schedules and actual operations at 5-minute intervals, with prices reflecting current system conditions. This two-settlement system enables the market to handle both long-term unit commitment needs and short-term balancing requirements.

Definition 1 (Generator Bid Strategy)**.** A bid strategy for generator g in zone i at time t is a pair $(b_{git}, q_{git}) \in \mathbb{R} \times \mathbb{R}_+$ where b_{git} is the bid price and q_{git} is the offered quantity. A generator bids a price-quantity pair for each hour.

Generator Cost Structure. Each generator's total cost comprises production, cycling, and fixed operational costs. The cycling cost function captures the complexity of output adjustments:

$$\Phi_{git}(q_{git}, q_{g,i,t-1}, u_{git}) = \kappa_g \max\{u_{git} - u_{g,i,t-1}, 0\} + \delta_g \max\{u_{g,i,t-1} - u_{git}, 0\} + \gamma_g |q_{git} - q_{g,i,t-1}| \quad (1)$$

[1] While NYISO's real-time market operates at 5-minute intervals, the theoretical framework uses hourly intervals for simplicity. This simplification does not affect the theoretical predictions, as the core economic mechanisms driving negative prices operate similarly at both time scales. The empirical analysis employs both 5-minute interval data and hourly aggregated data to ensure robustness across different time scales.

where u_{git} is a binary variable indicating whether generator g in zone i at time t is operational ($u_{git} = 1$) or not ($u_{git} = 0$). The first term κ_g represents startup costs, which may include expenses such as fuel consumption during warm-up periods, equipment wear from thermal stress, and associated labor requirements. The second term δ_g captures shutdown costs, reflecting the economic opportunity cost of ceasing operations and expenses required for eventual restart. The third term γ_g denotes ramping costs that arise when generators adjust output levels, which may stem from efficiency losses, equipment degradation, or operational constraints. These cycling costs create economic incentives for generators to maintain relatively stable output levels, particularly for baseload units where operational changes can impose substantial expenses. The fixed cost $\eta_g u_{git}$ applies when the generator is operational ($u_{git} = 1$).

Renewable Subsidies. Renewable generators receive production-based incentives defined by:

$$SUB_{git}(q_{git}) = s_g q_{git} + \xi_g \mathbb{1}\{q_{git} > 0\} \tag{2}$$

where s_g represents per-unit production subsidies and ξ_g denotes fixed renewable energy credits that apply whenever generation is positive.

Operational Constraints. Generator operations must satisfy key physical constraints:

$$\underline{q}_g u_{git} \leq q_{git} \leq \overline{q}_g u_{git} \qquad \text{(generation limits)} \tag{3}$$

$$-\underline{r}_g \leq q_{git} - q_{g,i,t-1} \leq \overline{r}_g \qquad \text{(ramping constraints)} \tag{4}$$

Equation 3 sets minimum ($\underline{q}_g$) and maximum ($\overline{q}_g$) generation levels. Equation 4 limits output changes to maximum ramp-down ($\underline{r}_g$) and ramp-up ($\overline{r}_g$) rates.

Assumption 1 (Cost Structure)**.** *Each generator g's cost structure satisfies:*

(a) *Production cost $A_g(q_{git})$, representing fuel and variable operating expenses, is continuous, strictly increasing, and convex in q_{git}*

(b) *Cycling costs Φ_{git} follow Equation 1 with $\kappa_g, \delta_g, \gamma_g \geq 0$*

(c) *Fixed costs $\eta_g \geq 0$ apply during operation*

(d) *Renewable subsidies SUB_{git} follow Equation 2 with $s_g, \xi_g \geq 0$*

Generator Optimization. Given this structure, each generator solves:

$$\max_{q_{git},u_{git}} \sum_{t\in\mathcal{T}} \beta^t [p_{it}q_{git} - A_g(q_{git}) - \Phi_{git}(q_{git}, q_{g,i,t-1}, u_{git}) - \eta_g u_{git} + SUB_{git}(q_{git})] \tag{5}$$

subject to constraints 3-4, where β is the intra-day discount factor and p_{it} is the market clearing price. While generators optimize over quantities q_{git} and commitment decisions u_{git} in this formulation, in practice they submit price-quantity pairs as bids. The optimal bid price reflects the generator's marginal willingness to produce at each quantity level, derived from the first-order conditions of this optimization problem with respect to q_{git}. This bidding strategy incorporates both direct production costs and the shadow values of operational constraints, ensuring that bids reflect the true economic cost of generation under prevailing market conditions.

Lemma 1 (Optimal Bidding Strategy)**.** *Under Assumption 1, for generator g facing operational constraints 3-4, the optimal bid price b^*_{git} satisfies:*

$$b^*_{git} = \frac{\partial A_g(q_{git})}{\partial q_{git}} + \frac{\partial \Phi_{git}}{\partial q_{git}} - \frac{\partial SUB_{git}}{\partial q_{git}} - (\lambda^{min}_{git} - \lambda^{max}_{git} + \mu^{down}_{git} - \mu^{up}_{git}) \tag{6}$$

where $\lambda^{min}_{git}, \lambda^{max}_{git}, \mu^{down}_{git}, \mu^{up}_{git} \geq 0$ are the Lagrange multipliers for constraints 3-4.[2]

Proposition 1 (Negative Bid Conditions)**.** *Under Assumption 1, a generator submits negative bids if and only if:*

$$\frac{\partial SUB_{git}}{\partial q_{git}} + \lambda^{min}_{git} + \mu^{down}_{git} > \frac{\partial A_g(q_{git})}{\partial q_{git}} + \frac{\partial \Phi_{git}}{\partial q_{git}} \tag{7}$$

Corollary 1 (Sources of Negative Bids)**.** *Negative bids arise through two distinct economic mechanisms. The first involves baseload generators with binding minimum generation constraints ($\lambda^{min}_{git} > 0$), typically during low demand periods when maintaining minimum output is technically necessary but economically costly. The second operates through renewable generators with production subsidies ($\partial SUB_{git}/\partial q_{git} > 0$) that exceed marginal production costs, enabling profitable operation even at negative market prices.*

Proposition 2 (Market Clearing Efficiency)**.** *Let p^*_{it} be the market clearing price determined by:*

$$\sum_{g\in\Omega^i_B\cup\Omega^i_R} q^*_{git}(p^*_{it}) = D_{it}(p^*_{it}) \tag{8}$$

[2]All proofs are provided in Appendix A.

where q^*_{git} are optimal quantities from Equation 5 and D_{it} is zonal demand. Under negative price allowance, the resulting allocation achieves:

$$\max_{q_{git}, u_{git}} \sum_{i \in \Omega} \sum_{t \in \mathcal{T}} \left[\int_0^{Q_{it}} P_{it}(q) dq - \sum_{g \in \Omega^i_B \cup \Omega^i_R} TC_g(q_{git}) - \sum_{\ell} Losses_\ell \right] \tag{9}$$

subject to constraints 3-4 and market clearing conditions, where TC_g includes all costs net of subsidies, and $Losses_\ell$ represents the system losses on transmission elements ℓ. This demonstrates that allowing negative prices is important for achieving economic efficiency in renewable-integrated electricity markets with operational constraints and renewable subsidies.

Price Formation. The market clearing price reflects both local conditions and network effects:

$$p^*_{it} = p^{sys}_t + C_{it} + L_{it} \tag{10}$$

where p^{sys}_t is the system-wide energy price, C_{it} represents marginal congestion costs, and L_{it} captures marginal loss components. When transmission constraints bind, congestion costs create price separation between zones:

$$|C_{it}| = \nu_{ijt}(\overline{F}_{ijt} - F_{ijt}) \tag{11}$$

where ν_{ijt} is the shadow price of transmission capacity $\overline{F}_{ijt}$ between zones i and j, and F_{ijt} is actual power flow.

This theoretical framework elucidates the two market scenarios illustrated in Figure 2. In the efficient equilibrium (Figure 2a), negative prices facilitate optimal market clearing, achieving a welfare-maximizing allocation characterized by:

$$p^* = P^d(Q^*) = \min_g MC_g(q^*_g) \tag{12}$$

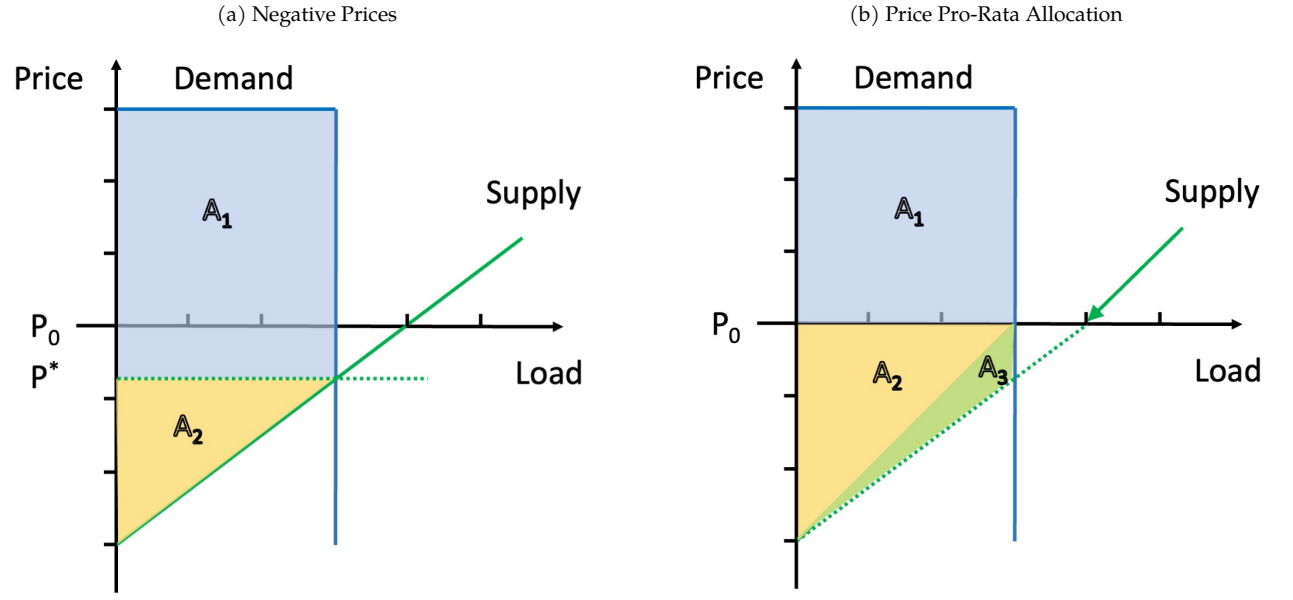
where $P^d(Q)$ is the inverse demand function, Q^* is the optimal quantity, and $MC_g(q_g)$ represents generator g's effective marginal cost including cycling costs and subsidies. Areas A_1 and A_2 represent consumer and producer surplus respectively.

Conversely, implementing non-negative price constraints requires alternative mechanisms to manage oversupply. Pro-rata allocation (Figure 2b) represents such a regulatory approach, wherein output reductions during oversupply periods are distributed proportionally across all

generators based on their capacity share, independent of their marginal costs or operational characteristics. Under this mechanism, system operators maintain a price floor at P_0 (typically zero) and mandate uniform percentage curtailments across all generating units when supply exceeds demand. This approach introduces market distortions through two primary channels. First, the artificial price floor prevents price signals from reflecting true opportunity costs of generation. Second, it distorts the merit order by forcing uniform output reductions across generators with heterogeneous cost structures, including baseload facilities facing high cycling costs ($\partial\Phi_{git}/\partial q_{git}$) and renewable facilities with production-linked subsidies ($\partial SUB_{git}/\partial q_{git}$). Area A_3 quantifies the resulting welfare loss from this inefficient allocation.

The economic rationale for negative prices largely derives from generators' opportunity costs (Cramton, 2004). Baseload generators confront significant costs beyond immediate variable expenses when reducing output, including startup costs (κ_g) and thermal cycling degradation (δ_g) from Equation 1. Similarly, renewable generators face opportunity costs during curtailment, such as foregone subsidies (s_g) and contractual penalties (ξ_g) from Equation 2. Including these opportunity costs in bidding strategies through Equation 6, even when resulting in negative prices, prevents suboptimal output reduction and maintains efficient production patterns as proven in Proposition 2.

Figure 2: Negative Prices vs. Price Pro-Rata Allocation



Notes: Panel (a) shows efficient market clearing with negative prices allowed. Areas A_1 and A_2 represent consumer and producer surplus respectively. Panel (b) shows distorted equilibrium under a non-negative price constraint. Area A_3 represents deadweight loss from inefficient allocation.

Economic Intuition. The optimization framework yields key insights into the economic logic of negative price formation. Negative prices emerge from economically rational responses to operational realities and represent efficient market signals responding to supply-demand imbalances. When generators face high costs from operational changes, such as cycling thermal units or curtailing subsidized renewable production, they may rationally prefer paying to produce and thereby avoid these opportunity costs. Allowing negative prices maintains economic efficiency by enabling appropriate incentive structures: generators with lowest total costs, including cycling costs and foregone subsidies, continue operating while higher-cost generators reduce output, achieving a cost-minimizing solution to balancing supply and demand. Conversely, suppressing negative prices through regulatory constraints creates economically inefficient outcomes, as alternative mechanisms like pro-rata curtailment treat all generators uniformly despite heterogeneous cost structures, reducing economic value by curtailing low-opportunity-cost and high-opportunity-cost units at equal rates.

2.3. From Theory to Empirical Predictions

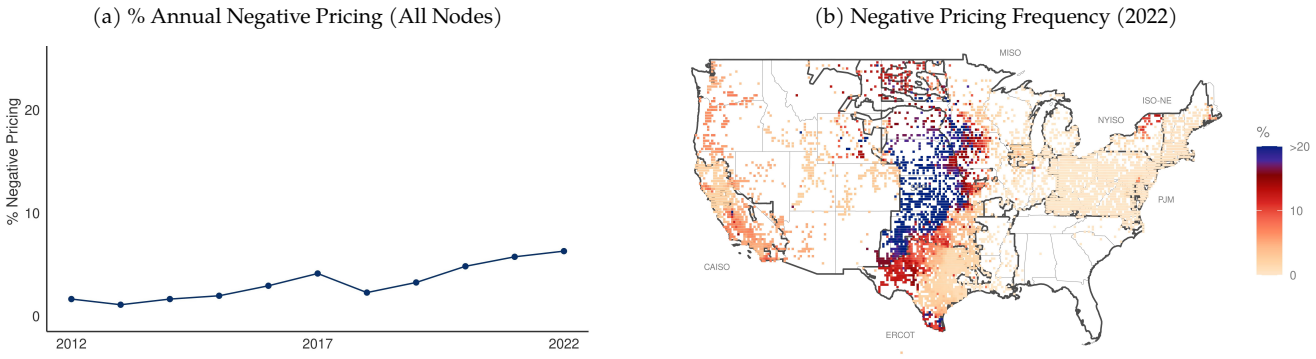
The micro-level generator optimization framework is bridged here to observable market patterns that enable empirical testing. While the theoretical model operates through unit commitment variables (u_{git}), cycling costs (Φ_{git}), and renewable subsidies (SUB_{git}), empirical analysis relies on publicly available market data.

The generator optimization problems in Equations 5-6 aggregate through market clearing to produce relationships between observable market characteristics and negative price formation. When multiple generators simultaneously face binding operational constraints or benefit from production subsidies, their individual bidding decisions create market-wide patterns detectable without proprietary unit commitment data. These theoretical mechanisms translate into empirical patterns through three channels: *renewable integration effects* operate through variable generation patterns and production-based subsidies, creating relationships between renewable potential and price formation; *weather-driven effects* capture both supply-side influences on renewable generation and demand-side impacts on system stress; and *grid constraint effects* emerge when transmission limitations create localized oversupply, leading generators to bid negative prices instead of reducing output below technical minimums.

These theoretical insights find compelling validation in real electricity markets. U.S. electricity markets have experienced a gradual increase in negative price occurrences that coincides precisely with the renewable energy transition predicted by the theoretical framework. As

illustrated in Figure 3, Panel (a) shows an upward trend in the nationwide frequency of negative prices, rising from approximately 2% in 2012 to nearly 7% by 2022—a period that corresponds with accelerated wind and solar deployment. The geographical distribution shown in Panel (b) provides further evidence of the theoretical mechanisms at work, indicating regional heterogeneity that reflects varying renewable penetration levels and transmission constraints. Central regions exceeding 20% frequency align with areas of high wind capacity and limited transmission export capability, while eastern markets, including NYISO, exhibit frequencies below 5%, consistent with more balanced generation portfolios and interconnected transmission networks.[3]

Figure 3: Negative Electricity Prices in the U.S.



Source: https://emp.lbl.gov/renewables-and-wholesale-electricity-prices-rewep

This empirical reality motivates a systematic approach to identifying causal relationships. Exogenous variation in the fundamental drivers of these mechanisms is exploited to avoid using endogenous generation data. This approach maintains causal interpretation while avoiding simultaneity bias, since the theoretical conditions determining negative price formation respond to observable, exogenous factors. Building on this foundation, the probability of negative prices is formalized and testable predictions are developed that transform the theoretical insights into empirically verifiable relationships:

Definition 2 (Negative Price Probability)**.** The probability of negative prices in zone i at time t is:

$$Pr(N_{it} = 1) = Pr(p^*_{it} < 0|\mathbf{X}_{it}) = F(\mathbf{X}^\top_{it}\boldsymbol{\beta}) \tag{13}$$

[3]The moderate frequency of negative prices in the NYISO market makes it suitable for empirical investigation. This measured level of occurrence provides adequate variation for statistical analysis while avoiding the potential confounding effects present in markets with more extreme frequencies.

where $\mathbf{X}_{it} = (R_{it}, W_{it}, G_{it})$ represents renewable integration, weather conditions, and grid constraints, and $F(\cdot)$ is a cumulative distribution function.

Renewable Integration Effects. The theoretical framework predicts that production-linked subsidies enable renewable generators to bid negative prices while maintaining profitability (Seel et al., 2021). From Equation 2, the marginal subsidy benefit s_g reduces the optimal bid price. This mechanism explains the temporal correlation observed in Figure 3, where increasing negative price frequency mirrors renewable capacity expansion. Antweiler and Muesgens (2021) demonstrate how renewable resources with near-zero marginal costs fundamentally alter wholesale price dynamics through merit order effects. Wind energy creates oversupply conditions during off-peak hours when production tax credits enable profitable operation at negative market prices (Bushnell and Novan, 2018), while solar generation aligns more predictably with daily demand patterns, reducing negative price occurrences during peak daylight hours (Woo et al., 2016).

Empirical Prediction 1 (Renewable Integration Impact)**.** The marginal effect of renewable generation from source $r \in \{Sol, Wind, Hydro\}$ on negative price probability is:

$$\frac{\partial Pr(N_{it} = 1)}{\partial R_{r,it}} = f(\mathbf{X}_{it}^{\top}\boldsymbol{\beta}) \left(\frac{\partial s_r}{\partial R_{r,it}} - \frac{\partial MC_r}{\partial R_{r,it}} \right) \tag{14}$$

where s_r is the subsidy rate and MC_r is the marginal cost of generation type r.

Weather Effects. The dual supply-demand channel predicted by theory manifests in the seasonal and regional patterns evident in the market data. Weather patterns influence negative price formation through both supply and demand channels. Mosquera-López et al. (2024) demonstrate how temperature and humidity affect renewable generation potential through nonlinear relationships, while simultaneously influencing electricity consumption patterns. Staffell and Pfenninger (2018) show how weather variability creates conditions where renewable output can exceed demand during moderate weather periods, particularly when heating and cooling loads are minimal while wind and solar resources remain abundant, affecting supply-demand imbalances that lead to negative pricing.

Empirical Prediction 2 (Weather Impact)**.** The marginal effect of weather condition $w \in \{T, H\}$

on negative price probability is:

$$\frac{\partial Pr(N_{it} = 1)}{\partial w_{it}} = f(\mathbf{X}_{it}^{\top}\boldsymbol{\beta}) \left(\frac{\partial Q_{it}^{S}}{\partial w_{it}} - \frac{\partial D_{it}}{\partial w_{it}} \right) \tag{15}$$

where Q_{it}^{S} and D_{it} represent aggregate supply and demand.

Grid Constraint Effects. The spatial heterogeneity documented in Figure 3 reflects the theoretical prediction that transmission constraints create localized oversupply when generation capacity exceeds local demand and export capabilities. LaRiviere and Lyu (2022) provide direct causal evidence using Texas's CREZ transmission expansion, demonstrating how eliminating interregional constraints reduces generation costs and negative pricing frequency. Schermeyer et al. (2018) distinguish between market-driven and grid-constraint curtailment, showing how transmission limitations force renewable generators to reduce output regardless of market prices. From Equation 10, marginal congestion costs (C_{it}) and loss components (L_{it}) create price separation between zones, explaining why wind-rich regions experience higher negative price frequencies than interconnected eastern markets.

Empirical Prediction 3 (Grid Constraint Impact)**.** The marginal effect of grid constraint measure $g \in \{C, L\}$ on negative price probability is:

$$\frac{\partial Pr(N_{it} = 1)}{\partial g_{it}} = f(\mathbf{X}_{it}^{\top}\boldsymbol{\beta})\nu_{ijt} \tag{16}$$

where ν_{ijt} is the shadow price of transmission capacity.

These empirical predictions, grounded in both theoretical insight and observed market patterns, provide the foundation for testing the framework using NYISO data while addressing key identification challenges.

3. Empirical Strategy

This section presents the empirical methodology for testing the theoretical predictions. Section 3.1 specifies the empirical framework linking theory to data. Section 3 develops binary response models for occurrence analysis and count data models for frequency analysis. Marginal effects are then computed to quantify the relative importance of each driver. Finally, key econometric challenges including endogeneity and omitted variable bias are addressed through identification strategies that support causal inference.

3.1. Empirical Framework

Building on the theoretical probability of negative prices from Equation 13, the empirical model is specified using the vector of characteristics $\mathbf{X}_{it}$ for each zone $i \in \Omega$ and time $t \in \mathcal{T}$:

$$\mathbf{X}_{it} = (1, \underbrace{C_{it}, L_{it}}_{\text{grid}}, \underbrace{Wind_{it}, Sol_{it}, Hydro_{it}}_{\text{renewable}}, \underbrace{T_{it}, H_{it}}_{\text{weather}})^\top \tag{17}$$

where C_{it} represents the marginal costs of congestion (MCC), L_{it} the marginal costs of losses (MCL), $Wind_{it}$ wind energy production, Sol_{it} solar energy production, $Hydro_{it}$ hydropower production, T_{it} temperature, and H_{it} humidity. This specification maps directly to the theoretical drivers in Equations 6 and 10.

3.2. Binary Response Model for Testing Occurrence Predictions

To test the theoretical framework established in Empirical Predictions 1-3 about negative price probability, electricity prices are first transformed into a binary variable:

$$N_{it} = \begin{cases} 0 & \text{if } LMP_{it} \geq 0 \\ 1 & \text{if } LMP_{it} < 0 \end{cases} \tag{18}$$

where LMP_{it} denotes the locational marginal price corresponding to market clearing price p^*_{it} from Equation 10. This transformation offers several methodological advantages. First, it focuses specifically on the factors causing negative prices rather than general price fluctuations. Second, specialized econometric techniques designed for analyzing rare occurrences can be applied by treating negative prices as discrete events. Third, the binary outcome helps isolate the comparative impact of different drivers.

Following the latent variable approach suggested by the theoretical framework, the propensity for negative prices is modeled as:

$$y^*_{it} = \mathbf{X}^\top_{it}\boldsymbol{\beta} + \gamma_i + \delta_t + \epsilon_{it} \tag{19}$$

Here, γ_i captures zone-specific factors affecting optimal bids through Equation 6, while δ_t controls for temporal variations in market conditions. The observed outcome follows:

$$N_{it} = \begin{cases} 1 & \text{if } y^*_{it} > 0 \\ 0 & \text{if } y^*_{it} \leq 0 \end{cases} \tag{20}$$

The probability of negative prices, corresponding to Equation 13, is:

$$Pr(N_{it} = 1|\mathbf{X}_{it}, \gamma_i, \delta_t) = F(\mathbf{X}_{it}^\top \boldsymbol{\beta} + \gamma_i + \delta_t) \tag{21}$$

where $F(\cdot)$ is a cumulative distribution function. Two specifications are considered that correspond to different assumptions about the distribution of shocks to optimal bids:

$$F(\cdot) = \Lambda(\cdot) = \frac{\exp(\cdot)}{1 + \exp(\cdot)} \quad \text{(Logit)} \tag{22}$$

$$F(\cdot) = \Phi(\cdot) = \int_{-\infty}^{(\cdot)} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz \quad \text{(Probit)} \tag{23}$$

3.3. Count Data Models for Testing Frequency Predictions

To complement the binary response analysis and examine the theoretical predictions about the intensity of negative price occurrences, count outcomes over different time horizons are analyzed. Let $Fr(N)_{it}$ represent the number of negative price events in zone i during period t. Building on the probability model and the theoretical framework from Section 2, the specification is:

$$E[Fr(N)_{it}|\mathbf{X}_{it}, \gamma_i, \delta_t] = \exp(\mathbf{X}_{it}^\top \boldsymbol{\theta} + \gamma_i + \delta_t) \tag{24}$$

where $\boldsymbol{\theta} = (\theta_0, \theta_1, \ldots, \theta_7)^\top$ represents the impacts of grid constraints, renewable integration, and weather conditions on the frequency of negative prices. The exponential function ensures non-negative predicted values, appropriate for count data.

The analysis begins with a Poisson regression specification:

$$Pr(Fr(N)_{it} = k|\mathbf{X}_{it}, \gamma_i, \delta_t) = \frac{e^{-\mu_{it}} \mu_{it}^k}{k!}, \quad k = 0, 1, 2, \ldots \tag{25}$$

where $\mu_{it} = \exp(\mathbf{X}_{it}^\top \boldsymbol{\theta} + \gamma_i + \delta_t)$. This model imposes equidispersion:

$$E[Fr(N)_{it}|\mathbf{X}_{it}, \gamma_i, \delta_t] = Var(Fr(N)_{it}|\mathbf{X}_{it}, \gamma_i, \delta_t) = \mu_{it} \tag{26}$$

To allow for overdispersion in negative price occurrences, which may arise from the clustering predicted by the theoretical framework, a Negative Binomial model is also specified:

$$Pr(Fr(N)_{it} = k|\mathbf{X}_{it}, \gamma_i, \delta_t) = \frac{\Gamma(k + \frac{1}{\alpha})}{\Gamma(k+1)\Gamma(\frac{1}{\alpha})} \left(\frac{\mu_{it}}{\mu_{it} + \frac{1}{\alpha}}\right)^k \left(\frac{\frac{1}{\alpha}}{\mu_{it} + \frac{1}{\alpha}}\right)^{\frac{1}{\alpha}} \tag{27}$$

with conditional moments:

$$\begin{aligned} E[Fr(N)_{it}|\mathbf{X}_{it}, \gamma_i, \delta_t] &= \mu_{it} \\ Var(Fr(N)_{it}|\mathbf{X}_{it}, \gamma_i, \delta_t) &= \mu_{it} + \alpha\mu_{it}^2 \end{aligned} \tag{28}$$

For additional robustness, a Generalized Negative Binomial model that introduces shape parameter p is also considered:

$$Var(Fr(N)_{it}|\mathbf{X}_{it}, \gamma_i, \delta_t) = \mu_{it} + \alpha\mu_{it}^{2-p} \tag{29}$$

3.4. Marginal Effects for Testing Theoretical Predictions

To assess the impact of each identified driver, average marginal effects of the variables are computed. For the binary response model, the marginal effect of characteristic k is:

$$\frac{\partial Pr(N_{it} = 1|\mathbf{X}_{it}, \gamma_i, \delta_t)}{\partial x_k} = f(\mathbf{X}_{it}^\top \boldsymbol{\beta} + \gamma_i + \delta_t)\beta_k \tag{30}$$

where $f(\cdot)$ is the density function corresponding to $F(\cdot)$. For the logit model, $f(\cdot) = \Lambda(\cdot)(1 - \Lambda(\cdot))$, while for the probit model, $f(\cdot) = \phi(\cdot)$, where $\phi(\cdot)$ denotes the standard normal density function.

For count models, the marginal effects are:

$$\frac{\partial E[Fr(N)_{it}|\mathbf{X}_{it}, \gamma_i, \delta_t]}{\partial x_k} = \exp(\mathbf{X}_{it}^\top \boldsymbol{\theta} + \gamma_i + \delta_t)\theta_k \tag{31}$$

To obtain a summary measure of these effects across all observations, the average marginal effect is calculated:

$$AME_k = \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} \frac{\partial y_{it}}{\partial x_k} \tag{32}$$

where y_{it} represents either $Pr(N_{it} = 1)$ or $E[Fr(N)_{it}]$ depending on the model. These marginal effects provide a standardized measure of the impact of each explanatory variable, facilitating direct comparison with the theoretical predictions.

3.5. Econometric Challenges

Two important challenges in the estimation strategy that could affect testing the theoretical predictions are addressed: omitted variable bias and simultaneity in renewable energy data. Both issues represent forms of endogeneity that could lead to biased and inconsistent estimates if not properly addressed.

3.5.1. Addressing Omitted Variable Bias

Omitted variable bias could arise if unobserved factors affect both the drivers and negative price occurrence. To address this concern, the high-frequency data are exploited through a comprehensive fixed effects strategy:

$$y_{it} = \mathbf{X}_{it}^{\top}\boldsymbol{\beta} + \gamma_i + \delta_y + \delta_m + \delta_d + \delta_w + \delta_h + \delta_5 + \epsilon_{it} \tag{33}$$

where y_{it} represents the dependent variable (either N_{it} or $Fr(N)_{it}$), and the fixed effects terms capture:

- γ_i: Zone fixed effects controlling for time-invariant characteristics affecting optimal bids through Equation 6
- δ_y: Year fixed effects absorbing long-term trends and policy changes
- δ_m: Month fixed effects accounting for seasonal patterns
- δ_d: Calendar day fixed effects capturing specific day variations within months (e.g., day 1, day 2, ..., day 31)
- δ_w: Day of the week fixed effects controlling for weekly cycles (Monday, Tuesday, ..., Sunday)
- δ_h: Hour fixed effects accounting for intra-day patterns
- δ_5: 5-minute fixed effects capturing high-frequency variations

This comprehensive set of fixed effects helps address omitted variable bias through several channels. First, zone fixed effects absorb time-invariant differences in market structures, geographical features, and local regulations that could influence negative price occurrences through generator cost functions and constraints. Second, the temporal fixed effects capture a range of common shocks and trends at different frequencies. Third, high-frequency fixed effects control for market design features like bidding windows and clearing processes.

3.5.2. Addressing Simultaneity in Renewable Energy Data

The second challenge concerns potential simultaneity between renewable energy production and electricity prices. Following the theoretical decomposition in Equations 6 and 2, actual renewable production can be decomposed as:

$$A_{it} = \bar{A}_i + \tilde{A}_{it} \tag{34}$$

where $\bar{A}_i$ represents spatiotemporal-fixed components captured by fixed effects (including installed capacity, geographical features, and technological advancement), and $\tilde{A}_{it}$ denotes variation driven by exogenous weather conditions.

The weather-based proxies $Wind_{it}$, Sol_{it}, and $Hydro_{it}$ capture this exogenous variation, satisfying:

$$E[\epsilon_{it}|Wind_{it}] = E[\epsilon_{it}|Sol_{it}] = E[\epsilon_{it}|Hydro_{it}] = 0 \tag{35}$$

This approach relies on two important assumptions supported by the literature: (1) a strong correlation between the proxies and the time-varying component of actual production (Katzenstein et al., 2010, Abrell et al., 2019), and (2) the proxies serve as unbiased predictors of this time-varying component on average (Staffell and Pfenninger, 2016, Chu et al., 2021).

One potential limitation of this approach warrants consideration: zonal weather variables may not adequately capture the interconnected structure of electricity markets, where generation in one zone can influence prices in other zones through inter-zonal power flows. However, several characteristics of the NYISO market and the empirical strategy support this approach. First, New York State's renewable energy imports remained relatively constrained during the study period, with indigenous renewable generation capacity primarily serving local demand. Second, the fixed effects strategy controls for inter-zonal power flow patterns.[4] Third, transmission constraints and energy losses are explicitly incorporated through C_{it} and L_{it}, which directly quantify the economic implications of inter-zonal power flows.

4. Data

To examine the economics of negative price phenomena in renewable-integrated electricity markets, an extensive dataset from New York's wholesale electricity markets spanning January

[4]In Appendix C, the empirical strategy is further validated through alternative fixed effects specifications, including zone-year, zone-month, and year-month interactions.

1, 2010, to June 30, 2022, is employed. This period captures an important phase of renewable energy expansion, providing over 14 million observations that enable detailed analysis of the economic drivers behind negative price occurrences during the clean energy transition.

New York's electricity market is selected as the empirical setting for several practical reasons. First, the market experiences moderate negative price frequencies (approximately 1% of all intervals), providing sufficient variation for robust statistical analysis while avoiding the more extreme conditions present in some regional markets that could confound results. Second, New York's diverse generation portfolio—including conventional thermal plants, hydroelectric facilities, and growing wind and solar capacity—creates the mixed resource conditions necessary to identify distinct economic effects across different generation technologies. Third, the state's ambitious renewable energy policies during the study period provide natural variation in renewable penetration levels that facilitates identification of integration effects on market outcomes.

The dataset is constructed by combining two primary data sources: real-time pricing information from the NYISO database and meteorological data from NASA POWER. This combination allows capture of both the market outcomes this study seeks to explain and the exogenous weather variation that drives renewable generation potential.

4.1. Electricity Market Data

Core pricing data are obtained from NYISO, including locational marginal prices (LMP), marginal congestion costs (MCC), and marginal losses costs (MCL) collected at 5-minute intervals across all 11 NYISO control area load zones.[5] This high-frequency pricing information captures the rapid market dynamics important for understanding negative price formation mechanisms predicted by the theoretical framework.

The 5-minute frequency provides analytical advantages. It preserves the native frequency of market operations, enabling identification of rapid price dynamics and intra-hour market responses that may not be visible in aggregated data. While this high frequency offers detailed insights into short-term market behavior, it naturally contains more noise that could obscure underlying economic relationships.

[5] Archived real-time pricing market data are publicly accessible at http://mis.nyiso.com/public/P-24Alist.htm.

4.2. *Weather and Renewable Generation Data*

To complement the pricing data, hourly renewable energy potential and weather variables for the corresponding NYISO zones and time period are incorporated, obtained from NASA POWER.[6] For each zone, the mean latitude and longitude of all generators serve as the geographic coordinates to derive the relevant meteorological data.

These weather-based proxies for renewable generation potential are used to address the endogeneity concerns discussed in the empirical strategy while capturing the exogenous variation needed to identify causal effects, with theoretical justification for these variables provided in Section 2.3 and methodological support for the proxy approach detailed in Section 3.5.[7]

Table 1 provides detailed descriptions of the renewable energy and weather variables obtained from NASA POWER.[8] These variables serve as instrumental proxies that capture weather-driven variation in renewable generation potential while avoiding the simultaneity bias that would arise from using actual generation data.

Table 1: Renewable and Weather Data Description from NASA POWER

Wind Speed	MERRA-2 Wind Speed at 10 Meters (m/s)
Sun Irradiance	CERES SYN1deg All Sky Surface Shortwave Downward Irradiance (kWh/m^2)
Precipitation	MERRA-2 Precipitation Corrected (mm/hour)
Temperature	MERRA-2 Temperature at 2 Meters (C)
Humidity	MERRA-2 Specific Humidity at 2 Meters (g/kg)

4.3. *Dataset Construction and Summary Statistics*

To ensure robust analysis across different time scales, two datasets with distinct temporal structures are constructed. The *5-minute dataset* preserves the native frequency of market operations, enabling identification of rapid price dynamics and intra-hour market responses that may not be visible in aggregated data. The *hourly dataset* aggregates 5-minute observations to their hourly means, potentially reducing noise and facilitating identification of broader

[6]These data are freely accessible through their web interface: https://power.larc.nasa.gov/data-access-viewer/ and feature a grid resolution of half an arc degree latitude by half an arc degree longitude. As per NASA POWER's request for acknowledgment: *"These data were obtained from the NASA Langley Research Center (LaRC) POWER Project funded through the NASA Earth Science/Applied Science Program."* For more information, see https://power.larc.nasa.gov/#contact.

[7]The analysis period primarily captures the effects of land-based wind generation, as offshore wind development in New York was still in planning stages. Future research will need to examine how the different generation profiles and geographic distribution of offshore wind might alter these dynamics.

[8]NASA POWER provides multiple data series for each variable, derived from distinct measurements. These series typically exhibit high correlation within each variable. A single series per variable is selected, prioritizing those with the highest variance.

economic trends while risking the loss of important short-term dynamics and introducing aggregation bias.

This dual-frequency approach enables cross-validation of findings and ensures consistent results across different time scales relevant to various market participants and policy decisions. Power system operators make real-time decisions at 5-minute intervals, while energy traders and renewable developers often focus on hourly patterns for planning and risk management.

Table 2 presents summary statistics revealing the economic characteristics of the variables. The pricing data shows the typical patterns of wholesale electricity markets, with positive mean prices but notable negative minimum values reaching -$4,444/MWh. The significant price volatility (standard deviation of $52.03/MWh) and high skewness (10.31) highlight the market conditions that create opportunities for negative price formation. Weather variables display the natural variation necessary to identify their economic effects on market outcomes, with wind speeds ranging from calm to storm conditions and solar irradiance following expected diurnal and seasonal patterns.[9]

Table 2: Summary Statistics of the Variables

Variable (Unit)	Min	Q1	Median	Mean	Q3	Max	SD	Kurt	Skew
LMP ($/MWh)	-4444.00	19.78	29.20	37.33	40.38	5974.00	52.03	352.30	10.31
MCC ($/MWh)	-5946.00	0.00	0.00	-5.70	0.00	4527.00	41.85	2048.00	-30.89
MCL ($/MWh)	-545.10	-0.03	1.05	1.46	2.46	293.90	3.87	1909.00	-15.35
Wind Speed(m/s)	0.00	1.74	2.54	3.03	3.85	26.61	1.85	6.94	1.61
Sun Irradiance (kWh/m^2)	0.00	0.00	0.07	0.15	0.24	1.00	0.23	4.37	1.56
Precipitation (mm/hour)	0.00	0.00	0.00	0.13	0.05	19.91	0.47	150.00	9.10
Temperature (C)	-33.12	0.77	9.79	9.49	18.70	38.65	11.04	2.28	-0.18
Humidity (g/kg)	0.18	3.42	5.98	7.02	10.25	22.52	4.35	2.36	0.61

LMP, MCC, MCL, Min, Q1, Q3, Max, SD, Kurt, and Skew are abbreviations for locational marginal price, marginal costs of congestion, marginal costs of losses, minimum, first quartile, third quartile, maximum, standard deviation, kurtosis, and skewness, respectively.

5. Estimation Results

This section presents the estimation results that test the theoretical predictions developed in Section 2. The results come from estimating Equation 19 (probability of negative price occurrence) and Equation 24 (frequency of negative price occurrence) using both 5-minute and hourly datasets. The analysis is organized to move from model selection through coefficient interpretation to economic significance assessment.

[9]Comprehensive descriptive analysis of negative price occurrence patterns across spatial, temporal, and frequency dimensions is provided in Appendix B, including detailed examination of zone-specific heterogeneity, annual trends, seasonal patterns, daily cycles, and intra-hour dynamics.

Results are displayed in Tables 3, 4, 5, and 6, which share three main characteristics. First, they employ a comprehensive set of zone and time fixed effects to control for unobserved time-invariant characteristics of each zone and common time-varying factors. Second, ordinary least squares (OLS) method results are presented in the first column of every table solely for comparison purposes. While OLS is widely used, it is not ideal for this analysis because it can produce heteroscedastic residuals and may generate illogical predictions for binary and count variables. Third, all estimation methods use robust standard errors clustered at the zonal level to address intra-cluster correlation arising from zone-specific factors. This clustering approach is consistent with Abadie et al. (2023), who establish that clustering is warranted when assignment mechanisms—here, the zone-specific deployment of renewable resources, transmission infrastructure, and locational market characteristics—generate dependence in outcomes within clusters over time.

5.1. Model Selection and Probability Analysis

To estimate the coefficients of the explanatory variables in Equation 19, dichotomous dependent variable regression models are employed, specifically Probit and Logit. Model selection is based on log-likelihood, Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC).[10]

Tables 3 and 4 show statistics that lead to selection of the Logit model, as all three criteria indicate a better fit. In the 5-minute data analysis (Table 3), the Logit model demonstrates better fit, shown by a higher log-likelihood and lower AIC and BIC values compared to the Probit model.

The coefficient analysis reveals patterns among the variables examined that align with the theoretical predictions. The MCC (C_{it}) exhibit positive but statistically insignificant coefficients across all models, consistent with the theoretical prediction that grid constraints have minimal impact. The MCL (L_{it}) demonstrate consistent negative effects across all models, with increasing magnitude and varying significance levels. Wind energy ($Wind_{it}$) exhibits a notable pattern, with positive and highly significant coefficients in the Probit and Logit models, while remaining

[10]The use of MLE naturally leads to the employment of log-likelihood, AIC, and BIC for model selection. Log-likelihood, fundamental to MLE, quantifies model fit. AIC and BIC, derived from the log-likelihood function, extend this by incorporating penalties for model complexity. Specifically, AIC $= 2k - 2\ln(L)$ and BIC $= \ln(n)k - 2\ln(L)$, where k denotes the number of parameters, n represents the sample size, and L is the maximized likelihood. These criteria facilitate balancing goodness-of-fit against model parsimony, guiding selection among nested and non-nested MLE-fitted models. Selection aims to maximize log-likelihood while minimizing AIC and BIC, providing assessment of model performance and complexity.

insignificant in the OLS model. This supports the theoretical prediction that wind generation increases negative price likelihood. Conversely, solar energy (Sol_{it}) shows a significant negative effect in the Probit and Logit models, but lacks significance in the OLS specification, supporting the prediction that solar generation reduces negative price occurrences.

The analysis of hourly data (Table 4) reveals patterns largely consistent with the 5-minute analysis, with some notable distinctions. Coefficients generally exhibit larger magnitudes, reflecting the aggregation effect when transitioning from 5-minute to hourly intervals. A notable difference emerges in the precipitation variable ($Hydro_{it}$), which attains marginal significance across all models, with a positive effect.

5.2. Frequency Analysis and Model Specification

For analyzing the frequency of negative price occurrences (Equation 24), a Poisson regression model is initially employed. However, recognizing the potential for overdispersion, formal tests are conducted which confirm its presence at the 1% significance level in both hourly and daily frequency data.[11] To address this overdispersion, the Negative Binomial model is employed, which relaxes the equal variance and mean assumption of the Poisson model. A Generalized Negative Binomial model is also estimated for comparison. The log-likelihood, AIC, and BIC criteria in Tables 5 and 6 indicate that the Negative Binomial model should be selected, as all three criteria show a better fit.

In the hourly frequency analysis (Table 5), the Negative Binomial model demonstrates better fit. The MCL (L_{it}) exhibit significant negative effects in the Poisson and Negative Binomial models, while remaining insignificant in the OLS and Generalized Negative Binomial specifications. Wind energy ($Wind_{it}$) demonstrates a consistent positive and highly significant effect across all count models. Solar energy (Sol_{it}) exhibits a significant negative effect across all count models, with consistent magnitudes.

The daily frequency models (Table 6) reveal different patterns compared to their hourly counterparts. Notably, the MCL (L_{it}) lose statistical significance across all models at the daily level. Wind energy ($Wind_{it}$) and solar energy (Sol_{it}) maintain their significant effects, with larger coefficients compared to the hourly models.

[11] Formal tests for overdispersion yield p-values < 0.001 for both hourly and daily datasets.

Table 3: Estimation Results for the Probability of 5-Minute Negative Price Occurrence
Dependent Variable: $Pr(N)_{it}$

Estimator	**OLS**	**Probit**	**Logit**
Variables	**(1)**	**(2)**	**(3)**
C_{it}	2.31e-05	0.00110	0.00686
	(6.59e-05)	(0.000958)	(0.00611)
L_{it}	-0.000974**	-0.0283***	-0.0940*
	(0.000430)	(0.0104)	(0.0554)
$Wind_{it}$	0.00391	0.140***	0.333***
	(0.00255)	(0.0323)	(0.0512)
Sol_{it}	-0.00172	-0.333***	-0.826***
	(0.00249)	(0.0228)	(0.142)
$Hydro_{it}$	-0.000122	-0.00147	-0.00322
	(0.000411)	(0.0115)	(0.0285)
T_{it}	0.000675*	0.0257***	0.0624***
	(0.000363)	(0.00119)	(0.00720)
H_{it}	-0.000613	-0.0382***	-0.0945***
	(0.000407)	(0.00399)	(0.0147)
Constant	0.0123*	-2.490***	-5.232***
	(0.00611)	(0.298)	(0.816)
Zone FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Month FE	Yes	Yes	Yes
Day FE	Yes	Yes	Yes
Day of Week FE	Yes	Yes	Yes
Hour FE	Yes	Yes	Yes
5-Minute FE	Yes	Yes	Yes
Log-Likelihood		-677,273.8	-673,333.4
AIC		1,354,570	1,346,689
BIC		1,354,729	1,346,848
Observations	14,458,752	14,458,752	14,458,752

***, ** and *: indicate 1%, 5% and 10% significance levels, respectively. Robust clustered standard errors are in parentheses.

Table 4: Estimation Results for the Probability of Hourly Negative Price Occurrence

Dependent Variable: $Pr(N)_{it}$

Estimator	**OLS**	**Probit**	**Logit**
Variables	**(1)**	**(2)**	**(3)**
C_{it}	1.90e-05	0.00181	0.00946
	(9.62e-05)	(0.00149)	(0.00820)
L_{it}	-0.000965*	-0.0427***	-0.129*
	(0.000523)	(0.0142)	(0.0675)
$Wind_{it}$	0.00314	0.124***	0.307***
	(0.00222)	(0.0343)	(0.0602)
Sol_{it}	-0.00387	-0.385***	-0.973***
	(0.00273)	(0.0381)	(0.165)
$Hydro_{it}$	0.00103*	0.0311*	0.0812*
	(0.000517)	(0.0171)	(0.0435)
T_{it}	0.000696*	0.0305***	0.0774***
	(0.000341)	(0.00291)	(0.0106)
H_{it}	-0.000876*	-0.0524***	-0.135***
	(0.000442)	(0.00703)	(0.0191)
Constant	0.00962	-2.507***	-5.226***
	(0.00543)	(0.264)	(0.732)
Zone FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Month FE	Yes	Yes	Yes
Day FE	Yes	Yes	Yes
Day of Week FE	Yes	Yes	Yes
Hour FE	Yes	Yes	Yes
Log-Likelihood		-52,381.77	-52,131.71
AIC		104,785.5	104,285.4
BIC		104,917.6	104,417.4
Observations	1,204,896	1,204,896	1,204,896

***, ** and *: indicate 1%, 5% and 10% significance levels, respectively. Robust clustered standard errors are in parentheses.

Table 5: Estimation Results for the Hourly Frequency of Negative Electricity Price Occurrence

Dependent Variable: $Fr(N)_{iT}$

Estimator	**OLS**	**Poisson**	**Generalized NB**	**NB**
Variables	**(1)**	**(2)**	**(3)**	**(4)**
C_{it}	0.000340	0.000740	0.000828	0.000749
	(0.00108)	(0.00175)	(0.00254)	(0.000543)
L_{it}	-0.00598	-0.0240***	-0.170	-0.0257***
	(0.00461)	(0.00645)	(0.198)	(0.00397)
$Wind_{it}$	0.0471	0.301***	0.308***	0.224***
	(0.0309)	(0.0356)	(0.0873)	(0.0395)
Sol_{it}	-0.0279	-0.714***	-0.772***	-0.773***
	(0.0263)	(0.147)	(0.141)	(0.0582)
$Hydro_{it}$	-0.00141	-0.00556	0.0387	-0.0222
	(0.00507)	(0.0259)	(0.0311)	(0.0173)
T_{it}	0.00849*	0.0538***	0.0506***	0.0565***
	(0.00409)	(0.00874)	(0.00743)	(0.00563)
H_{it}	-0.00811*	-0.0803***	-0.0824***	-0.0922***
	(0.00442)	(0.0182)	(0.0209)	(0.0126)
Constant	0.0645	-3.361***	-2.594***	-1.787***
	(0.0717)	(0.626)	(0.611)	(0.260)
Zone FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes
Day FE	Yes	Yes	Yes	Yes
Day of Week FE	Yes	Yes	Yes	Yes
Hour FE	Yes	Yes	Yes	Yes
Log-Likelihood		-454,920.6	-279,104.1	-278,939.3
AIC		909,861.1	558,228.1	557,898.7
BIC		909,981.1	558,348.1	558,018.7
Observations	1,204,896	1,204,896	1,204,896	1,204,896

***, ** and *: indicate 1%, 5% and 10% significance levels, respectively. Robust clustered standard errors are in parentheses. Negative Binomial is abbreviated as NB.

Table 6: Estimation Results for Daily Frequency of Negative Electricity Price Occurrence
Dependent Variable: $Fr(N)_{iT}$

Estimator	**OLS**	**Poisson**	**Generalized NB**	**NB**
Variables	**(1)**	**(2)**	**(3)**	**(4)**
C_{it}	0.00174	0.0111	0.00378	0.00346
	(0.00335)	(0.0182)	(0.00591)	(0.00547)
L_{it}	-0.000462	-0.112	-0.201	-0.215
	(0.0173)	(0.108)	(0.151)	(0.132)
$Wind_{it}$	0.0932	0.331***	0.285***	0.259***
	(0.0623)	(0.0565)	(0.0979)	(0.0682)
Sol_{it}	-0.568	-1.725***	-1.093**	-1.181***
	(0.436)	(0.417)	(0.476)	(0.393)
$Hydro_{it}$	-0.0256	0.0983	0.106	0.0526
	(0.0621)	(0.110)	(0.113)	(0.0652)
T_{it}	0.0193**	0.0735***	0.0781***	0.0748***
	(0.00844)	(0.0161)	(0.0168)	(0.0167)
H_{it}	-0.0273*	-0.133***	-0.141***	-0.129***
	(0.0129)	(0.0234)	(0.0408)	(0.0401)
Constant	0.0360	-2.590***	-1.780**	-1.679***
	(0.162)	(0.524)	(0.796)	(0.507)
Zone FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes
Day FE	Yes	Yes	Yes	Yes
Day of Week FE	Yes	Yes	Yes	Yes
Log-Likelihood		-27004.83	-20948.37	-20874.69
AIC		54029.66	41918.73	41769.38
BIC		54117.9	42015.8	41857.62
Observations	50,204	50,204	50,204	50,204

***, ** and *: indicate 1%, 5% and 10% significance levels, respectively. Robust clustered standard errors are in parentheses. Negative Binomial is abbreviated as NB.

5.3. *Economic Significance and Ranking of Effects*

Table 7 displays the average marginal effects, which are interpreted along with their associated significance levels to assess the economic importance of each driver. These marginal effects allow ranking of the relative impacts of different variables on negative price occurrences. The marginal effects are measured in percentage point changes in the probability or expected count of negative price events per unit change in each explanatory variable. For wind speed (measured in m/s), the 5-minute probability model shows that a 1 m/s increase in wind speed increases the probability of negative prices by 0.332 percentage points. For solar irradiance (measured in kWh/m²), a 1 kWh/m² increase decreases the probability by 0.824 percentage points. For temperature (measured in °C), a 1°C increase raises the probability by 0.062 percentage points. These magnitudes indicate economically meaningful effects, with solar energy showing the largest impact in absolute terms.

Based on the magnitude and consistency of effects across models, the following ranking of impacts is established:[12]

1. *Solar Energy*: Demonstrates the largest marginal effect in absolute terms, with a significant negative impact on both the probability and frequency of negative prices. Its effect is particularly pronounced in the daily frequency model, suggesting a cumulative impact over longer time horizons. This supports Empirical Prediction 1 that solar generation reduces negative price likelihood.
2. *Wind Energy*: Shows the second-largest impact, with a consistent positive effect on negative price occurrences. Its influence is robust across all time scales and model specifications, demonstrating the important role of wind power in driving negative price occurrences. This confirms the theoretical prediction about wind generation's positive relationship with negative pricing.
3. *Humidity*: Exhibits the third-largest impact, with a consistent negative effect on negative price occurrences. Its influence is significant across all models, highlighting the importance of weather conditions in shaping electricity price dynamics and supporting Empirical Prediction 2.

[12]This hierarchical relationship remains robust to alternative specifications. Appendix D demonstrates the stability of these results when controlling for system load across different temporal aggregations, while Appendix E shows their robustness when accounting for potential interactions between renewable sources, weather conditions, and grid constraints.

4. *Temperature*: Ranks fourth in impact, showing a consistent positive effect on negative price occurrences. While its marginal effect is smaller than humidity, it remains significant across all specifications, highlighting the role of weather in electricity markets and confirming the dual weather effects predicted theoretically.
5. *Marginal Costs of Losses*: Shows a significant negative effect in most models, particularly at shorter time scales (5-minute and hourly data). However, its impact diminishes in the daily frequency models, suggesting its relevance is primarily in short-term price dynamics rather than the longer-term effects emphasized in Empirical Prediction 3.
6. *Precipitation*: Demonstrates the smallest significant impact, showing marginal significance only in the hourly probability model. This suggests that hydropower potential, as proxied by precipitation, plays a limited role in driving negative price occurrences.
7. *Marginal Costs of Congestion*: Consistently shows the smallest and statistically insignificant effects across all models, indicating that grid congestion may not be a primary driver of negative price occurrences in the NYISO market during the study period. This result challenges conventional views about transmission constraints being an important driver of negative pricing.

This ranking confirms the hierarchy predicted by the theoretical framework and provides empirical support for the empirical predictions. The primary role of renewable energy sources, particularly solar and wind, in shaping negative price dynamics validates the renewable integration effects emphasized in the theoretical model. Weather factors also play an important role, while grid condition indicators show more limited impacts than commonly assumed in policy discussions. The varying impacts across different time scales highlight the temporal dynamics at play in electricity markets and suggest that policy interventions should account for these different time horizons.

6. Policy Implications

The theoretical framework and empirical results establish three key findings that inform policy design during the renewable energy transition. First, the theoretical analysis demonstrates that negative prices can achieve welfare-maximizing allocations under operational constraints and production subsidies, suggesting these price signals serve an efficient economic function rather than representing market failures requiring regulatory suppression. Second, the empirical analysis establishes a clear hierarchy of drivers: renewable energy integration emerges as

Table 7: Average Marginal Effects

C_{it}	L_{it}	$Wind_{it}$	Sol_{it}	$Hydro_{it}$	T_{it}	H_{it}
Probability of 5-Minute Occurrence - Logit Model						
0.0000684	-0.000938*	0.00332***	-0.00824***	-0.0000322	0.000622***	-0.000942***
Probability of Hourly Occurrence - Logit Model						
0.0000844	-0.00115*	0.00274***	-0.00868***	0.000724*	0.000690***	-.00120***
Hourly Frequency of Occurrence- NB Model						
0.0000983	-0.00338***	0.0293***	-0.101***	-0.00292	0.00741***	-0.0121***
Daily Frequency of Occurrence - NB Model						
0.000854	-0.0531	0.0639***	-0.292***	0.0130	0.0185***	-0.0319***

***, ** and *: indicate 1%, 5% and 10% significance levels, respectively. Negative Binomial is abbreviated as NB.

the primary factor influencing negative price formation, weather conditions rank second in importance, while grid constraints demonstrate minimal influence contrary to their prominence in policy discussions. Third, within renewable technologies, important heterogeneity exists, with solar generation reducing negative price occurrences while wind generation increases them, reflecting their distinct production patterns and subsidy structures. These findings offer evidence-based guidance for policy design that challenges conventional assumptions about managing negative electricity prices.

The finding that negative prices can achieve welfare-maximizing allocations under operational constraints and production subsidies suggests that market design may benefit from preserving these price signals rather than suppressing them. This perspective finds support in Halbrügge et al. (2024), who demonstrate that negative pricing contributes to increased system efficiency, and Korpås and Botterud (2020), who prove analytically that allowing negative prices achieves optimal investment portfolios. While preserving the efficiency benefits of negative price signals, policymakers can simultaneously work to address the underlying drivers that create excessive negative pricing frequency, which can undermine investment signals and market stability. Regulatory frameworks could therefore consider eliminating price floors and pro-rata curtailment mechanisms that may distort the efficient resource allocation demonstrated here.

The empirical hierarchy established in this analysis—renewable energy integration as the primary driver, weather conditions as secondary, and grid constraints as minimal—provides an evidence-based framework for policy prioritization that challenges conventional assumptions about negative pricing drivers. This hierarchy suggests that policy resources should be allocated according to demonstrated impact rather than conventional wisdom. The contrasting

technology-specific effects observed—with solar energy reducing and wind energy increasing negative price occurrences—indicate that differentiated renewable policies may prove more effective than uniform approaches. This finding aligns with Lehmann and Söderholm (2018), who argues that technology-specific support may be justified due to different learning curves and varying externalities by technology type.

These technology-specific findings have important precedent in policy practice, with Jerrentrup et al. (2019) documenting that 16 of 18 EU Member States used technology-discriminatory schemes despite neutrality requirements. This precedent supports differentiated approaches, and the findings suggest solar deployment merits particular attention given its unique market effects. The significant negative relationship between solar generation and negative price formation indicates that enhanced solar deployment policies could reduce negative pricing frequency while advancing environmental objectives. The observed alignment between solar generation patterns and demand cycles suggests that solar-specific incentives might simultaneously address market efficiency and renewable energy goals. Investment incentives and streamlined permitting for solar deployment merit consideration given these dual benefits, though broader system impacts including grid integration costs require careful evaluation.

In contrast, the positive relationship between wind generation and negative price occurrences suggests that current wind energy subsidies merit policy reconsideration. The results indicate that production-based subsidies enable wind generation during periods when market signals suggest reduced output would be economically efficient. This finding resonates with Mwampashi et al. (2021), who document how production tax credits enable wind generators to operate profitably at negative $35/MWh. Policy reforms might consider transitioning toward capacity payments or investment credits that decouple financial rewards from real-time generation decisions. The success that Beiter et al. (2024) document with contracts for difference—supporting over 50% of global offshore wind supply—offers a potential model for such reforms.

The identification of weather conditions as the second-most important driver suggests that incorporating enhanced meteorological forecasting into market operations could improve efficiency. The observed relationships between atmospheric conditions and negative pricing reveal specific patterns. Temperature increases negative price likelihood while humidity decreases it. These findings indicate that weather-responsive market mechanisms and demand response programs might reduce negative price frequency. Evidence from Tawn and Browell (2022)

showing 22% revenue increases from improved forecasting supports this recommendation for targeting the secondary driver. Seasonal rate structures reflecting weather-driven patterns in renewable generation and electricity demand throughout the year could complement forecasting improvements.

The finding that grid constraints show minimal influence on negative price formation represents a significant challenge to conventional policy assumptions about transmission expansion as the primary solution to negative pricing. This result indicates that resources currently directed toward transmission investments might achieve greater negative price reduction through alternative approaches. These alternatives should target the primary and secondary drivers rather than the minimal grid constraint effects observed here. The cost considerations that LaRiviere and Lyu (2022) document—transmission capital costs ranging from $1-10/MWh—support redirecting resources toward demand-side flexibility and storage applications that address renewable integration and weather effects.

Policy interventions should prioritize the drivers with greatest demonstrated impact, beginning with renewable energy integration through technology-specific policies emphasizing solar deployment incentives and wind subsidy reform. Addressing weather responsiveness through enhanced forecasting and demand response programs represents the logical second priority given its ranking as the secondary driver. Finally, attention could focus on storage deployment and demand flexibility rather than transmission expansion, given the minimal grid constraint effects observed. However, this impact-based prioritization would benefit from adaptation to local institutional capabilities and energy system characteristics. Successful implementation likely requires stakeholder education about the efficiency role of negative prices and coordination across policy domains to ensure resources target the mechanisms with demonstrated impact while maintaining system reliability during the renewable energy transition.

7. Conclusion

This paper examines the economic drivers of negative electricity prices in renewable-integrated markets, addressing which factors most significantly influence negative price formation and their relative importance for policy design. The analysis begins by developing a formal theoretical framework that models generator optimization under operational constraints and production subsidies, establishing the conditions under which negative prices emerge and proving that they can achieve welfare-maximizing allocations. These theoretical

predictions are then tested empirically using over 14 million observations from New York's wholesale electricity markets spanning 2010-2022. The empirical analysis confirms the theoretical framework, revealing a clear hierarchy: renewable energy integration is the primary driver of negative pricing, weather conditions rank second, and grid constraints show minimal influence. Within renewable technologies, solar generation reduces negative price occurrences while wind generation increases them, reflecting their distinct production patterns and subsidy structures.

These findings lead directly to policy implications that challenge conventional approaches to managing electricity markets during the renewable energy transition. The evidence indicates that transmission expansion—the conventional policy response—appears less effective than alternative strategies. Technology-specific renewable policies emerge as the most promising approach: leveraging solar energy's beneficial market effects while reforming wind energy subsidies that enable economically inefficient generation. Weather-responsive market mechanisms and enhanced forecasting capabilities represent secondary but useful interventions. The analysis suggests that preserving negative price signals while addressing their underlying drivers offers a more efficient approach than suppressing these prices through regulatory constraints.

Several limitations point toward important directions for future research. The analysis focuses on a single market during a specific period, limiting generalizability to different systems and market structures. Immediate extensions could examine how emerging technologies—particularly battery storage and offshore wind—alter these negative pricing dynamics. Building on such technology-specific insights, research investigating the optimal frequency of negative prices and evaluating the policy interventions recommended here would inform market design decisions. As renewable penetration increases globally, understanding these market dynamics becomes increasingly critical for efficient energy system operation during the clean energy transition.

References

3Degrees. Europe's Energy Market: Navigating Negative Prices and Geopolitical Shifts in H2 2024. Technical report, 2024. URL https://3degreesinc.com/insights/europes-energy-market-negative-prices-and-geopolitical-shifts-in-h2-2024/.

A. Abadie, S. Athey, G. W. Imbens, and J. M. Wooldridge. When Should You Adjust Standard Errors for Clustering? *The Quarterly Journal of Economics*, 138(1):1–35, 2023.

J. Abrell, M. Kosch, and S. Rausch. Carbon Abatement with Renewables: Evaluating Wind and Solar Subsidies in Germany and Spain. *Journal of Public Economics*, 169:172–202, 2019.

W. Antweiler and F. Muesgens. On the Long-Term Merit Order Effect of Renewable Energies. *Energy Economics*, 99:105275, 2021.

P. Beiter, J. Guillet, M. Jansen, E. Wilson, and L. Kitzing. The Enduring Role of Contracts for Difference in Risk Management and Market Creation for Renewables. *Nature Energy*, 9(1): 20–26, 2024.

BloombergNEF. Global Investment in the Energy Transition Exceeded $2 Trillion for the First Time in 2024. Technical report, Bloomberg New Energy Finance, 2025. URL https://about.bnef.com/blog/global-investment-in-the-energy-transition-exceeded-2-trillion-for-the-first-time-in-2024-according-to-bloombergnef-report/.

W. Bönte, S. Nielen, N. Valitov, and T. Engelmeyer. Price Elasticity of Demand in the EPEX Spot Market for Electricity—New Empirical Evidence. *Economics Letters*, 135:5–8, 2015.

J. Bushnell and K. Novan. Setting with the Sun: The Impacts of Renewable Energy on Wholesale Power Markets. Technical report, National Bureau of Economic Research, 2018.

Y. Chu, M. Li, C. F. Coimbra, D. Feng, and H. Wang. Intra-Hour Irradiance Forecasting Techniques for Solar Power Integration: A Review. *iScience*, 24(10), 2021.

P. Cramton. Competitive Bidding Behavior in Uniform-Price Auction Markets. In *Annual Hawaii International Conference on System Sciences (HICSS)*, pages 11–pp. IEEE, 2004.

European Commission. Electricity Market Design Reform. Technical report, European Commission, December 2023. URL https://energy.ec.europa.eu/topics/markets-and-consumers/electricity-market-design_en.

N. Farrell. The Increasing Cost of Ignoring Coase: Inefficient Electricity Tariffs, Welfare Loss and Welfare-Reducing Technological Change. *Energy Economics*, 97:104848, 2021.

F. Genoese, M. Genoese, and M. Wietschel. Occurrence of Negative Prices on the German Spot Market for Electricity and Their Influence on Balancing Power Markets. In *2010 7th International Conference on the European Energy Market*, pages 1–6. IEEE, 2010.

gridX. What Is Negative Energy Price in Europe. Technical report, 2024. URL https://www.gridx.ai/knowledge/what-is-negative-energy-price-in-europe.

S. Halbrügge, P. Heess, P. Schott, and M. Weibelzahl. Negative Electricity Prices as a Signal for Lacking Flexibility? On the Effects of Demand Flexibility on Electricity Prices. *International Journal of Energy Sector Management*, 18(3):596–616, 2024.

L. Hirth, T. M. Khanna, and O. Ruhnau. How Aggregate Electricity Demand Responds to Hourly Wholesale Price Fluctuations. *Energy Economics*, 135:107652, 2024.

IRC. The Independent System Operator (ISO) and Regional Transmission Operator (RTO) Markets. https://isorto.org/wp-content/uploads/2018/05/20170905_2017IRCMarketsCommitteeExecutiveSummaryFinal.pdf, 2017.

IRENA. Renewable Capacity Statistics 2025. Technical report, IRENA, March 2025. URL https://www.irena.org/Publications/2025/Mar/Renewable-capacity-statistics-2025.

M. Jerrentrup, M.-T. Lotz, S. Tiedemann, and L. Hirth. Technology-Neutral Auctions for Renewable Energy Support – More or Less Discriminatory? Technical report, Neon Neue Energieökonomik, 2019. URL https://neon.energy/Jerrentrup-Lotz-Tiedemann-Hirth-2019-Tech-neutral-auctions.pdf.

P. Joskow. Patterns of Transmission Investments. *Competitive Electricity Markets and Sustainability*, pages 131–186, 2006.

W. Katzenstein, E. Fertig, and J. Apt. The Variability of Interconnected Wind Plants. *Energy Policy*, 38(8):4400–4410, 2010.

J. C. Ketterer. The Impact of Wind Power Generation on the Electricity Price in Germany. *Energy Economics*, 44:270–280, 2014.

A. Khojaste, G. Pritchard, and G. Zakeri. Quantile Fourier Regressions for Decision Making under Uncertainty. *Annals of Operations Research*, pages 1–16, 2024.

A. Khojaste, J. Pearce, G. Zakeri, and Y. Sang. Electricity Price-Aware Scheduling of Data Center Cooling. *arXiv preprint arXiv:2508.03160*, 2025.

M. Korpås and A. Botterud. Optimality Conditions and Cost Recovery in Electricity Markets with VRE and Energy Storage. Technical report, MIT Center for Energy and Environmental Policy Research, 2020. URL https://ceepr.mit.edu/optimality-conditions-and-cost-recovery-in-electricity-markets-with-vre-and-energy-storage/.

J. LaRiviere and X. Lyu. Transmission Constraints, Intermittent Renewables and Welfare. *Journal of Environmental Economics and Management*, 112:102618, 2022.

P. Lehmann and P. Söderholm. Can Technology-Specific Deployment Policies Be Cost-Effective? The Case of Renewable Energy Support Schemes. *Environmental and Resource Economics*, 71 (2):475–505, 2018.

S. Mosquera-López, J. M. Uribe, and O. Joaqui-Barandica. Weather Conditions, Climate Change, and the Price of Electricity. *Energy Economics*, 137:107789, 2024.

M. M. Mwampashi, C. S. Nikitopoulos, O. Konstandatos, and A. Rai. Wind Generation and the Dynamics of Electricity Prices in Australia. *Energy Economics*, 103:105547, 2021.

D. M. Newbery. High Renewable Electricity Penetration: Marginal Curtailment and Market Failure Under "Subsidy-Free" Entry. *Energy Economics*, 126:107011, 2023.

M. Nicolosi. Wind Power Integration and Power System Flexibility–An Empirical Analysis of Extreme Events in Germany Under the New Negative Price Regime. *Energy Policy*, 38(11): 7257–7268, 2010.

NYISO. Introduction to the NYISO. https://www.nyiso.com/documents/20142/1392242/Introduction_to_the_NYISO.pdf/d027e637-20bf-2c9f-b3b6-43ce348a7595?t=1539227033083, 2011.

NYISO. Load & Capacity Data. https://www.nyiso.com/documents/20142/2226333/2022-Gold-Book-Final-Public.pdf, 2022a.

NYISO. The Path to a Reliable, Greener Grid for New York. https://www.nyiso.com/documents/20142/2223020/2022-Power-Trends-Report.pdf/d1f9eca5-b278-c445-2f3f-edd959611903?t=1654689893527, 2022b.

J. Pearce, A. Khojaste, G. Zakeri, and G. Pritchard. Baseline Hydropower Generation Offer Curves. *arXiv preprint arXiv:2508.04854*, 2025.

H. Schermeyer, C. Vergara, and W. Fichtner. Renewable Energy Curtailment: A Case Study on Today's and Tomorrow's Congestion Management. *Energy Policy*, 112:427–436, 2018.

J. Seel, D. Millstein, A. Mills, M. Bolinger, and R. Wiser. Plentiful Electricity Turns Wholesale Prices Negative. *Advances in Applied Energy*, 4:100073, 2021.

smartEn and LCP Delta. 2024 Market Monitor for Demand Side Flexibility. Technical report, Smart Energy Europe, 2024. URL https://smarten.eu/reports/2024-market-monitor-for-demand-side-flexibility/.

I. Staffell and S. Pfenninger. Using Bias-Corrected Reanalysis to Simulate Current and Future Wind Power Output. *Energy*, 114:1224–1239, 2016.

I. Staffell and S. Pfenninger. The Increasing Impact of Weather on Electricity Supply and Demand. *Energy*, 145:65–78, 2018.

R. Tawn and J. Browell. A Review of Very Short-Term Wind and Solar Power Forecasting. *Renewable and Sustainable Energy Reviews*, 153:111758, 2022.

Utility Dive. California's Solar, Wind Curtailment Jumped 29% in 2024: EIA. Technical report, 2025. URL https://www.utilitydive.com/news/solar-wind-curtailments-increasing-california-caiso/749420/.

E. Valor, V. Meneu, and V. Caselles. Daily Air Temperature and Electricity Load in Spain. *Journal of Applied Meteorology*, 40(8):1413–1421, 2001.

C.-K. Woo, I. Horowitz, J. Moore, and A. Pacheco. The Impact of Wind Generation on the Electricity Spot-Market Price Level and Variance: The Texas Experience. *Energy Policy*, 39(7): 3939–3944, 2011.

C.-K. Woo, J. Moore, B. Schneiderman, T. Ho, A. Olson, L. Alagappan, K. Chawla, N. Toyama, and J. Zarnikau. Merit-Order Effects of Renewable Energy and Price Divergence in California's Day-Ahead and Real-Time Electricity Markets. *Energy Policy*, 92:299–312, 2016.

Z. Zhu, S. Bu, K. W. Chan, F. Li, Y. Ye, C. Y. Chung, and G. Strbac. Designing the Future Electricity Spot Market with High Renewables via Reliable Simulations. *Nature Reviews Electrical Engineering*, pages 1–18, 2025.

Appendix A Proofs

Proof of Lemma 1. The generator's optimization problem from Equation 5 is:

$$\max_{q_{git},u_{git}} \sum_{t\in\mathcal{T}} \beta^t [p_{it}q_{git} - A_g(q_{git}) - \Phi_{git}(q_{git}, q_{g,i,t-1}, u_{git}) - \eta_g u_{git} + SUB_{git}(q_{git})] \tag{A.1}$$

subject to constraints 3-4.

The Lagrangian for this constrained optimization problem is formulated as:

$$\begin{aligned}\mathcal{L} = &\sum_{t\in\mathcal{T}} \beta^t [p_{it}q_{git} - A_g(q_{git}) - \Phi_{git}(q_{git}, q_{g,i,t-1}, u_{git}) - \eta_g u_{git} + SUB_{git}(q_{git})] \\ &+ \lambda^{min}_{git}(\underline{q}_g u_{git} - q_{git}) + \lambda^{max}_{git}(q_{git} - \overline{q}_g u_{git}) \\ &+ \mu^{down}_{git}(-\underline{r}_g - q_{git} + q_{g,i,t-1}) + \mu^{up}_{git}(q_{git} - q_{g,i,t-1} - \overline{r}_g)\end{aligned} \tag{A.2}$$

Taking the first-order condition with respect to q_{git}:

$$\frac{\partial \mathcal{L}}{\partial q_{git}} = \beta^t p_{it} - \beta^t \frac{\partial A_g(q_{git})}{\partial q_{git}} - \beta^t \frac{\partial \Phi_{git}}{\partial q_{git}} + \beta^t \frac{\partial SUB_{git}}{\partial q_{git}} - \lambda^{min}_{git} + \lambda^{max}_{git} - \mu^{down}_{git} + \mu^{up}_{git} = 0 \tag{A.3}$$

For electricity markets operating at short time scales, setting $\beta = 1$ is standard practice since intra-day discounting effects are minimal. Rearranging the first-order condition:

$$p_{it} = \frac{\partial A_g(q_{git})}{\partial q_{git}} + \frac{\partial \Phi_{git}}{\partial q_{git}} - \frac{\partial SUB_{git}}{\partial q_{git}} + (\lambda^{min}_{git} - \lambda^{max}_{git} + \mu^{down}_{git} - \mu^{up}_{git}) \tag{A.4}$$

The optimal bid price b^*_{git} reflects the generator's marginal willingness to produce, accounting for both direct costs and the shadow values of operational constraints:

$$b^*_{git} = \frac{\partial A_g(q_{git})}{\partial q_{git}} + \frac{\partial \Phi_{git}}{\partial q_{git}} - \frac{\partial SUB_{git}}{\partial q_{git}} - (\lambda^{min}_{git} - \lambda^{max}_{git} + \mu^{down}_{git} - \mu^{up}_{git}) \tag{A.5}$$

This gives Equation 6. The non-negativity of all Lagrange multipliers follows from the inequality constraints in the optimization problem and the Karush-Kuhn-Tucker conditions. □

Proof of Proposition 1. From Lemma 1, the bid price is negative ($b^*_{git} < 0$) if and only if:

$$\frac{\partial A_g(q_{git})}{\partial q_{git}} + \frac{\partial \Phi_{git}}{\partial q_{git}} - \frac{\partial SUB_{git}}{\partial q_{git}} - (\lambda^{min}_{git} - \lambda^{max}_{git} + \mu^{down}_{git} - \mu^{up}_{git}) < 0 \tag{A.6}$$

When considering the economic conditions under which generators submit negative bids, the following complementary slackness conditions are relevant:

- If $q_{git} = \underline{q}_g u_{git}$ (minimum generation constraint binds), then $\lambda^{min}_{git} \geq 0$
- If $q_{git} < \overline{q}_g u_{git}$ (maximum generation constraint doesn't bind), then $\lambda^{max}_{git} = 0$
- If $q_{git} - q_{g,i,t-1} > -\underline{r}_g$ (down-ramping constraint doesn't bind), then $\mu^{down}_{git} = 0$
- If $q_{git} - q_{g,i,t-1} < \overline{r}_g$ (up-ramping constraint doesn't bind), then $\mu^{up}_{git} = 0$

For a generator operating at minimum output levels or facing binding operational constraints, typically $\lambda^{max}_{git} = 0$ and $\mu^{up}_{git} = 0$. Under these conditions, substituting these values and rearranging terms yields the condition in Equation 7:

$$\frac{\partial SUB_{git}}{\partial q_{git}} + \lambda^{min}_{git} + \mu^{down}_{git} > \frac{\partial A_g(q_{git})}{\partial q_{git}} + \frac{\partial \Phi_{git}}{\partial q_{git}} \tag{A.7}$$

This condition demonstrates that negative bidding occurs when the combined value of subsidies and constraint shadow prices exceeds the marginal cost of production and cycling. □

Proof of Corollary 1. From Proposition 1, negative bids require:

$$\frac{\partial SUB_{git}}{\partial q_{git}} + \lambda^{min}_{git} + \mu^{down}_{git} > \frac{\partial A_g(q_{git})}{\partial q_{git}} + \frac{\partial \Phi_{git}}{\partial q_{git}} \tag{A.8}$$

For baseload generators, $SUB_{git} = 0$ implies $\partial SUB_{git}/\partial q_{git} = 0$. Therefore, negative bids require $\lambda^{min}_{git} + \mu^{down}_{git} > 0$, which occurs when at least one operational constraint is binding. Most commonly, this involves minimum generation constraints ($\lambda^{min}_{git} > 0$) during low demand periods when maintaining minimum output is technically necessary but economically costly due to low market prices.

For renewable generators, from Equation 2, $\partial SUB_{git}/\partial q_{git} = s_g > 0$ for generators receiving production-based subsidies. This positive subsidy term can make the left side of the inequality exceed marginal costs even when operational constraints are non-binding ($\lambda^{min}_{git} = \mu^{down}_{git} = 0$), enabling negative bids solely through the subsidy mechanism. This allows renewable generators to profitably operate even when market prices are negative, as long as the absolute value of negative prices does not exceed their subsidy rates.

These two mechanisms—constraint-driven negative bidding by conventional generators and subsidy-enabled negative bidding by renewables—represent the primary economic pathways through which negative prices emerge in wholesale electricity markets. □

Proof of Proposition 2. The social planner's problem that maximizes total welfare in Equation 9 is formulated as:

$$\max_{q_{git},u_{git}} \sum_{i\in\Omega}\sum_{t\in\mathcal{T}}\left[\int_0^{Q_{it}} P_{it}(q)dq - \sum_{g\in\Omega_B^i\cup\Omega_R^i} TC_g(q_{git}) - \sum_{\ell}\text{Losses}_\ell\right] \tag{A.9}$$

subject to:

$$\sum_{g\in\Omega_B^i\cup\Omega_R^i} q_{git} = D_{it}(p_{it}) \qquad \forall i,t \quad (\pi_{it}) \tag{A.10}$$

$$|F_{ijt}| \le \overline{F}_{ijt} \qquad \forall i,j,t \quad (\nu_{ijt}) \tag{A.11}$$

plus constraints 3-4, where π_{it} and ν_{ijt} are the respective Lagrange multipliers.

The first-order conditions with respect to q_{git} yield:

$$-\frac{\partial TC_g(q_{git})}{\partial q_{git}} - \frac{\partial \text{Losses}_\ell}{\partial q_{git}} + \pi_{it} + \sum_j \nu_{ijt}\cdot \text{PTDF}_{ij} = 0 \tag{A.12}$$

where PTDF_{ij} represents the power transfer distribution factor between zones i and j. Rearranging:

$$\pi_{it} + \sum_j \nu_{ijt}\cdot \text{PTDF}_{ij} = \frac{\partial TC_g(q_{git})}{\partial q_{git}} + \frac{\partial \text{Losses}_\ell}{\partial q_{git}} = MC_g(q_{git}) + L_{it} \tag{A.13}$$

From consumer optimization in zone i, the marginal utility of consumption equals the zone price: $P_{it}(Q_{it}) = \pi_{it}$ in the absence of transmission constraints. With transmission constraints:

$$P_{it}(Q_{it}) = \pi_{it} + \sum_j \nu_{ijt}\text{PTDF}_{ij} = p_t^{sys} + C_{it} + L_{it} = p_{it}^* \tag{A.14}$$

This establishes that the market clearing price equals the marginal cost of the price-setting generator, adjusted for congestion and losses, supporting the welfare-maximizing allocation characterized in Equation 12.

With negative prices allowed, no additional constraints restrict the optimization. By the First Welfare Theorem, the competitive market equilibrium with negative price allowance achieves the social optimum. Under a price floor constraint $p_{it} \geq 0$, the constrained optimum yields strictly lower welfare. The welfare loss equals:

$$\Delta W = \sum_{i,t} \left[\int_{Q_{it}^{floor}}^{Q_{it}^{*}} P_{it}(q)dq - \sum_{g} \int_{q_{git}^{floor}}^{q_{git}^{*}} MC_g(q)dq \right] \tag{A.15}$$

where superscripts $floor$ and $*$ denote quantities under price floor and efficient allocation respectively. This difference corresponds to area A_3 in Figure 2.

Production efficiency follows from the first-order conditions, which ensure that marginal costs are equated across all operating generators within transmission constraints. The market clearing price p_{it}^{*} equals the marginal cost of the marginal generator adjusted for congestion and losses as in Equation 10, demonstrating that negative price allowance preserves the efficiency properties of competitive electricity markets. □

Proof of Empirical Prediction 1. From Equations 2 and 6, renewable generator g's optimal bid includes the term $\partial SUB_{git}/\partial q_{git} = s_r$. The probability of negative prices is:

$$Pr(N_{it} = 1) = Pr(p_{it}^{*} < 0) = F(\mathbf{X}_{it}^{\top}\boldsymbol{\beta}) \tag{A.16}$$

where $F(\cdot)$ is a cumulative distribution function and $\mathbf{X}_{it} = (R_{r,it}, W_{it}, G_{it})$ includes renewable generation from source $r \in \{Sol, Wind, Hydro\}$, weather conditions, and grid constraints as defined in the theoretical section.

Taking the derivative with respect to renewable generation $R_{r,it}$ and applying the chain rule:

$$\frac{\partial Pr(N_{it} = 1)}{\partial R_{r,it}} = f(\mathbf{X}_{it}^{\top}\boldsymbol{\beta}) \frac{\partial(\mathbf{X}_{it}^{\top}\boldsymbol{\beta})}{\partial R_{r,it}} \tag{A.17}$$

where $f(\cdot)$ is the probability density function.

From the optimal bid condition and its impact on market clearing prices, the effect of renewable generation on the probability of negative prices operates through two channels: the subsidy effect (which tends to increase negative bidding) and the marginal cost effect (which

represents the direct production cost). Thus:

$$\frac{\partial(\mathbf{X}_{it}^{\top}\boldsymbol{\beta})}{\partial R_{r,it}} = \frac{\partial s_r}{\partial R_{r,it}} - \frac{\partial MC_r}{\partial R_{r,it}} \tag{A.18}$$

This gives Equation 14. While subsidy rates and marginal costs operate at the generator level, their combined effect influences market-clearing prices, which are captured through observed renewable generation quantities $R_{r,it} \in \{Sol_{it}, Wind_{it}, Hydro_{it}\}$ in the empirical analysis. For renewables with near-zero marginal costs, this effect reflects primarily the subsidy mechanism, as increased renewable generation with production-based subsidies tends to increase the likelihood of negative market-clearing prices. □

Proof of Empirical Prediction 2. Weather conditions affect both supply and demand sides of the market. On the supply side, weather variables influence renewable generation potential through the functional relationships between weather conditions and renewable output:

$$\frac{\partial Q_{it}^{S}}{\partial w_{it}} = \sum_{r \in \{Sol, Wind, Hydro\}} \frac{\partial R_{r,it}}{\partial w_{it}} \tag{A.19}$$

where $R_{r,it} \in \{Sol_{it}, Wind_{it}, Hydro_{it}\}$ represent solar, wind, and hydro generation respectively, and their derivatives capture how weather conditions affect output (e.g., wind speed affects wind generation, solar irradiance affects solar generation).

On the demand side, weather affects consumption through heating and cooling requirements:

$$\frac{\partial D_{it}}{\partial w_{it}} = \frac{\partial D_{it}^{H}}{\partial w_{it}} + \frac{\partial D_{it}^{C}}{\partial w_{it}} \tag{A.20}$$

where D_{it}^{H} and D_{it}^{C} represent heating and cooling demand components respectively.

Applying the chain rule to the probability of negative prices:

$$\frac{\partial Pr(N_{it} = 1)}{\partial w_{it}} = f(\mathbf{X}_{it}^{\top}\boldsymbol{\beta}) \cdot \left(\frac{\partial(\mathbf{X}_{it}^{\top}\boldsymbol{\beta})}{\partial Q_{it}^{S}} \cdot \frac{\partial Q_{it}^{S}}{\partial w_{it}} - \frac{\partial(\mathbf{X}_{it}^{\top}\boldsymbol{\beta})}{\partial D_{it}} \cdot \frac{\partial D_{it}}{\partial w_{it}} \right) \tag{A.21}$$

Since both supply and demand adjustments affect the likelihood of negative prices through their impact on supply-demand balance, this can be expressed as:

$$\frac{\partial Pr(N_{it} = 1)}{\partial w_{it}} = f(\mathbf{X}_{it}^{\top}\boldsymbol{\beta}) \cdot \left(\frac{\partial Q_{it}^{S}}{\partial w_{it}} - \frac{\partial D_{it}}{\partial w_{it}} \right) \tag{A.22}$$

This yields Equation 15, showing how weather variables affect negative price probability through their dual impact on supply and demand conditions. The sign of this effect depends on whether weather-induced supply changes dominate demand changes or vice versa. □

Proof of Empirical Prediction 3. From Equation 10, grid constraints affect locational marginal prices through:

$$p_{it}^{*} = p_{t}^{sys} + C_{it} + L_{it} = p_{t}^{sys} + \sum_{j} \nu_{ijt}(\overline{F}_{ijt} - F_{ijt}) + L_{it} \tag{A.23}$$

Taking the derivative with respect to a grid constraint measure g_{it}:

$$\frac{\partial p_{it}^{*}}{\partial g_{it}} = \frac{\partial p_{t}^{sys}}{\partial g_{it}} + \frac{\partial}{\partial g_{it}} \left[\sum_{j} \nu_{ijt}(\overline{F}_{ijt} - F_{ijt}) \right] + \frac{\partial L_{it}}{\partial g_{it}} \tag{A.24}$$

When g_{it} represents a binding transmission constraint, the primary effect operates through the congestion cost component. For a binding constraint between zones i and j, $\frac{\partial p_{it}^{*}}{\partial g_{it}} = \nu_{ijt}$, where the sign of ν_{ijt} depends on the direction of the constraint and power flow. This can either increase or decrease the nodal price depending on whether the zone is importing power (positive effect when constraint tightens) or exporting power (negative effect when constraint tightens).

Applying the chain rule to the probability of negative prices:

$$\frac{\partial Pr(N_{it} = 1)}{\partial g_{it}} = f(\mathbf{X}_{it}^{\top}\boldsymbol{\beta}) \cdot \frac{\partial(\mathbf{X}_{it}^{\top}\boldsymbol{\beta})}{\partial p_{it}^{*}} \cdot \frac{\partial p_{it}^{*}}{\partial g_{it}} = f(\mathbf{X}_{it}^{\top}\boldsymbol{\beta}) \cdot \nu_{ijt} \tag{A.25}$$

This establishes Equation 16, showing how grid constraints influence negative price probability through their impact on spatial price separation. The effect can be positive or negative depending on the zone's position relative to binding transmission constraints and the direction of power flows. □

Appendix B Descriptive Analysis of Negative Price Occurrence Patterns

This appendix provides comprehensive descriptive analysis of negative price occurrence patterns across temporal, spatial, and frequency dimensions. Table B.1 quantifies the frequency of negative price occurrences in the NYISO market during 2010-2022. Negative prices occur in 1.09% of 5-minute intervals (158,103 observations) and 0.96% of hourly periods (11,573 observations), demonstrating that while uncommon, these events occur with sufficient frequency to

warrant rigorous economic analysis. The distribution of negative price counts within hours and days shows that these events often cluster temporally, suggesting common underlying economic causes rather than random market anomalies. Specifically, among hours with at least one negative price event, 41,096 hours experienced 1-3 occurrences, 4,801 hours experienced 4-6 occurrences, and progressively fewer hours experienced higher frequencies, indicating concentration during particular market conditions.

Table B.1: Frequency of Negative Price Occurrences

Occurrence	**5-Minute Data**	**Hourly Data**
Negative Prices	158,103 (1.09%)	11,573 (0.96%)
Non-Negative Prices	14,300,649 (98.91%)	1,193,323 (99.04%)
Frequency	**Number of Hours**	**Number of Days**
0	1,151,414	45,328
1-3	41,096	4,046
4-6	4,801	493
7-9	2,863	185
10-12	4,722	83
13+	–	69

Figure B.1 reveals the economic patterns underlying negative price formation across multiple dimensions that connect directly to the theoretical predictions, particularly focusing on the temporal regularities that demonstrate the cyclostationary nature of electricity price dynamics and enable systematic identification of negative pricing episodes (Khojaste et al., 2025).

Spatial Distribution. The spatial distribution (top row of Figure B.1) shows notable heterogeneity across the 11 NYISO zones, with negative price frequencies ranging from less than 0.5% to over 2% depending on the zone. This spatial variation suggests that local economic factors—such as generation mix, transmission capacity, and load patterns—play important roles in price formation. Zones with higher renewable penetration or transmission constraints tend to exhibit higher negative price frequencies. This spatial heterogeneity provides identifying variation for the econometric analysis and aligns with the theoretical prediction that grid constraints create localized oversupply conditions.

Temporal Evolution. The annual trend (second row) reveals an upward trajectory from 2010 to 2022, with negative price frequency increasing from approximately 0.3% in early years to over 1.5% in recent years. This trend coincides precisely with New York's renewable energy expansion, particularly the growth of wind capacity from approximately 1,300 MW to over 2,000 MW and solar capacity from less than 50 MW to over 2,000 MW during this period, supporting the hypothesis about renewable integration effects on negative price formation.

Seasonal Patterns. The monthly distribution (third row) shows clear seasonal patterns, with spring months (April and May) displaying the highest negative price frequencies, exceeding 1.5% in both datasets. Summer and winter months show lower frequencies, while fall months exhibit intermediate levels. This pattern aligns with economic theory: spring conditions often combine high renewable generation potential (moderate temperatures, adequate precipitation, and good wind conditions) with relatively low electricity demand due to minimal heating and cooling needs, creating the supply-demand imbalances conducive to negative pricing.

Weekly and Daily Cycles. Daily and weekly cycles (fourth and fifth rows) demonstrate market operation effects that reflect demand-side factors in the theoretical framework. The calendar day pattern shows relatively uniform distribution across days of the month, suggesting day-of-month effects are minimal. However, the day-of-week pattern reveals that weekends show slightly higher negative price frequencies compared to weekdays. This pattern reflects the economic reality of reduced industrial demand during weekends while renewable generation continues based on weather conditions rather than demand patterns.

Diurnal Patterns. The hourly breakdown (sixth row) reveals a clear diurnal pattern, with peaks during early morning hours (4-6 AM exceeding 1.5%) and a secondary peak in the early afternoon (12-2 PM around 1.2%). These timing patterns correspond to periods when renewable generation potential (particularly wind and solar) may exceed local demand, creating the supply-demand imbalances that economic theory predicts will drive negative prices. The early morning peak reflects overnight wind generation during low demand periods, while the afternoon peak captures midday solar generation. These diurnal-seasonal regularities in electricity market dynamics reflect the cyclostationary nature of renewable generation and demand processes (Khojaste et al., 2024).

Intra-Hour Dynamics. The intra-hour variation captured in the 5-minute data (bottom left) provides additional insight into the rapid market dynamics that characterize modern electricity markets with high renewable penetration. The pattern shows relative uniformity across 5-minute intervals within hours, with slight variations suggesting that negative prices can occur at any point within the hour rather than being concentrated at specific dispatch intervals. The consistency of patterns between 5-minute and hourly datasets validates the dual-frequency approach while the 5-minute data captures more detailed variations in market behavior.

These descriptive patterns provide preliminary support for the theoretical predictions about the economic drivers of negative price formation in renewable-integrated electricity markets.

The spatial heterogeneity supports the role of local market conditions, the temporal evolution aligns with renewable capacity expansion, the seasonal and diurnal patterns reflect weather-driven supply-demand dynamics, and the clustering suggests systematic rather than random causes. These observations motivate the formal econometric analysis in the main text.

Appendix C Robustness Analysis: Alternative Fixed Effects Specifications

To assess the robustness of the findings and examine the sensitivity of results to various unobserved factors, a series of alternative fixed effects specifications is implemented. These specifications help address concerns about inter-zonal power flows, renewable capacity evolution, seasonal patterns, and broader market dynamics. Building upon the baseline specifications—the Logit model for probability and the Negative Binomial model for frequency analyses, which demonstrated better performance in the main analysis—three alternative fixed effects structures are employed, as shown in Tables C.2–C.5. The first introduces zone-year interactions, capturing annual variations in transmission patterns, capacity changes, and market developments across zones. The second implements zone-month interactions, controlling for seasonal variations in power flows, demand patterns, and operational characteristics. The third incorporates year-month interactions, accounting for temporal variations in market conditions common across zones. Throughout all specifications, baseline controls for day and day-of-week fixed effects are maintained, and where applicable, zone, hour, and 5-minute fixed effects.

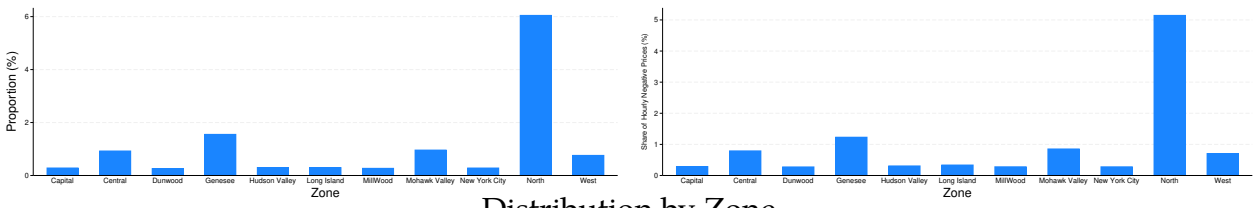
The stability of grid condition variables across probability model specifications provides support for the empirical strategy. In the 5-minute data (Table C.2), marginal congestion costs (C_{it}) maintain consistent positive coefficients across all specifications. This stability persists despite increasingly granular controls, suggesting the baseline Logit specification effectively captures underlying market dynamics. Marginal loss costs (L_{it}) demonstrate similarly consistent patterns, showing negative effects across all specifications with comparable magnitudes. The significance patterns of these coefficients vary predictably with the temporal structure of market operations. These patterns persist across hourly specifications (Table C.3), where coefficients maintain similar magnitudes and significance levels. The weather-based renewable proxies exhibit resilience in the probability models, with wind speed ($Wind_{it}$) coefficients remaining consistently positive and highly significant, while solar irradiance (Sol_{it}) maintains negative coefficients throughout.

The Negative Binomial frequency models reveal similar patterns across different time

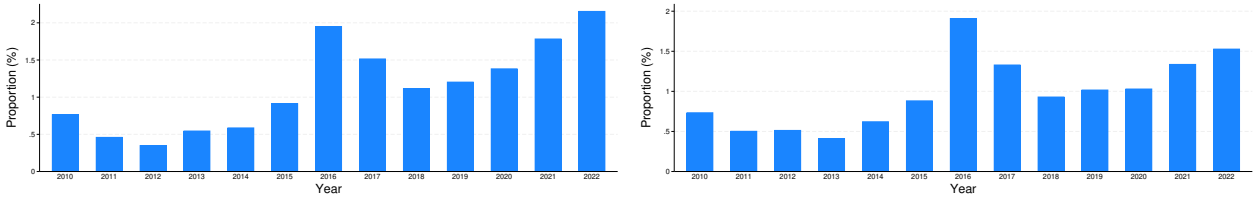
Figure B.1: Proportion of Negative Prices in 5-Minute and Hourly Data

(a) 5-Minute Data

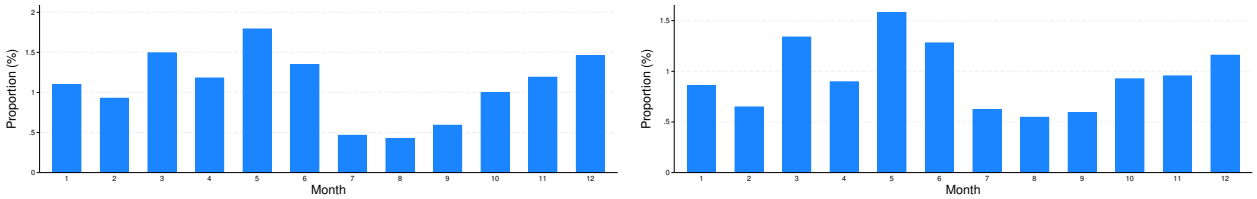
(b) Hourly Data



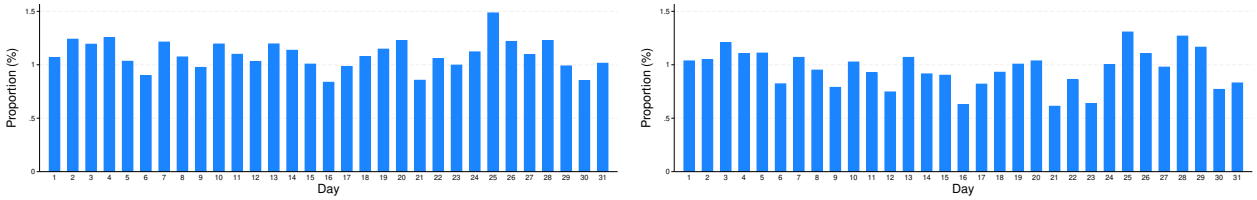
Distribution by Zone



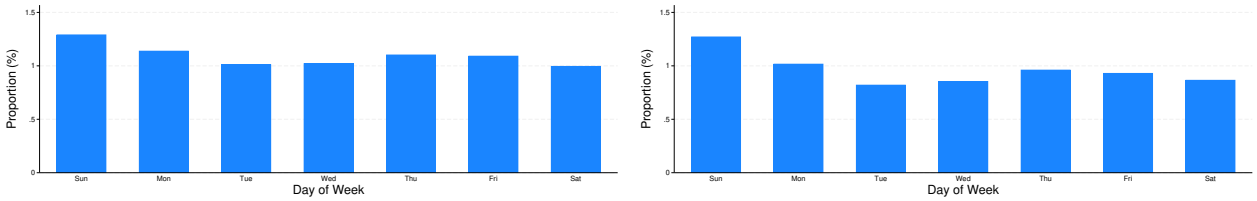
Annual Trend (2010-2022)



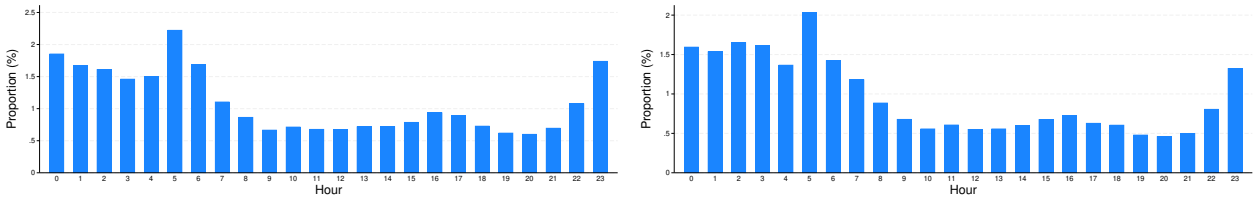
Monthly Pattern



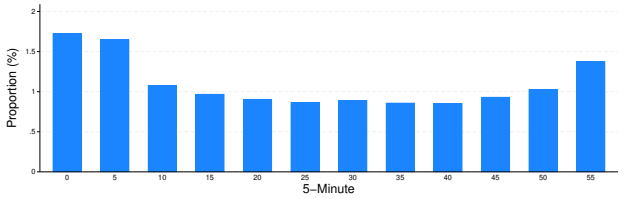
Daily Pattern (Day of Month)



Day of Week Pattern



Hourly Pattern



5-Minute Interval Pattern

horizons. In the hourly specifications (Table C.4), the grid condition variables maintain stable patterns, with marginal congestion costs (C_{it}) showing consistently small positive coefficients and marginal loss costs (L_{it}) maintaining significant negative effects across specifications. The renewable energy proxies demonstrate similar stability, with wind and solar effects remaining significant and consistent in magnitude. The weather variables show effects, with temperature (T_{it}) maintaining positive and significant coefficients, and humidity (H_{it}) showing consistent negative effects across specifications. The daily frequency results (Table C.5) reveal larger coefficient magnitudes, particularly for solar effects, suggesting cumulative impacts over longer periods, while maintaining the same pattern of significance and sign stability observed in the hourly specifications.

The dataset of over 14 million observations ensures identification of effects, even with parameter increases across specifications. The model fit statistics provide additional validation of the empirical approach. Log-likelihood values for both probability and frequency models improve consistently with more controls. The corresponding improvements in AIC and BIC suggest these additional parameters may capture variation in market patterns and dynamics, supporting the robustness of findings to different approaches in controlling for unobserved heterogeneity.

Appendix D Robustness Analysis: Controlling for System Load

The relationship between electricity load and price formation has been documented in the literature (Bönte et al., 2015, Hirth et al., 2024). However, real-time load is deliberately excluded as an explanatory variable in the baseline specifications for methodological and theoretical reasons. First, load demonstrates endogeneity with prices in real-time markets, as demand response programs and price-sensitive consumers adjust consumption based on price signals. This simultaneity poses identification challenges that could bias coefficient estimates. Second, load exhibits temporal patterns already captured through the fixed effects structure, including hour, day-of-week, and seasonal controls. Third, and importantly, load serves as a transmission mechanism through which variables of interest—particularly weather conditions and renewable generation—affect price formation. Including load as a control variable could therefore mask the total effects the analysis aims to measure, potentially leading to underestimation of these impacts.

To validate this methodological choice while addressing potential concerns about omitted

Table C.2: Logit Model Estimates with Alternative Fixed Effects: 5-Minute Data
Dependent Variable: $Pr(N_{it})$

Variables	(1) Baseline	(2) Zone×Year	(3) Zone×Month	(4) Year×Month
C_{it}	0.00686	0.00668	0.00722	0.00720
	(0.00611)	(0.00667)	(0.00661)	(0.00637)
L_{it}	-0.0940*	-0.0945	-0.0973	-0.0959*
	(0.0554)	(0.0609)	(0.0600)	(0.0580)
$Wind_{it}$	0.333***	0.334***	0.347***	0.356***
	(0.0512)	(0.0528)	(0.0468)	(0.0617)
Sol_{it}	-0.826***	-0.779***	-0.886***	-0.605***
	(0.142)	(0.135)	(0.186)	(0.125)
$Hydro_{it}$	-0.00322	-0.00466	-0.00668	-0.0246
	(0.0285)	(0.0291)	(0.0273)	(0.0301)
T_{it}	0.0624***	0.0609***	0.0697***	0.0621***
	(0.00720)	(0.00700)	(0.0113)	(0.00698)
H_{it}	-0.0945***	-0.0895***	-0.112***	-0.0901***
	(0.0147)	(0.0150)	(0.0243)	(0.0142)
Constant	-5.232***	-4.025***	-4.782***	-4.461***
	(0.816)	(0.448)	(0.844)	(0.828)
Zone FE	Yes	No	No	Yes
Year FE	Yes	No	Yes	No
Month FE	Yes	Yes	No	No
Day FE	Yes	Yes	Yes	Yes
Day of Week FE	Yes	Yes	Yes	Yes
Hour FE	Yes	Yes	Yes	Yes
5-Minute FE	Yes	Yes	Yes	Yes
Zone×Year FE	No	Yes	No	No
Zone×Month FE	No	No	Yes	No
Year×Month FE	No	No	No	Yes
Log-Likelihood	-673333.40	-656164.40	-668973.90	-645753
AIC	1346689	1312351	1337970	1291528
BIC	1346848	1312510	1338129	1291687
Observations	14,458,752	14,458,752	14,458,752	14,458,752

***, ** and *: indicate 1%, 5% and 10% significance levels, respectively. Robust clustered standard errors are in parentheses.

Table C.3: Logit Model Estimates with Alternative Fixed Effects: Hourly Data
Dependent Variable: $Pr(N_{it})$

Variables	(1) Baseline	(2) Zone×Year	(3) Zone×Month	(4) Year×Month
C_{it}	0.00946	0.00905	0.00998	0.0104
	(0.00820)	(0.00882)	(0.00885)	(0.00870)
L_{it}	-0.129*	-0.133*	-0.133*	-0.136*
	(0.0675)	(0.0743)	(0.0742)	(0.0729)
$Wind_{it}$	0.307***	0.306***	0.318***	0.330***
	(0.0602)	(0.0619)	(0.0581)	(0.0713)
Sol_{it}	-0.973***	-0.912***	-1.051***	-0.768***
	(0.165)	(0.162)	(0.224)	(0.122)
$Hydro_{it}$	0.0812*	0.0778*	0.0779*	0.0362
	(0.0435)	(0.0408)	(0.0414)	(0.0380)
T_{it}	0.0774***	0.0755***	0.0903***	0.0748***
	(0.0106)	(0.0105)	(0.0155)	(0.00895)
H_{it}	-0.135***	-0.130***	-0.165***	-0.127***
	(0.0191)	(0.0202)	(0.0300)	(0.0171)
Constant	-5.226***	-4.187***	-4.814***	-4.565***
	(0.732)	(0.380)	(0.822)	(0.738)
Zone FE	Yes	No	No	Yes
Year FE	Yes	No	Yes	No
Month FE	Yes	Yes	No	No
Day FE	Yes	Yes	Yes	Yes
Day of Week FE	Yes	Yes	Yes	Yes
Hour FE	Yes	Yes	Yes	Yes
Zone×Year FE	No	Yes	No	No
Zone×Month FE	No	No	Yes	No
Year×Month FE	No	No	No	Yes
Log-Likelihood	-52131.71	-50493.72	-51504.19	-49300.53
AIC	104285.40	101009.40	103030.40	98623.06
BIC	104417.40	101141.20	103162.40	98754.79
Observations	1,204,896	1,174,344	1,204,896	1,172,952

***, ** and *: indicate 1%, 5% and 10% significance levels, respectively. Robust clustered standard errors are in parentheses.

Table C.4: Negative Binomial Estimates with Alternative Fixed Effects: Hourly Data

Dependent Variable: $Fr(N)_{it}$

Variables	(1) Baseline	(2) Zone×Year	(3) Zone×Month	(4) Year×Month
C_{it}	0.000749	0.000707	0.000847*	0.00104
	(0.000543)	(0.000436)	(0.000505)	(0.000749)
L_{it}	-0.0257***	-0.0260***	-0.0253***	-0.0260***
	(0.00397)	(0.00385)	(0.00384)	(0.00521)
$Wind_{it}$	0.224***	0.224***	0.227***	0.233***
	(0.0395)	(0.0388)	(0.0391)	(0.0458)
Sol_{it}	-0.773***	-0.791***	-0.825***	-0.594***
	(0.0582)	(0.0589)	(0.0778)	(0.0678)
$Hydro_{it}$	-0.0222	-0.0192	-0.0150	-0.0319*
	(0.0173)	(0.0168)	(0.0152)	(0.0181)
T_{it}	0.0565***	0.0567***	0.0641***	0.0587***
	(0.00563)	(0.00615)	(0.00878)	(0.00655)
H_{it}	-0.0922***	-0.0925***	-0.110***	-0.0923***
	(0.0126)	(0.0141)	(0.0194)	(0.0139)
Constant	-1.787***	-1.110***	-1.566***	-1.099***
	(0.260)	(0.0937)	(0.284)	(0.282)
Zone FE	Yes	No	No	Yes
Year FE	Yes	No	Yes	No
Month FE	Yes	Yes	No	No
Day FE	Yes	Yes	Yes	Yes
Day of Week FE	Yes	Yes	Yes	Yes
Hour FE	Yes	Yes	Yes	Yes
Zone×Year FE	No	Yes	No	No
Zone×Month FE	No	No	Yes	No
Year×Month FE	No	No	No	Yes
Log-Likelihood	-278939.3	-274563.4	-278096.6	-271363.1
AIC	557898.70	549148.70	556213.20	542748.20
BIC	558018.70	549280.80	556333.20	542880.30
Observations	1,204,896	1,204,896	1,204,896	1,204,896

***, ** and *: indicate 1%, 5% and 10% significance levels, respectively. Robust clustered standard errors are in parentheses.

Table C.5: Negative Binomial Estimates with Alternative Fixed Effects: Daily Data

Dependent Variable: $Fr(N)_{it}$

Variables	(1) Baseline	(2) Zone×Year	(3) Zone×Month	(4) Year×Month
C_{it}	0.00346	0.00204	0.00197	0.00579
	(0.00547)	(0.00570)	(0.00549)	(0.00636)
L_{it}	-0.215	-0.294*	-0.207	-0.177**
	(0.132)	(0.175)	(0.135)	(0.0893)
$Wind_{it}$	0.259***	0.251***	0.269***	0.292***
	(0.0682)	(0.0659)	(0.0729)	(0.0833)
Sol_{it}	-1.181***	-1.436***	-1.400***	-0.696*
	(0.393)	(0.402)	(0.422)	(0.367)
$Hydro_{it}$	0.0526	0.0108	0.0939*	-0.0834
	(0.0652)	(0.0635)	(0.0559)	(0.0761)
T_{it}	0.0748***	0.0751***	0.106***	0.0656***
	(0.0167)	(0.0191)	(0.0154)	(0.0137)
H_{it}	-0.129***	-0.128***	-0.194***	-0.0935***
	(0.0401)	(0.0455)	(0.0348)	(0.0299)
Constant	-1.679***	-0.930**	-1.185*	-1.106**
	(0.507)	(0.387)	(0.647)	(0.531)
Zone FE	Yes	No	No	Yes
Year FE	Yes	No	Yes	No
Month FE	Yes	Yes	No	No
Day FE	Yes	Yes	Yes	Yes
Day of Week FE	Yes	Yes	Yes	Yes
Zone×Year FE	No	Yes	No	No
Zone×Month FE	No	No	Yes	No
Year×Month FE	No	No	No	Yes
Log-Likelihood	-20874.69	-20225.50	-20644.58	-19773.91
AIC	41769.38	40473.01	41309.17	39569.82
BIC	41857.62	40570.07	41397.4	39666.88
Observations	50,204	50,204	50,204	50,204

***, ** and *: indicate 1%, 5% and 10% significance levels, respectively. Robust clustered standard errors are in parentheses.

variable bias, additional analyses incorporating real-time load as an explanatory variable at both 5-minute and hourly frequencies are conducted.[D.1] Augmented versions of the baseline models are estimated:

$$y_{it} = \mathbf{X}_{it}^{\top}\boldsymbol{\beta} + \theta Q_{it} + \gamma_i + \delta_t + \epsilon_{it} \tag{D.26}$$

where Q_{it} represents the real-time load in zone i at time t, and all other variables retain their previous definitions. To address potential endogeneity concerns, different lag structures for occurrence and frequency analyses are employed.

For the probability analysis of occurrence, three specifications at each frequency are considered. For 5-minute data: (1) contemporaneous load, (2) one-period (5-minute) lagged load, and (3) twelve-period (1-hour) lagged load. For hourly data: (1) contemporaneous load, (2) one-period (1-hour) lagged load, and (3) twenty-four-period (24-hour) lagged load. For the frequency analysis examining the count of negative price events, longer lag structures are employed. For 1-hour data: (1) contemporaneous load, (2) one-period (1-hour) lagged load, and (3) twenty-four-period (24-hour) lagged load. For daily data: (1) contemporaneous load, (2) one-period (1-day) lagged load, and (3) seven-period (7-day) lagged load. Tables D.6, D.7, D.8, and D.9 present the results of these augmented models. The probability models maintain the Logit specification that demonstrated better performance in the main analysis, while the frequency models employ the Negative Binomial framework.

The 5-minute probability models (Table D.6) reveal that real-time load exhibits a significant negative relationship with negative price occurrence, with the effect moderating as the lag structure is extended. The impact is strongest for contemporaneous load and gradually diminishes with longer lags, suggesting that immediate load conditions have a pronounced effect on price formation at high frequencies. Importantly, variables of interest maintain their significance and relative magnitudes, with wind energy ($Wind_{it}$) retaining its positive effect and solar energy (Sol_{it}) its negative impact across all specifications.

The hourly probability models (Table D.7) demonstrate similar stability in the core coefficient estimates but with some differences in the temporal structure of load effects. The contemporaneous load maintains statistical significance at the 10% level, with similar effects

[D.1] When aggregating data to hourly frequency, the average is taken for all other variables in the analysis except for load values. For load, the 5-minute readings are summed since each 5-minute interval represents energy consumed (kWh) during that period. This distinction in treatment is necessary because load measures actual energy consumption over time, unlike the instantaneous measurements captured by other variables.

observed for the one-hour lag, while the 24-hour lag captures daily cyclical patterns in load-price relationships. This suggests that load effects may accumulate and persist differently at lower frequencies.

The hourly frequency analysis (Table D.8) reveals patterns in load effects across different temporal scales. The coefficients show a consistent decline from contemporaneous to longer lags, suggesting a gradual attenuation of load effects over time. Wind energy ($Wind_{it}$) demonstrates a consistent positive and highly significant effect, while solar energy (Sol_{it}) exhibits a negative effect across all specifications.

The daily frequency results (Table D.9) reveal patterns in load effects over longer horizons. The contemporaneous load effect is modest, and the relationship evolves across longer lags, with the seven-day lag showing a significant positive effect. This sign reversal suggests temporal dynamics in load-price relationships over weekly horizons. The renewable energy variables maintain their significant effects, with wind energy ($Wind_{it}$) showing consistent positive impacts and solar energy (Sol_{it}) demonstrating negative effects, though with larger magnitudes compared to the hourly specifications.

These results suggest that while load plays a role in price formation, its inclusion does not alter the main findings regarding the impacts of renewable integration, weather conditions, and grid constraints. The robustness of results across different lag specifications and time resolutions validates the original empirical strategy while providing insights into the temporal dynamics of load-price relationships in electricity markets.

Appendix E Robustness Analysis: Variable Interactions

The dynamics of electricity markets suggest that renewable integration, weather conditions, and grid constraints likely interact in ways that influence price formation beyond their individual impacts. These potential interaction effects are investigated using theoretically motivated specifications that capture economic mechanisms and market dynamics.

The following specification with interaction terms is estimated:

$$y_{it} = \mathbf{X}_{it}^{\top}\boldsymbol{\beta} + \sum_{j,k} \theta_{jk}(x_{j,it} \times x_{k,it}) + \gamma_i + \delta_t + \epsilon_{it} \tag{E.27}$$

where $x_{j,it} \times x_{k,it}$ represents interactions between variables j and k, and θ_{jk} captures their joint effect. Based on economic theory and market characteristics, three categories of interactions are examined:

Table D.6: Logit Model Estimates Including Load Effects: 5-Minute Data
Dependent Variable: $Pr(N_{it})$

Variables	(1) Contemp.	(2) 5-min lag	(3) 1-hour lag
Q_{it}	-0.00117**	-0.00118**	-0.000992**
	(0.000566)	(0.000565)	(0.000479)
C_{it}	0.00681	0.00681	0.00680
	(0.00618)	(0.00619)	(0.00617)
L_{it}	-0.0942*	-0.0942*	-0.0938*
	(0.0556)	(0.0556)	(0.0555)
$Wind_{it}$	0.334***	0.334***	0.334***
	(0.0517)	(0.0517)	(0.0514)
Sol_{it}	-0.688***	-0.687***	-0.704***
	(0.0941)	(0.0942)	(0.0957)
$Hydro_{it}$	-0.00586	-0.00580	-0.00546
	(0.0290)	(0.0290)	(0.0291)
T_{it}	0.0484***	0.0483***	0.0501***
	(0.00588)	(0.00589)	(0.00528)
H_{it}	-0.0670***	-0.0668***	-0.0699***
	(0.0104)	(0.0104)	(0.00932)
Constant	-3.972***	-3.961***	-4.114***
	(1.060)	(1.061)	(1.040)
Zone FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Month FE	Yes	Yes	Yes
Day FE	Yes	Yes	Yes
Day of Week FE	Yes	Yes	Yes
Hour FE	Yes	Yes	Yes
5-Minute FE	Yes	Yes	Yes
Log-Likelihood	-671008.80	-670984.80	-671486.80
AIC	1342040	1341992	1342996
BIC	1342199	1342151	1343155
Observations	14,458,752	14,458,741	14,458,620

***, ** and *: indicate 1%, 5% and 10% significance levels, respectively. Robust clustered standard errors are in parentheses.

Table D.7: Logit Model Estimates Including Load Effects: Hourly Data
Dependent Variable: $Pr(N_{it})$

Variables	(1) Contemp.	(2) 1-hour lag	(3) 24-hour lag
Q_{it}	-0.000133*	-0.000126**	-4.11e-05
	(7.01e-05)	(6.32e-05)	(3.40e-05)
C_{it}	0.00925	0.00922	0.00942
	(0.00824)	(0.00827)	(0.00814)
L_{it}	-0.130**	-0.129*	-0.129*
	(0.0662)	(0.0668)	(0.0666)
$Wind_{it}$	0.306***	0.307***	0.307***
	(0.0624)	(0.0618)	(0.0604)
Sol_{it}	-0.776***	-0.787***	-0.911***
	(0.0976)	(0.0982)	(0.130)
$Hydro_{it}$	0.0808*	0.0809*	0.0844*
	(0.0417)	(0.0417)	(0.0439)
T_{it}	0.0583***	0.0591***	0.0729***
	(0.00641)	(0.00610)	(0.00816)
H_{it}	-0.0969***	-0.0980***	-0.126***
	(0.0128)	(0.0120)	(0.0133)
Constant	-3.503***	-3.524***	-4.670***
	(1.185)	(1.153)	(0.861)
Zone FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Month FE	Yes	Yes	Yes
Day FE	Yes	Yes	Yes
Day of Week FE	Yes	Yes	Yes
Hour FE	Yes	Yes	Yes
Log-Likelihood	-51811.43	-51835.85	-52079.60
AIC	103644.90	103693.70	104181.20
BIC	103776.90	103825.70	104313.20
Observations	1,204,896	1,204,885	1,204,632

***, ** and *: indicate 1%, 5% and 10% significance levels, respectively. Robust clustered standard errors are in parentheses.

Table D.8: Negative Binomial Estimates Including Load Effects: Hourly Data
Dependent Variable: $Fr(N)_{it}$

Variables	(1) Contemp.	(2) 1-hour lag	(3) 24-hour lag
Q_{it}	-7.58e-05**	-7.47e-05**	-3.07e-05**
	(3.60e-05)	(3.33e-05)	(1.42e-05)
C_{it}	0.000605	0.000598	0.000716
	(0.000596)	(0.000585)	(0.000564)
L_{it}	-0.0252***	-0.0254***	-0.0255***
	(0.00378)	(0.00379)	(0.00392)
$Wind_{it}$	0.225***	0.225***	0.224***
	(0.0401)	(0.0400)	(0.0397)
Sol_{it}	-0.596***	-0.597***	-0.698***
	(0.0874)	(0.0835)	(0.0533)
$Hydro_{it}$	-0.0292*	-0.0291*	-0.0227
	(0.0173)	(0.0176)	(0.0173)
T_{it}	0.0452***	0.0453***	0.0508***
	(0.00525)	(0.00509)	(0.00600)
H_{it}	-0.0592***	-0.0592***	-0.0818***
	(0.00933)	(0.00907)	(0.00974)
Constant	-0.875*	-0.827*	-1.437***
	(0.454)	(0.450)	(0.246)
Zone FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Month FE	Yes	Yes	Yes
Day FE	Yes	Yes	Yes
Day of Week FE	Yes	Yes	Yes
Hour FE	Yes	Yes	Yes
Log-Likelihood	-278138	-278129.10	-278492.80
AIC	556296	556278.20	557005.60
BIC	556416	556398.20	557125.70
Observations	1,204,896	1,204,885	1,204,632

***, ** and *: indicate 1%, 5% and 10% significance levels, respectively. Robust clustered standard errors are in parentheses.

Table D.9: Negative Binomial Estimates Including Load Effects: Daily Data
Dependent Variable: $Fr(N)_{it}$

Variables	(1) Contemp.	(2) 1-day lag	(3) 7-day lag
Q_{it}	-1.85e-05	-3.83e-06	4.73e-05***
	(3.43e-05)	(2.36e-05)	(1.35e-05)
C_{it}	0.00331	0.00344	0.00316
	(0.00563)	(0.00554)	(0.00539)
L_{it}	-0.210	-0.214	-0.217
	(0.130)	(0.131)	(0.134)
$Wind_{it}$	0.259***	0.259***	0.261***
	(0.0690)	(0.0684)	(0.0677)
Sol_{it}	-1.085**	-1.171***	-1.267***
	(0.426)	(0.405)	(0.399)
$Hydro_{it}$	0.0410	0.0518	0.0692
	(0.0541)	(0.0621)	(0.0591)
T_{it}	0.0685***	0.0738***	0.0833***
	(0.00937)	(0.0117)	(0.0146)
H_{it}	-0.113***	-0.126***	-0.149***
	(0.0268)	(0.0303)	(0.0335)
Constant	-1.473**	-1.626***	-2.306***
	(0.615)	(0.629)	(0.693)
Zone FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Month FE	Yes	Yes	Yes
Day FE	Yes	Yes	Yes
Day of Week FE	Yes	Yes	Yes
Log-Likelihood	-20871.04	-20869.28	-20797.76
AIC	41762.09	41758.57	41615.52
BIC	41850.33	41846.80	41703.74
Observations	50,204	50,193	50,127

***, ** and *: indicate 1%, 5% and 10% significance levels, respectively. Robust clustered standard errors are in parentheses.

1. *Supply-side complementarities*: Interactions between different renewable sources ($Wind_{it} \times Sol_{it}$) capture potential complementarities or substitution effects in generation profiles
2. *Grid condition interactions*: Products of transmission constraints and renewable generation ($C_{it} \times Wind_{it}$, $C_{it} \times Sol_{it}$) reflect how congestion impacts renewable resource utilization
3. *Weather-renewable interactions*: Terms combining weather variables with generation ($T_{it} \times Wind_{it}$, $T_{it} \times Sol_{it}$, $H_{it} \times Wind_{it}$, $H_{it} \times Sol_{it}$) capture how atmospheric conditions affect renewable production efficiency[E.2]

Tables E.10, E.11, E.12, and E.13 present results from these interaction models for both probability and frequency analyses at different time resolutions. For each time scale, supply-side interactions are progressively introduced (Column 1), grid condition interactions are added (Column 2), and weather-renewable interactions are incorporated (Column 3). The probability models maintain the Logit specification, while the frequency models employ the Negative Binomial framework that demonstrated better performance in the main analysis.

The 5-minute probability analysis (Table E.10) reveals mechanisms. First, the wind-solar interaction ($Wind_{it} \times Sol_{it}$) shows significant positive coefficients across all specifications, with the effect moderating from supply-side to weather interactions. This suggests that the combined presence of wind and solar generation has a complementary effect on negative price formation. Second, the grid condition interactions reveal that transmission constraints moderately amplify both wind and solar effects, though with greater impact on wind generation. The weather-renewable interactions demonstrate patterns, with temperature showing stronger effects when interacting with solar generation compared to wind generation, while humidity interactions exhibit the opposite pattern.

The hourly probability models (Table E.11) maintain similar patterns but with some differences in magnitude. The wind-solar interaction remains significant but shows a diminishing effect when weather interactions are included, suggesting that some of the complementarity operates through weather channels. The congestion-renewable interactions demonstrate effects, with both wind and solar interactions maintaining significance across specifications. Weather interactions reveal stronger temperature effects compared to the 5-minute data, particularly for solar generation.

[E.2]While the baseline model includes precipitation ($Hydro_{it}$) as a control variable, it is excluded from interaction terms because its effect on electricity markets operates through hydropower generation, which has different operational characteristics. Unlike wind and solar resources that exhibit contemporaneous relationships with weather conditions, hydropower's relationship with precipitation involves temporal lags and reservoir management decisions. Moreover, the baseline results consistently show limited direct effects of precipitation.

The frequency analyses (Tables E.12 and E.13) provide insights into how these interactions affect the count of negative price events. The hourly frequency results show that wind-solar complementarity persists but weakens with the inclusion of weather effects. Notably, the congestion-renewable interactions maintain significance only for wind generation, while weather interactions show asymmetric effects across renewable sources. The daily frequency models reveal patterns, with interaction effects generally showing larger magnitudes but lower statistical significance, suggesting that some interaction effects may dissipate over longer horizons.

These findings illuminate aspects of market operation. First, they highlight temporal dependencies in renewable resource complementarity, operating both through direct supply-side effects and indirect weather channels. Second, they demonstrate that transmission constraints have heterogeneous effects across renewable technologies, with impacts on wind generation. Third, they reveal weather-dependent patterns in renewable generation impacts that vary across time scales. These insights suggest that market design and renewable integration policies should account for these interaction effects to better manage negative price occurrences.

The inclusion of interaction terms improves model fit across specifications, as evidenced by higher log-likelihood values and lower AIC and BIC measures. However, the patterns identified in the main analysis remain robust. The direct effects of variables maintain their signs and significance, with magnitudes largely unchanged after accounting for interactions. This stability supports the baseline findings while providing additional nuance about the mechanisms through which various factors influence negative price formation.

Table E.10: Logit Model Estimates with Interaction Terms: 5-Minute Data
Dependent Variable: $Pr(N_{it})$

Variables	(1) Supply	(2) Grid	(3) Weather
C_{it}	0.00687	-0.00369	-0.00381
	(0.00612)	(0.0114)	(0.0114)
L_{it}	-0.0944*	-0.0849	-0.0858
	(0.0555)	(0.0624)	(0.0628)
$Wind_{it}$	0.306***	0.306***	0.301***
	(0.0439)	(0.0341)	(0.0388)
Sol_{it}	-1.846***	-1.799***	-1.061***
	(0.130)	(0.121)	(0.0867)
$Hydro_{it}$	0.0162	0.0211	-0.00684
	(0.0290)	(0.0305)	(0.0314)
T_{it}	0.0687***	0.0693***	0.0469**
	(0.00677)	(0.00673)	(0.0200)
H_{it}	-0.106***	-0.109***	-0.116***
	(0.0137)	(0.0119)	(0.0294)
$Wind_{it} \times Sol_{it}$	0.293***	0.280***	0.127***
	(0.0517)	(0.0357)	(0.0196)
$C_{it} \times Wind_{it}$		0.00432	0.00439
		(0.00583)	(0.00581)
$C_{it} \times Sol_{it}$		0.00526	0.00539
		(0.0106)	(0.0106)
$T_{it} \times Wind_{it}$			0.00871***
			(0.00286)
$T_{it} \times Sol_{it}$			-0.0804***
			(0.0287)
$H_{it} \times Wind_{it}$			-0.00495
			(0.00576)
$H_{it} \times Sol_{it}$			0.159***
			(0.0494)
Constant	-5.050***	-5.052***	-4.899***
	(0.780)	(0.735)	(0.759)
Zone FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Month FE	Yes	Yes	Yes
Day FE	Yes	Yes	Yes
Day of Week FE	Yes	Yes	Yes
Hour FE	Yes	Yes	Yes
5-Minute FE	Yes	Yes	Yes
Log-Likelihood	-672587.60	-668123.20	-666842.50
AIC	1345197	1336268	1333707
BIC	1345356	1336428	1333866
Observations	14,458,752	14,458,752	14,458,752

***, ** and *: indicate 1%, 5% and 10% significance levels, respectively. Robust clustered standard errors are in parentheses.

Table E.11: Logit Model Estimates with Interaction Terms: Hourly Data
Dependent Variable: $Pr(N_{it})$

Variables	(1) Supply	(2) Grid	(3) Weather
C_{it}	0.00947	-0.00936	-0.00932
	(0.00823)	(0.0151)	(0.0153)
L_{it}	-0.129*	-0.0952	-0.0973
	(0.0677)	(0.0811)	(0.0830)
$Wind_{it}$	0.283***	0.282***	0.268***
	(0.0519)	(0.0377)	(0.0531)
Sol_{it}	-1.848***	-1.729***	-0.638***
	(0.126)	(0.140)	(0.148)
$Hydro_{it}$	0.0972**	0.107***	0.0751
	(0.0399)	(0.0372)	(0.0458)
T_{it}	0.0832***	0.0839***	0.0626***
	(0.00981)	(0.00883)	(0.0216)
H_{it}	-0.145***	-0.151***	-0.166***
	(0.0177)	(0.0130)	(0.0359)
$Wind_{it} \times Sol_{it}$	0.263***	0.230***	0.0381
	(0.0554)	(0.0410)	(0.0290)
$C_{it} \times Wind_{it}$		0.00730	0.00732
		(0.00633)	(0.00631)
$C_{it} \times Sol_{it}$		0.0130	0.0132
		(0.0146)	(0.0142)
$T_{it} \times Wind_{it}$			0.00980***
			(0.00226)
$T_{it} \times Sol_{it}$			-0.108***
			(0.0415)
$H_{it} \times Wind_{it}$			-0.00513
			(0.00562)
$H_{it} \times Sol_{it}$			0.196**
			(0.0820)
Constant	-5.062***	-5.093***	-4.850***
	(0.686)	(0.665)	(0.751)
Zone FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Month FE	Yes	Yes	Yes
Day FE	Yes	Yes	Yes
Day of Week FE	Yes	Yes	Yes
Hour FE	Yes	Yes	Yes
Log-Likelihood	-52,130.2	-52,127.8	-52,125.6
AIC	104,282.4	104,277.6	104,273.2
BIC	104,414.4	104,409.6	104,405.2
Observations	1,204,896	1,204,896	1,204,896

***, ** and *: indicate 1%, 5% and 10% significance levels, respectively. Robust clustered standard errors are in parentheses.

Table E.12: Negative Binomial Estimates with Interaction Terms: Hourly Data
Dependent Variable: $Fr(N)_{it}$

Variables	(1) Supply	(2) Grid	(3) Weather
C_{it}	0.000765	-0.00326	-0.00327
	(0.000544)	(0.00260)	(0.00244)
L_{it}	-0.0259***	-0.0137	-0.0140
	(0.00393)	(0.0145)	(0.0139)
$Wind_{it}$	0.207***	0.214***	0.192***
	(0.0337)	(0.0324)	(0.0383)
Sol_{it}	-1.451***	-1.457***	-0.890**
	(0.189)	(0.170)	(0.404)
$Hydro_{it}$	-0.00998	-0.00703	-0.0261
	(0.0177)	(0.0193)	(0.0204)
T_{it}	0.0608***	0.0604***	0.0586***
	(0.00493)	(0.00499)	(0.0129)
H_{it}	-0.0998***	-0.101***	-0.134***
	(0.0113)	(0.0106)	(0.0219)
$Wind_{it} \times Sol_{it}$	0.216***	0.218***	0.113*
	(0.0610)	(0.0602)	(0.0642)
$C_{it} \times Wind_{it}$		0.00187*	0.00186*
		(0.00111)	(0.00103)
$C_{it} \times Sol_{it}$		-0.000144	4.89e-05
		(0.00998)	(0.00914)
$T_{it} \times Wind_{it}$			0.00351**
			(0.00157)
$T_{it} \times Sol_{it}$			-0.0935***
			(0.0362)
$H_{it} \times Wind_{it}$			0.00244
			(0.00332)
$H_{it} \times Sol_{it}$			0.187***
			(0.0544)
Constant	-1.674***	-1.706***	-1.503***
	(0.229)	(0.246)	(0.280)
Zone FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Month FE	Yes	Yes	Yes
Day FE	Yes	Yes	Yes
Day of Week FE	Yes	Yes	Yes
Hour FE	Yes	Yes	Yes
Log-Likelihood	-278809.20	-278491.80	-278217.10
AIC	557638.30	557003.50	556454.20
BIC	557758.30	557123.60	556574.20
Observations	1,204,896	1,204,896	1,204,896

***, ** and *: indicate 1%, 5% and 10% significance levels, respectively. Robust clustered standard errors are in parentheses.

Table E.13: Negative Binomial Estimates with Interaction Terms: Daily Data
Dependent Variable: $Fr(N)_{it}$

Variables	(1) Supply	(2) Grid	(3) Weather
C_{it}	0.00340	0.00179	0.00260
	(0.00553)	(0.0109)	(0.0109)
L_{it}	-0.214	-0.194	-0.190
	(0.132)	(0.140)	(0.137)
$Wind_{it}$	0.242***	0.283***	0.187*
	(0.0669)	(0.0636)	(0.0970)
Sol_{it}	-1.618	-1.526	-0.861
	(1.052)	(1.014)	(1.226)
$Hydro_{it}$	0.0593	0.0525	-0.00468
	(0.0620)	(0.0683)	(0.0721)
T_{it}	0.0761***	0.0771***	0.104***
	(0.0177)	(0.0159)	(0.0211)
H_{it}	-0.131***	-0.144***	-0.248***
	(0.0419)	(0.0357)	(0.0573)
$Wind_{it} \times Sol_{it}$	0.157	0.0510	-0.169
	(0.252)	(0.262)	(0.233)
$C_{it} \times Wind_{it}$		0.00581	0.00557
		(0.00473)	(0.00478)
$C_{it} \times Sol_{it}$		-0.0698**	-0.0717**
		(0.0279)	(0.0283)
$T_{it} \times Wind_{it}$			-0.00416
			(0.00641)
$T_{it} \times Sol_{it}$			-0.122
			(0.0803)
$H_{it} \times Wind_{it}$			0.0266*
			(0.0138)
$H_{it} \times Sol_{it}$			0.197
			(0.212)
Constant	-1.611***	-1.611***	-1.235**
	(0.466)	(0.439)	(0.520)
Zone FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Month FE	Yes	Yes	Yes
Day FE	Yes	Yes	Yes
Day of Week FE	Yes	Yes	Yes
Log-Likelihood	-20874.04	-20812.41	-20798.36
AIC	41768.07	41644.82	41616.71
BIC	41856.31	41733.05	41704.95
Observations	50,204	50,204	50,204

***, ** and *: indicate 1%, 5% and 10% significance levels, respectively. Robust clustered standard errors are in parentheses.

Appendix F Nomenclature

This appendix presents key notation used throughout the paper, including electricity market-specific variables and central theoretical and empirical parameters.

Symbol	Description
Sets and Indices	
Ω	Set of zones, indexed by i
$\mathcal{T}$	Set of time periods, indexed by t
Ω_B^i	Set of baseload generators in zone i
Ω_R^i	Set of renewable generators in zone i
g	Generator index
Generator Decision Variables	
q_{git}	Generation quantity of generator g in zone i at time t (MWh)
u_{git}	Binary commitment status (1 = operational, 0 = offline)
b_{git}^*	Optimal bid price (\$/MWh)
Market Prices	
p_{it}^*, LMP_{it}	Locational marginal price in zone i at time t (\$/MWh)
p_t^{sys}	System-wide energy component of price (\$/MWh)
C_{it}, MCC_{it}	Marginal congestion cost (\$/MWh)
L_{it}, MCL_{it}	Marginal loss cost (\$/MWh)
D_{it}	Electricity demand (MWh)
Q_{it}	System load (MWh)
Generator Cost and Subsidy Parameters	
$A_g(\cdot)$	Variable production cost function (\$)
$\Phi_{git}(\cdot)$	Cycling cost function (startup, shutdown, ramping) (\$)
κ_g	Startup cost for generator g (\$)
δ_g	Shutdown cost for generator g (\$)
γ_g	Ramping cost per unit change in output (\$/MW)
η_g	Fixed operational cost (\$)
$SUB_{git}(\cdot)$	Subsidy function (\$)
s_g	Per-unit production subsidy (\$/MWh)
ξ_g	Fixed renewable energy credit (\$)
Operational Constraints	
$\underline{q}_g, \overline{q}_g$	Minimum and maximum generation capacity (MW)
$\underline{r}_g, \overline{r}_g$	Maximum ramp-down and ramp-up rates (MW/period)

Continued on next page

Continued from previous page

Symbol	Description
$\overline{F}_{ijt}$	Transmission capacity limit between zones i and j (MW)
ν_{ijt}	Shadow price of transmission constraint (\$/MW)
Empirical Variables	
N_{it}	Binary indicator for negative prices (1 = negative, 0 = otherwise)
$Fr(N)_{it}$	Frequency (count) of negative price occurrences
$\mathbf{X}_{it}$	Vector of explanatory variables
$Wind_{it}$	Wind speed (m/s)
Sol_{it}	Solar irradiance (kWh/m^2)
$Hydro_{it}$	Precipitation (mm/hour)
T_{it}	Temperature (°C)
H_{it}	Humidity (g/kg)
Econometric Parameters	
y_{it}	Generic dependent variable (N_{it} or $Fr(N)_{it}$)
y^*_{it}	Latent propensity variable in binary response models
$\boldsymbol{\beta}$	Coefficient vector in binary response models
$\boldsymbol{\theta}$	Coefficient vector in count data models
γ_i	Zone fixed effects
δ_t	Time fixed effects (year, month, day, hour, 5-minute)
ϵ_{it}	Error term
μ_{it}	Conditional mean in count data models
α	Overdispersion parameter in Negative Binomial model
AME_k	Average marginal effect of variable k