Do Consumers Need Hourly Prices? Simulation in Belgian Electricity Markets

Elise Viadere[*†]

October 31, 2025

Abstract

Decarbonizing electricity requires shifting consumption toward hours when renewable generation is abundant and costs are low. Yet most consumers face flat prices with no incentive to respond. Real-time pricing is the theoretical first-best but sees limited demand response. Time-of-use pricing offers a simpler alternative. This paper evaluates their relative efficiency at equilibrium, accounting for how consumer responses feeds back to wholesale markets. Simulations using Belgian data show a well-designed two-period time-of-use captures nearly 100% of real-time pricing's efficiency gains. The model assumes equal consumer attention to both pricing schemes, isolating the pure efficiency cost of price simplification from behavioral frictions. These results help rationalize weak real-time pricing response: under perfect attention consumers make nearly identical consumption decisions under simple or complex prices, questioning whether real-time pricing remains first-best under realistic attention constraints.

Keywords: Time-of-use pricing, retail electricity pricing, demand response, electric vehicles, load shifting

[*]European Center of Advanced Research in Economics and Statistics (ECARES), Université Libre de Bruxelles (ULB)

[†]I am grateful to Nicolas Astier, Quentin Hoarau, Marten Ovaere, Guntram Wolff and seminar and conferences participants at CentralSupélec, Paris School of Economics, University of Gent for helpful comments and suggestions. Remaining errors are mine. The research leading to these results has received funding from the SPF Économie through the Belgian Energy Transition Fund as part of the DemandFlex research project. The results of this paper do not reflect the official position of the SPF Économie.

1 Introduction

Decarbonizing the economy requires both transforming electricity generation toward intermittent renewables and electrifying end uses such as heating and transport. These two processes exert opposing pressures on wholesale electricity prices. For a given level of demand, greater renewable penetration, with its near-zero marginal costs, tends to lower equilibrium prices and increase intraday volatility. Conversely, for a given supply capacity, electrification raises aggregate demand, amplifying price pressure when dispatchable generation must balance residual load. Figure 1 illustrates this transformation: intraday price variation has intensified substantially between 2018 and 2024, reflecting the amplified influence of intermittent generation.

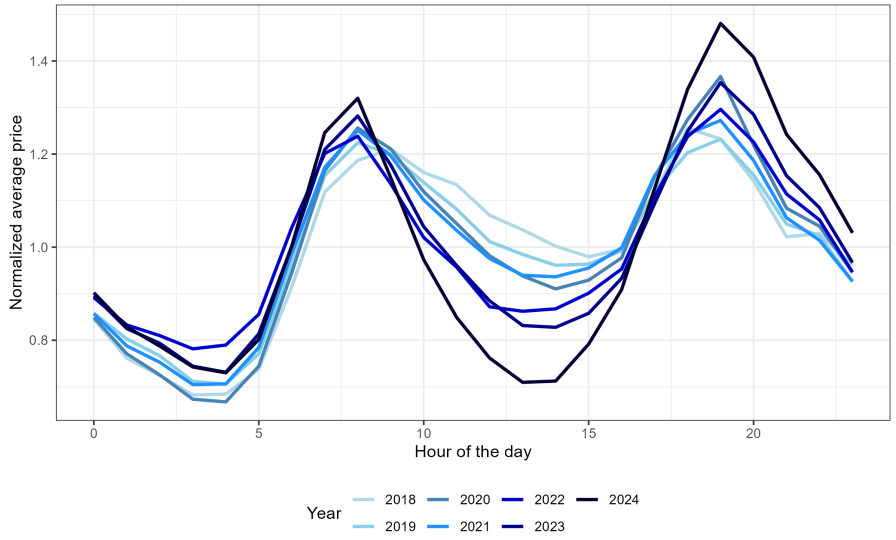


Figure 1: Rising intraday volatility in wholesale electricity prices (2018–2024, normalized by year)

Note: Hourly day-ahead prices are averaged by year using weekday data. Each year's curve is normalized by its own annual mean to highlight changes in intraday price patterns over time. Data from Belgian wholesale markets illustrate broader trends in systems with growing renewable penetration. **Source:** ENTSOE Transparency Platform: Day-ahead prices, Belgian bidding zone (2018–2024), weekdays only.

This price variability creates opportunities for demand-side flexibility, enabling consumers to shift electricity consumption toward cheaper hours and prevent curtailment of renewable energy during surplus periods. As electrification proceeds, this flexibility becomes increasingly valuable. Electric vehicles and heat pumps represent growing shares of electricity consumption: between 2019 and 2023, European markets experienced rapid electrification of both heating and transport, with heat pump sales rising 88% and electric vehicles sales surging 486%.[1] Where and when this consumption occurs has first-order implications for system costs, infrastructure

[1]European annual heat pump sales increased from 1.4 million units in 2019 to 2.63 million in 2023, while electric vehicles sales rose from 140,000 to 820,000 units over the same period (IEA (2024a), IEA (2024b).

requirements, and emission reductions

This paper addresses a fundamental question in retail electricity market design: Do consumers need hourly real-time prices? Electricity consumers today typically face flat electricity prices that remain constant throughout the day, despite wholesale prices varying substantially across hours. Flat electricity prices abstract from intertemporal variation in marginal costs and thus violate the core logic of peak-load pricing (Boiteux (1949)). As Joskow and Wolfram (2012) emphasize, suppressing time-varying retail prices distorts consumption and blunts investment signals.

Policymakers typically consider two alternatives for exposing consumers to these intraday relative prices. Under *real-time pricing*, retail prices track wholesale market prices hour-by-hour, updating daily to reflect actual supply and demand conditions. Under *time-of-use pricing*, regulators divide the day into a small number of periods (typically two to four), each with a fixed price designed to reflect average costs during those periods.[2] These prices remain constant for months or years, providing predictability but sacrificing the hour-by-hour precision of real-time pricing. More precisely, this paper asks: can simplified time-of-use pricing structures achieve outcomes comparable to real-time pricing in aligning demand with variable renewable supply, or does the discretization inherent in time-of-use pricing impose substantial efficiency costs?

The theoretical case for real-time pricing appears compelling. When retail prices track wholesale markets hour-by-hour, consumers face the true marginal cost of their electricity consumption at each moment. This alignment between prices and costs should, in principle, guide efficient resource allocation: consumers shift load away from high-cost peak hours toward low-cost off-peak periods, reducing the need for expensive backup generation while preventing curtailment of renewable energy during surplus hours. Economic theory predicts that real-time pricing would reduce peak demand (Borenstein (2005)), lower system-wide emissions (Holland and Mansur (2008)), and limit the exercise of market power in wholesale markets (Poletti and Wright (2020)).

Yet real-time pricing has proven remarkably difficult to implement in practice. Despite being available in many markets, few consumers voluntarily adopt real-time pricing contracts. The volatility in wholesale electricity prices (particularly dramatic during the 2022 energy crisis) creates financial risk that makes consumers hesitant to expose themselves to hour-by-hour

[2]For example, a three-period structure might charge high prices during evening peak hours (18:00–21:00), medium prices during daytime hours (09:00–18:00), and low prices overnight (21:00–09:00).

price fluctuations (Pébereau and Remmy (2023)). Even when regulators mandate real-time pricing as the default, consumer responses remain surprisingly limited. (Fabra et al. (2021)) shows no significant demand reponse from large-scale implementation in Spain, partly because the complexity of hourly price signals makes them difficult for consumers to understand and act upon. Jessoe and Rapson (2014) demonstrate experimentally that consumers only respond when information about price surges is made easily accessible. More fundamentally, the potential bill savings for individual households are typically small relative to the cognitive effort required to actively manage consumption hour-by-hour. (Ito (2014)) find that most households respond primarily to average price levels rather than to hourly variation, substantially limiting the efficiency gains from real-time pricing in practice. Enrich et al. (2024) show that Spain's time-of-use program induced habit formation and gradual adaptation, ultimately producing permanent changes in consumption behavior—suggesting that even simpler time-varying pricing can generate persistent demand response

These practical challenges have led policymakers to consider time-of-use pricing as an alternative approach.[3] By simplifying the price signal into a small number of predictable daily periods, time-of-use pricing trade some allocative precision for transparency and ease of comprehension. The idea is that this simplification might facilitate broader consumer participation. But this raises a fundamental question: does simplifying prices from 24 hourly values to three or four daily blocks come at acceptable efficiency cost, or does it fundamentally undermine the potential for price-responsive demand to support the energy transition?

Researchers have tackled this question from multiple angles, yet no consensus has emerged. Existing research has approached this question through three distinct lenses. One strand evaluates time-of-use designs by measuring how well they approximate hourly wholesale prices, typically using statistical measures such as in-sample R-squared (Hogan (2014)). Jacobsen et al. (2020) find that even complex time-of-use structures with 24 daily price periods explain only 13% of wholesale price variation in the PJM market in 2012. In their most extreme specification of hour-of-day interacted with day-of-week and month-of-year fixed effects, which abandons the simplicity objective of time-of-use pricing—the explanatory power reaches only 43%. Astier (2021) show that optimized pricing with three or four pricing periods explain 15–25% of French wholesale price variation between 2012 and 2016, and 20–30% of California wholesale price variation between 2015 and 2018. Astier (2021) further demonstrates that

[3]Automated demand response via smart devices can sidestep behavioral frictions, especially those created by default appliance-operation settings (Bailey et al., 2024; Byrne and Martin, 2025).), but requires substantial upfront technology investment that may not be cost-effective compared to simpler pricing-based solutions.

increasing the number of daily periods yields diminishing returns, as the fit metric flattens considerably after three daily pricing periods. Correlation-based welfare approximations rely on near-linear demand and supply curves and no systematic relationship between demand elasticity and pricing errors across hours. Hinchberger et al. (2024) address these limitations through an adjusted R-squared measure and simulations of wholesale market outcomes under alternative retail pricing. Their results indicate that time-of-use pricing correct about 10% of welfare losses from flat-rate pricing.

A second approach challenges the focus on price replication, arguing instead that what matters for demand response is preserving the relative prices between hours rather than matching absolute price levels. Studies using rank correlation metrics find that well-designed time-of-use structures can match 60–80% of real-time pricing's incentive strength in systems with high solar penetration and flexible loads (Schittekatte et al. (2024)). Yet this approach conflates two distinct frictions: information costs from discretizing continuous prices and behavioral costs from limited consumer attention. Moreover, these results prove sensitive to market characteristics: in systems dominated by inflexible nuclear generation or with highly weather-sensitive heating demand, correlations fall to 25–35% (Cabot and Villavicencio (2024)).

A third body of work uses structural models to directly simulate welfare outcomes under alternative pricing regimes. These studies typically show that real-time pricing can reduce the variance, both within- and across-days (Holland and Mansur (2006)), with estimates ranging widely depending on model assumptions about generation technology, demand elasticity, and whether the analysis considers short-run or long-run adjustments. Borenstein (2005) demonstrates that even with modest demand elasticity, moving large customers to real-time pricing generates substantial efficiency gains by reducing the need for peaker capacity that operates only during the highest-demand hours.

Yet despite this substantial body of research, two critical gaps remain. First, existing analyses ignore how retail pricing feeds back to wholesale markets. The sole exception is Cabot and Villavicencio (2024), which accounts for one round of adjustment and represents the closest work in the current literature. When consumers shift load in response to time-of-use prices, aggregate demand patterns change, altering which generators run and at what prices. These new wholesale prices should, in principle, inform revised time-of-use structures, inducing further consumer adjustments. This process continues iteratively until retail prices, consumer behavior, and wholesale outcomes reach mutual consistency. Failing to model this complete feedback loop can substantially misrepresent both the equilibrium allocation, its efficiency

properties, and the stability of pricing schemes. The magnitude of this omission grows with the scale of flexible loads: as electric vehicles and heat pumps come to represent larger shares of residential consumption, their collective response to pricing signals increasingly influences wholesale market clearing, making the feedback effect non-negligible.

Second, at the exception of structural models, prior research has predominantly evaluated time-of-use pricing by comparing price series: measuring statistical correlations between hourly wholesale prices and simplified time-of-use pricing. This approach treats price signal quality as an inherent property of the retail pricing structure, independent of how consumers respond to this signal. But the efficiency of any retail pricing depends on the complete chain from pricing design through consumer behavior to wholesale market impacts. The critical question is not whether time-of-use pricing accurately replicate wholesale prices statistically, but whether they induce consumer responses that efficiently align aggregate demand with variable renewable supply.

This paper addresses both gaps by developing an integrated simulation framework that jointly models time-of-use pricing design, consumer load-shifting decisions for any shiftable load, and wholesale market dispatch. The framework iterates to equilibrium, capturing the feedback loops between retail pricing, consumer response, and wholesale price formation. Critically, the approach assumes no behavioral frictions: consumers understand both real-time pricing and time-of-use pricing equally well, face identical time preferences and constraints under both regimes, and respond optimally to price signals in both cases. This design isolates the pure informational cost of price discretization, measuring how much efficiency is lost from simplifying 24 hourly prices into fewer daily blocks when consumers are equally capable of responding to either pricing structure. Any differences in outcomes therefore reflect the informational content of the pricing regime itself, not differential behavioral responses or comprehension difficulties. Moreover, the welfare comparison between time-of-use and real-time pricing is robust to assumptions about consumer population size and baseline load profiles. Because both pricing schemes are evaluated using the same underlying consumer load-shifting behavior, the relative performance depends on the structure of the pricing rather than on these calibration parameters.

The analysis yields three main findings. First, well-designed two-period time-of-use pricing achieve nearly 100% of the efficiency a,d welfare gains reached under real-time pricing at equilibrium. This near-equivalence reflects a fundamental property of the load-shifting problem

that depends on the relative prices across feasible shiftable load starting times rather than precise hourly price levels. Time-of-use structures are optimized to maximize rank correlation and preserve this ordering structure where it matters for consumer decisions.

Second, the analysis helps resolve an empirical puzzle in the literature. Simulation studies consistently find large efficiency gaps between time-of-use pricing and real-time pricing, while field experiments show minimal behavioral response to real-time prices. This apparent contradiction reflects different counterfactuals. Simulation studies typically assume optimal consumer response to all price signals, making real-time pricing's superior informational content deliver large efficiency gains. Field experiments compare actual consumer behavior under real-time pricing, which is minimal due to complexity, limited salience and attention, and small potential savings, to flat-rate baselines. My framework shows that when behavioral response capability *is held constant across pricing regimes*, the informational advantage of real-time pricing becomes negligible for well-designed time-of-use structures. Combined with extensive field evidence that real-time pricing fails to elicit meaningful response while simpler time-of-use schemes achieve higher participation, this suggests that time-of-use pricing may dominate in practice not despite its simplicity, but because of it.

Third, the empirical application exploits detailed data that allow realistic calibration of both electricity supply and demand. On the supply side, I use wholesale market outcomes combined with firms generation capacity by technology, enabling calibration of strategic markup behavior across different market conditions. On the demand side, I use benchmark residential consumption patterns including electric vehicle and heat pump ownership rates and charging behavior. While the application uses data from Belgium, the methodological framework and substantive insights apply broadly to electricity systems integrating intermittent renewables and electrifying transportation and heating.

The remainder of this paper proceeds as follows. Section 2 develops the theoretical framework, specifying consumer optimization problems, wholesale market dispatch, and time-of-use design optimization. Section 3 presents the datasets used in the simulations. Section 4 presents the simulation implementation. Section 5 reports equilibrium outcomes under alternative retail pricing regimes and decomposes efficiency differences. Section 6 concludes.

2 Theoretical framework

The model comprises three interacting components: consumer demand for electricity distinguishing inelastic baseload from flexible loads that can be shifted in response to prices (Section

2.1), wholesale price formation under oligopolistic competition (Section 2.2), and regulatory design of time-of-use pricing structures (Section 2.3). These components interact sequentially: the regulator sets prices, consumers respond by shifting flexible loads, and wholesale prices clear based on aggregate demand. The process iterates until the wholesale price distribution converges.

2.1 Consumer demand response

Demand structure. Consumer i has two types of electricity demand. Baseline consumption $Q_{i,t}^{\text{base}}$ (lighting, appliances) is inelastic to prices, where $t \in \{1, \ldots, 24\}$ indexes hours. Flexible load L_i (electric vehicle charging, heat pump operation) must be consumed within a D_i-hour window but can be optimally scheduled across the day.

Discrete choice problem. The consumer observes 24-hour retail prices $\mathbf{p}^r = (p_1^r, \ldots, p_{24}^r)$ where p_t^r is the retail price at hour t, and chooses when to begin the D_i-hour consumption block. Crucially, the consumer understands the retail pricing structure perfectly and responds optimally in both cases. The choice problem is:

$$h_i^* = \arg\min_{h \in \mathcal{H}} \left\{ \underbrace{\sum_{s=0}^{D_i-1} p_{h+s}^r \cdot \frac{L_i}{D_i}}_{\text{electricity cost}} + \underbrace{\lambda_i \sum_{s=0}^{D_i-1} \omega_{i,h+s}}_{\text{timing disutility}} \right\} \tag{1}$$

where h is a potential starting hour, $\mathcal{H} = \{1, \ldots, 24 - D_i + 1\}$ is the set of feasible starting hours, s indexes hours within the consumption block, $\lambda_i \geq 0$ measures sensitivity to inconvenient scheduling, and $\omega_{i,t}$ are hour-specific timing penalties with $\omega_{i,t_i^*} = 0$ at the most preferred time t_i^*.

The consumer selects the D_i-hour window minimizing the sum of electricity cost and timing disutility. The program reduces consumption timing to a discrete choice among at most 24 alternatives rather than a continuous allocation problem across hours. Each alternative corresponds to a different starting time for the consumption block. The consumer evaluates the average price across the D_i consecutive hours in each feasible window and selects the minimum-cost option, accounting for timing preferences.

Consumption profile. The optimal choice h_i^* determines the flexible load profile:

$$x_{i,t}^* = \begin{cases} \dfrac{L_i}{D_i} & \text{if } t \in \{h_i^*, h_i^* + 1, \ldots, h_i^* + D_i - 1\} \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

where $x_{i,t}^*$ is consumer i's optimal flexible load at hour t. Total consumption at hour t combines baseline and flexible components:

$$q_{i,t}^* = Q_{i,t}^{\text{base}} + x_{i,t}^* \tag{3}$$

Aggregate demand. With N consumers characterized by $(L_i, D_i, \lambda_i, \boldsymbol{\omega}_i)$ and baseline profiles $Q_{i,t}^{\text{base}}$, aggregate flexible demand at hour t is:

$$Q_t^{\text{flex}}(\mathbf{p}^r) = \sum_{i=1}^{N} x_{i,t}^*(\mathbf{p}^r) \tag{4}$$

Total system demand is:

$$Q_t(\mathbf{p}^r) = Q_t^{\text{base}} + Q_t^{\text{flex}}(\mathbf{p}^r), \quad \text{where } Q_t^{\text{base}} = \sum_{i=1}^{N} Q_{i,t}^{\text{base}} \tag{5}$$

Timing preferences and behavioral frictions. The parameter λ_i governs the trade-off between cost minimization and scheduling convenience. When $\lambda_i = 0$, consumer i exhibits perfect flexibility, shifting load frictionlessly to the lowest-cost window. As λ_i increases, timing preferences increasingly constrain load shifting, with $\lambda_i \to \infty$ yielding perfect inertia where consumption remains fixed at t_i^* regardless of prices.

2.2 Wholesale market equilibrium

From retail to wholesale demand. Given retail prices $\mathbf{p}^r = (p_1^r, \ldots, p_{24}^r)$, consumers choose consumption profiles $x_{i,t}^*(\mathbf{p}^r)$ as described above. This determines the 24-hour system load vector

$$\mathbf{Q}(\mathbf{p}^r) \equiv (Q_1, \ldots, Q_{24}), \qquad Q_t = Q_t^{\text{base}} + \sum_{i=1}^{N} x_{i,t}^*(\mathbf{p}^r)$$

that the wholesale market must serve. Retail pricing thus reallocates load intertemporally while holding each consumer's daily energy requirement L_i fixed.

Market structure and timing. The day-ahead wholesale market clears 24 hours before delivery. Generators submit bids for all 24 hours simultaneously, based on known market

conditions (weather forecasts, expected demand). These bids represent binding commitments: once submitted, generators cannot adjust offers in response to realized prices within the day.[4] Generators are assumed to perfectly anticipate the hourly load profile $\mathbf{Q}(\mathbf{p}^r)$ implied by retail prices. The market then clears independently for each hour.

Consider an oligopolistic market with M firms. Firm $j \in \{1, \ldots, M\}$ owns a portfolio of generation units with heterogeneous technologies. Each unit g has capacity $\bar{k}_{g,j}$ (MW) and marginal cost $c_{g,j}$ (€/MWh). To focus on short-run price formation, I abstract from ramp constraints, unit unavailability, and transmission congestion.

Strategic pricing. Under perfect competition, generators bid at marginal cost and the market clears at $p_t^w = MC(Q_t)$ where p_t^w is the wholesale price and $MC(\cdot)$ is the marginal cost function. In practice, concentrated markets with near-inelastic demand create opportunities for strategic markup. Based on typical Cournot oligopoly model, firm j producing q_j units at market price p^w earns profit:

$$\pi_j(q_j) = p^w(Q)q_j - C_j(q_j) \tag{6}$$

where $Q = q_j + \sum_{j' \neq j} q_{j'}$ is total market output, $p^w(Q)$ is the inverse demand curve, and $C_j(\cdot)$ is firm j's cost function.

The first-order condition for profit maximization yields:

$$p^w(Q) = MC_j(q_j) - q_j \frac{dp^w}{dq_j} \tag{7}$$

where $MC_j(q_j)$ is firm j's marginal cost. The markup term $-q_j \frac{dp^w}{dq_j} > 0$ captures firm j's incentive to raise price: withholding a marginal unit loses profit on that unit but raises revenue on all infra-marginal capacity. With $\frac{dp^w}{dq_j} < 0$, larger producers have stronger incentives to mark up above marginal cost.

The slope $\frac{dp^w}{dq_j}$ depends on the firm's residual demand curve. From market clearing, firm j faces residual demand:

$$q_j = Q - S_{-j}(p^w) \tag{8}$$

where $S_{-j}(p^w)$ is aggregate supply from all competitors excluding firm j. Inverting this relationship gives $p^w = P_j^{RD}(q_j)$ where $P_j^{RD}(\cdot)$ is firm j's inverse residual demand function, and

[4]This day-ahead structure implies that demand facing the wholesale market is effectively inelastic even under real-time retail pricing. Although $p_t^r = p_t^w$ would create a direct link between wholesale prices and retail consumption, the sequential timing breaks this feedback: wholesale prices are determined 24 hours in advance based on forecast demand, before consumers observe prices and make consumption decisions. Formally, $\frac{dQ_t}{dp_t^w} \approx 0$ at the wholesale clearing stage because Q_t is predetermined by the day-ahead forecast.

the slope of this inverse residual demand is:

$$\frac{dp^w}{dq_j} = \frac{1}{\frac{dQ}{dp^w} - \frac{dS_{-j}}{dp^w}} \tag{9}$$

Because short-run demand is nearly inelastic ($\frac{dQ}{dp^w} \approx 0$), this simplifies to:

$$\theta_j \equiv \frac{dp^w}{dq_j} \approx -\frac{1}{\frac{dS_{-j}}{dp^w}} \tag{10}$$

The parameter $\theta_j < 0$ summarizes firm j's market power: when competitors' supply is less elastic, $|\theta_j|$ is large, meaning small output reductions substantially raise prices.

Bidding strategy. Equation (7) determines the pricing rule; the bidding strategy implements this rule through the day-ahead auction. Firm j can rank its generation units from cheapest to most expensive. Let unit g have marginal cost $c_{g,j}$ and capacity $\bar{k}_{g,j}$. Define cumulative capacity of all cheaper units as $K_{g,j} = \sum_{m<g} \bar{k}_{m,j}$ where m indexes units cheaper than g. This capacity is infra-marginal: it will be dispatched before unit g, earning the market clearing price on all these units.

Equation (7) implies that firm j should bid unit g at:

$$B_{g,j} = \max\{c_{g,j} - \theta_j K_{g,j}, c_{g,j}\} \tag{11}$$

where $B_{g,j}$ is the bid price for unit g by firm j. The markup term $-\theta_j K_{g,j} > 0$ increases with infra-marginal capacity $K_{g,j}$, reflecting that higher prices apply to a larger base of already-dispatched units. The $\max\{\cdot\}$ operator prevents bidding below marginal cost. This bidding rule directly implements the pricing condition (7): the markup scales with the firm's residual demand slope θ_j and with how much capacity the firm has already placed in the merit order.

Market clearing. Given bid schedules $\{B_{g,j,t}\}$ for all units, generators, and hours, the market clears independently each hour. The system operator constructs the aggregate merit-order supply curve by horizontally summing firm-level curves and dispatches units from cheapest to most expensive until total capacity meets demand Q_t:

$$p_t^w = \min\left\{P : \sum_{j=1}^{M} \sum_{g} \bar{k}_{g,j} \mathbf{1}\{B_{g,j,t} \leq P\} \geq Q_t\right\} \tag{12}$$

where $\mathbf{1}\{\cdot\}$ is the indicator function and P is a candidate price. All units with $B_{g,j,t} \leq p_t^w$ dispatch fully; those with $B_{g,j,t} > p_t^w$ do not run. The resulting price p_t^w reflects both underlying generation costs and strategic markups embedded in the bid functions.

The connection between equations (7) and (12) is direct: the pricing rule determines optimal markups given market structure, while the bidding strategy implements these markups through day-ahead commitments. Market clearing then selects the price consistent with these strategic bids and realized demand.

2.3 Time-of-use pricing design

Design objective. The regulator (or retail supplier under regulatory mandate) designs a time-of-use pricing structure to align consumer incentives with wholesale cost variation. The goal is to stimulate demand response, facilitate renewable integration, and reduce system costs from electrification while maintaining simplicity for consumer comprehension. Given the consumer response program from Section 2.1, consumers selecting charging windows care primarily about the rank ordering of average prices across feasible starting times rather than absolute price levels. The regulator therefore maximizes rank correlation between time-of-use retail prices and wholesale prices, preserving the relative ordering that drives efficient load shifting. At the first iteration, the regulator's prior reflects 2023 wholesale prices under flat retail pricing, absent any load shifting. As consumers adjust their charging behavior in response, new wholesale market-clearing prices emerge. The regulator iteratively updates the pricing structure based on these clearing prices until the wholesale price distribution stabilizes, signaling convergence to equilibrium.

Time-of-use structure. A time-of-use pricing scheme partitions the 24-hour day into $F \in \{3, 4, 5\}$ blocks of consecutive hours, where F balances simplicity (fewer periods) against allocative precision (more periods). Let $\mathcal{P} = \{P_1, P_2, \ldots, P_F\}$ denote a partition, where each block P_f contains at least 4 consecutive hours and $f \in \{1, \ldots, F\}$ indexes blocks. Blocks may wrap around midnight; for example, an overnight block might include hours $\{22, 23, 24, 1, 2, 3\}$.

For any partition $\mathcal{P}$, the time-of-use price in block f equals the load-weighted average wholesale price across all hours in that block:

$$p_f^r = \frac{\sum_{d \in \mathcal{D}} \sum_{t \in P_f} p_{t,d}^w Q_{t,d}}{\sum_{d \in \mathcal{D}} \sum_{t \in P_f} Q_{t,d}} \tag{13}$$

where d indexes days, $p_{t,d}^w$ is the day-ahead wholesale price at hour t on day d, $Q_{t,d}$ is observed

load at hour t on day d, and $\mathcal{D}$ is the set of historical days used for calibration. Load-weighting ensures that prices reflect actual costs incurred during those periods.

This pricing rule creates a 24-hour retail price vector $\mathbf{p}^r(\mathcal{P}) = (p_1^r, \ldots, p_{24}^r)$ where $p_t^r = p_f^r$ for all $t \in P_f$. Each hour within a block faces the same price, but prices vary across blocks.

Optimal partition. The regulator's problem is:

$$\begin{aligned} \max_{\mathcal{P}} \quad & \mathbb{E}_d \left[\rho_S \left(\mathbf{p}^r(\mathcal{P}), \mathbf{p}_d^w \right) \right] \\ \text{subject to} \quad & \mathcal{P} = \{P_1, \ldots, P_F\} \text{ partitions } \{1, \ldots, 24\} \\ & |P_f| \geq 4 \quad \forall f \\ & \text{each } P_f \text{ consists of consecutive hours (modulo 24)} \\ & F \in \{3, 4, 5\} \end{aligned} \tag{14}$$

where $\rho_S(\cdot, \cdot)$ denotes Spearman rank correlation between hourly price vectors, $\mathbf{p}_d^w = (p_{1,d}^w, \ldots, p_{24,d}^w)$ is the wholesale price vector on day d, and $\mathbb{E}_d[\cdot]$ denotes expectation over the empirical distribution of historical days in $\mathcal{D}$.

Spearman correlation measures how well the time-of-use structure preserves the rank ordering of wholesale prices. Since consumers solve discrete choice problems comparing average costs across feasible windows, preserving this ordering ensures that time-of-use pricing replicates the allocation decisions consumers would make under real-time pricing.

3 Data

The empirical analysis uses Belgian electricity market data from 2020-2023, combining publicly available wholesale market information with calibrated parameters for flexible residential loads and generation costs. This section describes three novel datasets constructed for the analysis.

3.1 Wholesale prices and system Load

Hourly day-ahead wholesale electricity prices and total system load come from the ENTSO-E Transparency Platform via Elia, the Belgian transmission system operator. The dataset covers 2020-2023 (35,040 hours total), partitioned into a training sample (2020-2022, 26,280 hours) for time-of-use design optimization and an out-of-sample evaluation period (2023, 8,760 hours).

Table 1 presents summary statistics for 2023. Wholesale prices exhibit substantial variation both across days (inter-day standard deviation of €36/MWh) and within days (intra-day

standard deviation of €29/MWh). The median peak-to-trough price ratio within days is 2.8, providing clear signals for efficient load shifting. System load similarly varies across days (standard deviation 1,002 MW) and within days (standard deviation 945 MW), reflecting both weather-driven demand fluctuations and regular diurnal patterns.

Table 1: Summary statistics: Belgian wholesale prices and system load, 2023

Variable	Mean	SD	Min	Max
Inter-day variation (daily means):				
Day-ahead price (€/MWh)	97	36	-2	205
Total load (MW)	9,555	1,002	6,855	12,339
Intra-day variation (deviations from daily means):				
Day-ahead price (€/MWh)	0	29	-388	578
Total load (MW)	0	945	-2,670	6,115

Notes: Inter-day statistics use daily means (average of 24 hourly values for each day). Intra-day statistics summarize hourly deviations from each day's mean, pooled across all days; these deviations average to zero within each day by construction. *Source:* ENTSO-E Transparency Platform, Belgian bidding zone, 2023.

3.2 Flexible Load Parameters

The consumer demand model requires parameters for electric vehicles and heat pumps. Table 2 summarizes the calibration, which draws on multiple sources. As of 2023, Belgium had 138,745 registered electric vehicles **?** and approximately 209,000 heat pumps Gils et al. (2024); Joint Research Centre (JRC) (2024).

For electric vehicles, I assume 43 kWh per charging cycle (80% of a typical 54 kWh battery capacity) with 6-hour charging duration. Following Rangaraju et al. (2015), approximately 57% of vehicles charge daily, yielding 79,282 active charging cycles per day. For heat pumps, I assume 12 kWh per heating cycle International Energy Agency (2023) with 4-hour operating duration Nolting and Praktiknjo (2019), with all units operating daily during the winter evaluation period. Baseline consumption profiles under flat pricing—showing when consumers would charge or heat absent price incentives—come from load patterns reported in Elia's Adequacy and Flexibility Study Elia (2024). Figure **??** displays these profiles, which concentrate EV charging in evening hours (18:00-24:00) and heat pump operation during morning (06:00-09:00) and evening (18:00-22:00) peaks.

The analysis considers a 20% participation rate among technically eligible loads. The

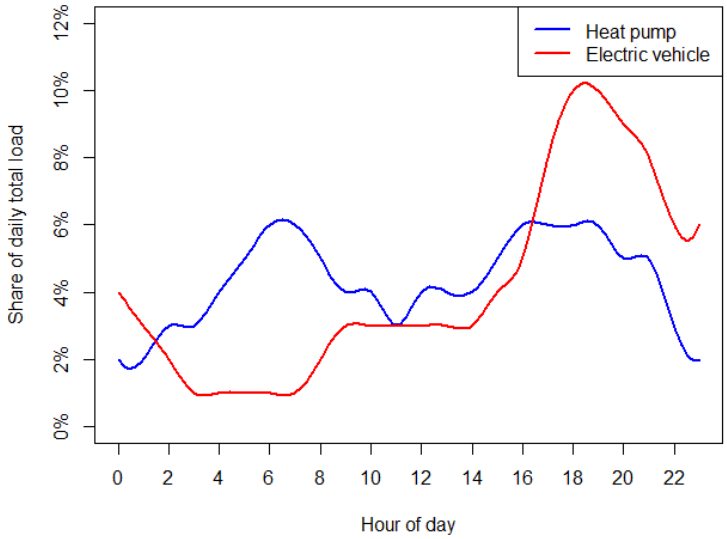


Figure 2: Baseline hourly usage profile of EVs and HPs under flat pricing (2023).

Note: The normalized profiles are used as time-preference weights for residential heat-pump operation and EV charging in the simulation. The estimated population across daily cycles is initially distributed over these profiles, which serve as the baseline before any shifting. Source: Elia (2023), Adequacy and Flexibility Study 2024–2034.

Table 2: Flexible load parameters and assumptions used in the Belgium 2023 simulation

Parameter	Electric Vehicles	Heat Pumps
Installed base	138,745[a]	209,000[b]
Daily active units	79,282[c]	209,000
Energy per cycle (kWh)	43[d]	12[e]
Cycle duration (hours)	6[c]	4[f]
Total daily energy (MWh)	3,409	2,508

Sources: [a]Statbel (2023); [b]Gils et al. (2024); Joint Research Centre (JRC) (2024); [c]Rangaraju et al. (2015); [d]80% of 54 kWh battery capacity; [e]International Energy Agency (2023); [f]Nolting and Praktiknjo (2019).

remaining 80% of potential flexible demand is either technologically constrained or chooses not to enroll in the tested retail pricing (real-time pricing or time-of-use), reflecting adoption frictions and heterogeneous adjustment costs.

3.3 Generation Capacity and Costs

The wholesale market model requires unit-level data on installed capacity and marginal costs for strategic bidding simulation. Elia provides installed capacity by generation unit and fuel type for all Belgian generators. As of 2023, 17 firms controlled 23.8 GW total capacity. Market structure is highly concentrated: the three largest firms (Electrabel/ENGIE, Luminus, Total-

Energies) control 87.5% of dispatchable capacity, consistent with the oligopolistic framework in Section 2.2.

To construct time-varying marginal costs, I combine four data sources: (i) technology-specific thermal efficiencies from the Joint Research Centre database; (ii) daily natural gas prices from TTF futures (accessed via Datastream); (iii) EU ETS carbon allowance prices for 2023; and (iv) emission factors by fuel type from RTE (French TSO). This yields hourly marginal costs ranging from €0/MWh for renewables to €120-180/MWh for gas peaking units during high fuel price periods.

For intermittent renewable generation, I use hourly wind and solar output data from Elia's Open Data platform. Belgium's installed wind capacity (1.7 GW onshore, 2.3 GW offshore) and solar capacity (6.8 GW) exhibit substantial hourly and seasonal variation, with solar capacity factors averaging 12% annually and wind averaging 28%.

These datasets enable complete simulation from retail price design through consumer response to wholesale market outcomes, with all parameters grounded in observed market conditions.

4 Methodology

I apply the theoretical framework to Belgian electricity market data from 2023 in a counterfactual simulation exercise. The simulation solves for equilibrium through an iterative algorithm that integrates retail pricing design, consumer demand response, and strategic wholesale market clearing, identifying a fixed point where retail prices, consumer load-shifting decisions, and wholesale market outcomes are mutually consistent. The counterfactual analysis compares three alternative retail pricing regimes: flat-rate, time-of-use, and real-time pricing—under identical baseline conditions. All regimes employ the same market fundamentals (demand, generation capacity, fuel costs) and strategic conduct parameters calibrated from observed 2023 outcomes, differing solely in their retail pricing structure and the resulting consumer optimization problem. Figure 3 illustrates the equilibrium structure under each regime and how the iterative solution procedure adapts to regime-specific differences in consumer response.

Baseline (flat pricing): No iteration required. Consumers face uniform retail prices throughout the day, replicating the observed 2023 Belgian system. With no price variation across hours, flexible loads remain at their observed baseline profiles, no load shifting occurs. The wholesale market clears using strategic bidding calibrated to 2023 data (described below in Stage

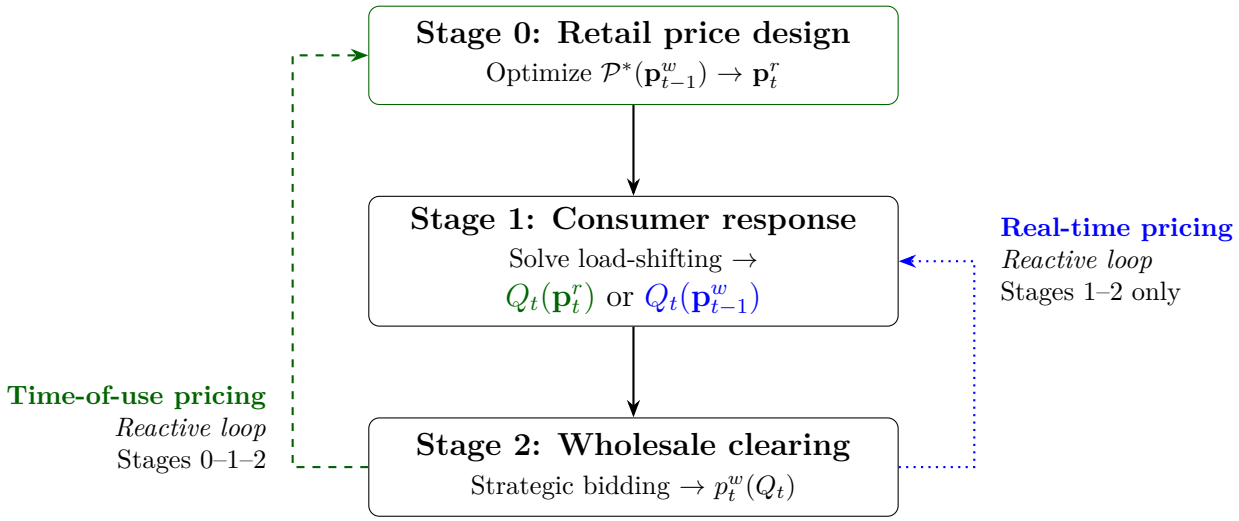


Figure 3: Simulation workflow under alternative retail pricing

Note: The figure depicts how the simulation is executed under three retail time-of-use pricing regimes. *Stage 0* optimizes time-of-use prices design based on previous period wholesale prices p^w_{t-1}. *Stage 1* computes load shifting given those prices p^r_t simulation on optimized time-of-use prices or p^w_{t-1} if real-time pricing. *Stage 2* determines wholesale prices via strategic bidding based on the results total demand Q_t. *Baseline (flat pricing):* one pass through Stages 1–2 with fixed flat retail prices (no iteration, no load shifting fed back to Stage 0). *Time-of-use (green dashed loop):* Stages 0–1–2 iterate—wholesale outcomes feed back to redesign the ToU partition, until the partition stabilizes. *Real-time pricing (blue dotted bypass):* Stage 0 is bypassed ($p^r_t = p^w_{t-1}$); only Stages 1–2 iterate until prices/quantities converge.

2), generating baseline wholesale prices. This regime serves as the counterfactual representing current market outcomes.

Time-of-use pricing: Consumers face simplified pricing with 3 or 4 daily blocks, optimized as described in Stage 0. Price variation induces load shifting: flexible loads reallocate to minimize costs given the time-of-use structure. The wholesale market clears using the same strategic bidding parameters as baseline, but with shifted demand profiles from consumer response. The algorithm iterates Stages 0-1-2 until the optimal partition stabilizes.

Real-time pricing: Consumers face hourly wholesale prices directly ($p^r_t = p^w_{t-1}$ from previous iteration). Price variation induces load shifting across all 24 hours. The wholesale market clears using the same strategic bidding parameters, but with shifted demand profiles. The algorithm iterates Stages 1-2 only (Stage 0 is bypassed) until demand and prices converge.

All three regimes use identical baseline conditions: 2023 observed hourly load, renewable generation (wind and solar), dispatchable capacity by fuel type, and flexible load parameters (138,745 electric vehicles, 209,000 heat pumps from Section 3). Crucially, the strategic conduct parameters $\{\theta_{j,b}\}$ calibrated from 2023 baseline prices remain fixed across all three regimes.

This isolates the pure effect of retail pricing design on load allocation and wholesale market outcomes, holding strategic behavior constant.

4.1 Stage 0: Time-of-use design

The algorithm enumerates all partitions of 24 hours into $F \in 3, 4$ contiguous blocks, with each block containing at least four consecutive hours. This restriction limits artificial load concentration at block boundaries and aligns with typical flexible-load cycle durations.[5]

To ensure comparability across pricing regimes, block prices for each partition $\mathcal{P}$ are calibrated on the 2023 strategic-bidding baseline. Prices equal load-weighted averages of the strategic-bidding baseline baseline wholesale price profile.

The optimal partition maximizes expected Spearman rank correlation between the time-of-use price vector and the hourly baseline wholesale prices:

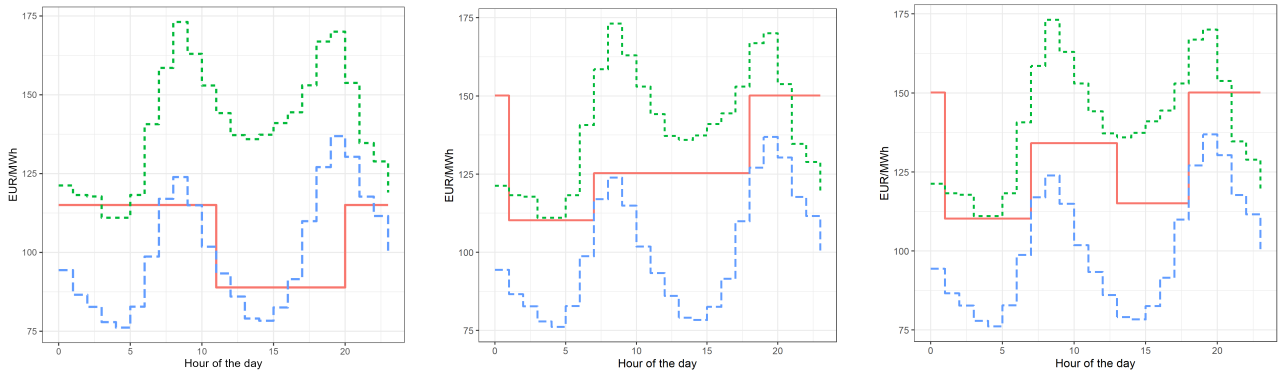
$$\mathcal{P}^* = \arg\max_{\mathcal{P}} \mathbb{E}_{d \in 2023} \left[\rho_S \left(\mathbf{p}^r(\mathcal{P}), \mathbf{p}_d^w \right) \right] \tag{15}$$

where the expectation is taken over the 2023 evaluation period, using the strategic-bidding baseline to maintain full consistency of outcomes across counterfactual pricing structures.

Figure 4 display the resulting structures for electric vehicles. The three-block design defines night (00:00–07:00), day (07:00–20:00), and peak (20:00–24:00) periods. The four-block design adds a shorter shoulder window. These structures remain stable across all scenarios and iterations.[6]

[5] Detailed explanation of time-of-use optimisation in Appendix C.

[6] From panel (a) of figure 4, the optimized two-period schedule departs from Belgium's standard "peak/oof-peak" design (07:00–22:00), which shows little intraday variation (see Appendix A). Equivalent results for heat pumps are documented in Appendix E.



(a) Two-block ToU (b) Three-block ToU (c) Four-block ToU

Figure 4: Optimized time-of-use (ToU) pricing structures after convergence

Notes: Red solid line: optimized time-of-use designs and prices after three iterations of 20% of electric vehiule charging cycle shifting calibrated on 2023 Belgian data. Green dotted line: strategic-bidding wholesale price for March 8, 2023 (illustrative day). Blue dotted line: baseline strategic-bidding wholesale hourly average over 2023.*Source:* ENTSO-E Transparency Platform; authors' calculations. Time-of-use prices can be higher since they are computed as load-weighted averages.

4.2 Stage 1: Consumer demand response

Given retail prices $\mathbf{p}^r$, the algorithm solves discrete choice problems for all flexible loads. The implementation proceeds in two steps: baseline profile construction and optimal reallocation.

Baseline profile construction. I first distribute the calibrated flexible load quantities across hours using observed consumption patterns under flat pricing from Elia's Adequacy Study. For electric vehicles, 79,282 daily charging cycles are allocated across the 24-hour day according to the baseline profile: if 12% of observed EV charging occurs at 18:00 under flat pricing, then $0.12 \times 79,282 = 9,514$ charging cycles are assigned to that hour. This yields the baseline flexible load profile $\{Q_t^{\text{EV,base}}\}_{t=1}^{24}$ representing where consumers charge absent price incentives.[7]

The observed total load in 2023 incorporates both this baseline flexible consumption and all other inelastic demand (lighting, appliances, industrial use). I decompose observed load as:

$$Q_t^{\text{observed}} = Q_t^{\text{base}} + Q_t^{\text{EV,base}} + Q_t^{\text{HP,base}} \tag{16}$$

where Q_t^{base} is inelastic baseline consumption, assumed fixed across all pricing regimes. Only flexible loads can shift in response to price signals; all other consumption remains constant.

[7]The same procedure distributes 209,000 heat pump cycles across hours using baseline heating patterns, yielding $\{Q_t^{\text{HP,base}}\}_{t=1}^{24}$.

Optimal reallocation. For each hour h containing baseline flexible load, the algorithm evaluates whether consumers initially charging at hour h would prefer to shift to an alternative starting time given retail prices $\mathbf{p}^r$. Each consumer at baseline hour h evaluates all feasible starting times $h' \in \mathcal{H}$ (those allowing the full consumption window to complete within the day) and selects the minimum-cost option. If the optimal starting time h^* differs from baseline h, that consumer's load reallocates from hour h to the window beginning at h^*.

This generates shifted flexible load profiles $\{Q_t^{\text{EV,shift}}(\mathbf{p}^r)\}$ and $\{Q_t^{\text{HP,shift}}(\mathbf{p}^r)\}$, which differ from baseline when price incentives induce reallocation. Total system demand becomes:

$$Q_t(\mathbf{p}^r) = Q_t^{\text{base}} + Q_t^{\text{EV,shift}}(\mathbf{p}^r) + Q_t^{\text{HP,shift}}(\mathbf{p}^r) \tag{17}$$

Under flat pricing (baseline regime), $\mathbf{p}^r$ is constant across hours, so no consumer has incentive to shift: $Q_t^{\text{EV,shift}} = Q_t^{\text{EV,base}}$ and $Q_t^{\text{HP,shift}} = Q_t^{\text{HP,base}}$ for all t, reproducing observed 2023 load. Under time-of-use or real-time pricing, price variation induces shifting: $Q_t^{\text{flex,shift}} \neq Q_t^{\text{flex,base}}$ for some hours t.

Participation rate scenarios. The analysis is parametized by participation rate of the calibrated availble shiftable load α. At 10%, only 7,928 active EV charging cycles shift daily (0.34 GWh); at 50%, 39,641 cycles shift (1.7 GWh). Results in Section 5 focus on the 20% case ($\alpha = 0.2$) as a central scenario, with 10% producing effects too small to identify system-level patterns and 30–50% likely overstating near-term behavioral response.

4.3 Stage 2: Wholesale market clearing

The wholesale clearing stage simulates strategic bidding in Belgium's oligopolistic market using the conduct parameter approach from Section 2.2. The conduct parameters are estimated parametrically from observed market outcomes to calibrate the degree of market power exercised by generators. This subsection describes implementation details.

Merit order construction. Net load at hour t equals total demand minus renewable generation:

$$Q_t^{\text{net}} = Q_t(\mathbf{p}^r) - \text{Wind}_t - \text{Solar}_t \tag{18}$$

where wind and solar output come from Elia's hourly generation data.

For each thermal unit, marginal cost is:

$$c_{g,j,t} = \frac{\text{Fuel Price}_{f(g),t}}{\text{Efficiency}_g} + \frac{\text{CO}_2 \text{ Factor}_g}{\text{Efficiency}_g} \times \text{EU ETS Price}_t \tag{19}$$

where $f(g)$ indexes fuel type (gas, coal, oil), thermal efficiencies come from JRC power plant data, emission factors from RTE technical documentation, fuel prices from TTF futures, and carbon prices from EU ETS allowances.

Under perfect competition, the market clears at the marginal cost of the last dispatched unit:

$$p_t^{\text{comp}} = \min \left\{ c_{g,j,t} : \sum_{(g',j'):\text{MC}_{g',j',t} \leq c_{g,j,t}} \bar{k}_{g',j'} \geq Q_t^{\text{net}} \right\} \tag{20}$$

For hours when renewable generation exceeds total demand ($Q_t^{\text{net}} \leq 0$), the conduct parameter calibration uses observed 2023 market outcomes to estimate firm behavior during these periods. The model does not explicitly generate negative wholesale prices but incorporates these hours in the strategic parameter estimation.

Strategic conduct calibration. The competitive merit order systematically underpredicts observed prices, reflecting strategic behavior. Three firms (Electrabel/ENGIE, Luminus, TotalEnergies) control 87.5% of dispatchable thermal capacity, creating substantial market power. To capture strategic conduct, I calibrate firm–market-condition markup parameters from observed price–cost gaps by constructing empirical residual-demand curves for each firm, which measure how much prices must rise for that firm to supply additional capacity.

I partition the 2023 net-load distribution into 20 equally spaced bins $\{\mathcal{B}_1, \ldots, \mathcal{B}_{20}\}$ spanning the observed range, thereby defining distinct market situations.[8] For each firm–bin pair (j, b), I (i) collect hours with net load in $\mathcal{B}_b$; (ii) assemble competitors' supply at marginal cost; (iii) recover firm j's residual demand; (iv) estimate the local inverse slope $\theta_{j,b} = dp/dq$ at observed clearing prices via finite differences; and (v) average within bin.

The calibration yields 340 parameters (17 firms $\times$ 20 bins), of which 187 are nonzero. Table 9 summarizes the distribution. Median markup coefficients are 0.36 cents per MWh of infra-marginal capacity. Market power concentrates in mid-load bins (8–14), where thermal capacity is neither slack nor fully exhausted. Details about the parameters values and estimation are provided in Appendix D

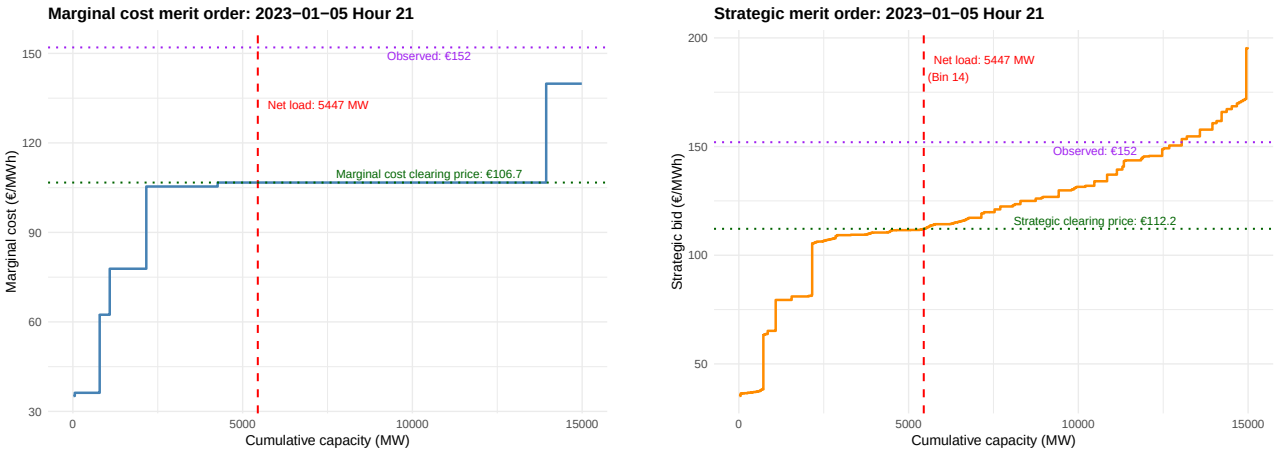
[8]This approach is inspired by Holland and Mansur (2008) nonparametric classification of demand using joint deciles of mean daily and hourly load. In this paper, it allows to avoid hour-specific calibration and mitigates overfitting.

Strategic bid construction and market clearing. The calibrated parameters enter through modified bid functions. Firm j's bid for unit g at hour t is:

$$B_{g,j,t} = \max\left\{c_{g,j,t} - \theta_{j,b(t)} \cdot K^{\text{infra}}_{g,j,t},\ c_{g,j,t}\right\} \tag{21}$$

where $\theta_{j,b(t)} < 0$ is the inverse residual demand slope for net load bin $b(t)$, and $K^{\text{infra}}_{g,j,t}$ denotes cumulative capacity from firm j with lower marginal costs than unit g.

Figure 5 illustrates these effects for a representative hour. The left panel shows marginal cost ordering under perfect competition; the right incorporates strategic markups.



(a) Marginal cost merit order

(b) Strategic bid merit order

Figure 5: Merit order construction (5 January 2023, 21:00)

Note: Illustrative example of merit order use for the simulation of Belgian electricity wholesale market. Left panel shows generation technology and fuel marginal costs. Right panel incorporates calibrated strategic markups used for the simulaitons. *Source:* Author's computation and Elia installed capacity by unit data.

Given strategic bid functions $\{B_{g,j,t}\}$ and net load Q^{net}_t, the market clears independently each hour. Each firm submits a stepwise supply curve:

$$S_j(p;t) = \sum_{g \in \mathcal{G}_j} \bar{k}_{g,j} \cdot \mathbf{1}\{B_{g,j,t} \leq p\} \tag{22}$$

The system operator determines the minimum price satisfying net load:

$$p^w_t = \min\left\{p : \sum_{j=1}^{M} S_j(p;t) \geq Q^{\text{net}}_t\right\} \tag{23}$$

The calibrated conduct parameters $\{\theta_{j,b}\}$ remain fixed across counterfactuals, conditional on net load bin.

Validation. Table 3 compares predicted strategic-bidding prices to observed day-ahead prices. In-sample fit for 2023 is strong (Spearman 0.61; RMSE €42/MWh). Out-of-sample performance is solid for 2020 and weak for 2021–2022, especially during the 2022 crisis when the model underpredicts by €138/MWh on average. The evidence supports using the calibrated model for counterfactuals under conditions similar to 2023 and cautions against extrapolation to extreme price environments.

Table 3: Strategic bidding model validation (2020–2023)

Year	N	Pearson	Spearman	RMSE (€/MWh)	MAE (€/MWh)	Bias (€/MWh)
2020 (out-of-sample)	8,760	0.60	0.62	34	29.39	29
2021 (out-of-sample)	8,760	0.25	0.45	76	57.89	0.29
2022 (out-of-sample)	8,760	0.21	0.20	190	149	-138
2023 (in-sample)	8,760	0.45	0.61	42	32	7

Notes: Statistics validate a wholesale market model with strategic bidding. Firm-specific strategic parameters are calibrated by net-load bin using 2023 data; 2020–2022 are evaluated out-of-sample and 2023 in-sample. Correlations compare predicted to observed day-ahead prices. RMSE = root mean squared error; MAE = mean absolute error; Bias = mean(predicted − observed).

The model closely replicates 2023 outcomes and performs reasonably in 2020, indicating that the conduct calibration captures structural features of the Belgian market. Performance degrades in 2021–2022 due to crisis-driven shocks. Counterfactual ToU results are therefore interpreted as effects under conditions akin to 2023.

Convergence and computational notes Convergence is defined solely on the optimal partition. The algorithm terminates once the optimal design remains unchanged across iterations, i.e. $\mathcal{P}^{(n)} = \mathcal{P}^{(n-1)}$ for two consecutive iterations. A stable partition implies stable relative price differences across feasible charging windows. Because flexible consumption responds only to these relative differences, a fixed $\mathcal{P}$ induces the same load-shifting pattern, which in turn yields the same wholesale clearing outcome. Once the partition is invariant, the system has reached a fixed point of the retail design, demand response, and market clearing operators.

5 Results

To evaluate the relative efficiency of real-time and time-of-use pricing at equilibrium, I simulate the iterative algorithm described in Section 4 until convergence. The simulationspecification assumes 20% participation among consumers with electric vehicles who can shift loads, with

strategic bidding calibrated to 2023 Belgian market data. The simulation results should be interpreted as illustrative counterfactual outcomes under the maintained assumptions of the model rather than as precise quantitative predictions. I present five pricing regimes: flat-rate pricing (status quo), real-time pricing (RTP), optimized two-period time-of-use (ToU2), optimized three-period time-of-use (ToU3), and optimized four-period time-of-use (ToU4). I first compare system costs and consumer surplus across pricing regimes as my main results. I then examine impacts on total load and peak demand, wholesale price adjustments, and distributional implications for generators.

5.1 Welfare and system costs

Tables 10 and 11 present equilibrium simulation outcomes across four pricing regimes. The baseline reflects observed 2023 Belgian market conditions under flat-rate retail pricing, where consumers face no incentive to shift flexible loads.

Supply-side efficiency. In Table 10, real-time and time-of-use pricing both achieve substantial reductions in equilibrium generation costs relative to flat-rate pricing. Real-time pricing reduces total generation costs by €999 million (16.0%), lowering average supply costs from €143/MWh to €123/MWh. Remarkably, two-period and three-period time-of-use pricing achieve nearly identical and marginally superior outcomes: €1,023 million in savings (16.4%), with the same average costs of €123/MWh. The absolute cost difference between real-time pricing and ToU2 or ToU3 is €7.8 million, representing 0.2 percentage points. Four-period time-of-use performs similarly at €1,016 million in savings (16.2%).

The marginal superiority of ToU2 and ToU3 over real-time pricing, while counterintuitive, reflects the interaction between consumers' time preferences and time-of-use desings' price block structure. Consider a consumer who prefers evening charging. Under real-time pricing, the cheapest hour may occur in the early morning, imposing substantial time-preference costs. Under time-of-use pricing, the off-peak block averages prices across multiple hours, and the closest low-price window to the consumer's preferred timing may offer a superior trade-off between cost minimization and scheduling convenience. This mechanism can induce marginally greater aggregate load shifting, reducing equilibrium system costs despite coarser price signals.

More fundamentally, the near-equivalence across pricing schemes arises because all induce consumers to shift flexible loads from high-cost peak hours to low-cost off-peak periods. Since consumers respond to relative price rankings when optimizing consumption timing (equa-

tion 1), a two-period structure conveys sufficient information to replicate real-time pricing's load-shifting incentives. Adding a third or fourth period yields negligible additional efficiency gains: ToU3 saves only €1 million more than ToU2, while ToU4 performs slightly worse than both. [9] Discretizing real-time hourly prices into as few as 2 daily blocks imposes minimal efficiency costs at equilibrium, with additional price granularity offering sharply diminishing returns. This finding aligns with Astier (2021), who document that statistical fit metrics flatten as the number of time-of-use periods increases beyond two or three.

Table 4: Equilibrium system costs by pricing regime from simulation calibrated on Belgian electricity markets in 2023

	Supply-side efficiency		
Pricing regime	Average cost (€/MWh)	System cost (M€)	Δ vs. baseline %
Baseline (flat)	143	6,247.6	—
Real-time pricing (RTP)	122.61	5,232.6	−16.2
ToU2	122.70	5,224.8	−16.4
ToU3	122.70	5,224.8	−16.4
ToU4	122.66	5,233.4	−16.2
Efficiency comparisons:			
ToU2 or ToU3 − RTP	≈ 0	−7.8	−0.2 pp
ToU4 − RTP	≈ 0	−0.8	0 pp

Notes: Baseline (flat) has no iteration. **Real-time pricing (RTP)** sets retail prices in each iteration equal to the wholesale prices from the previous iteration; only Stages 1–2 iterate to convergence (see Fig. 3). **Time-of-use (ToUF)** for $F = 2, 3, 4$ optimizes a F-block schedule each iteration to best match the *relative* hourly wholesale price profile from the previous iteration. *System cost* is total costs of serving the demand at market clearing price; *Average cost* is system cost divided by total energy served. Results shown at convergence after three iterations with 20% participation of estimated electric vehicle daily cycles. Percentage changes are relative to the baseline; comparison rows report percentage-point differences. ToU2 and ToU3 yield identical efficiency outcomes. Source: refer to Section 3.

Consumer welfare. Table 11 presents two measures of consumer welfare. Expenditure calculations multiply retail prices by total load (baseline or shifted) in each hour, treating aggregate consumption as a single representative consumer who responds optimally to prices. This approach masks important heterogeneity: in the simulation, only the 20% of consumers owning electric vehicles optimize daily charging cycles based on time preferences and prices, while the remaining 80% face altered retail prices without adjusting consumption. Consequently, the reported aggregate savings and welfare gains concentrate among price-responsive

[9] This result is similar to Cabot and Villavicencio (2024) who find that the simple two-period pricing is almost as well correlated with wholesale prices as the more complex pricing.

consumers, while non-responsive consumers experience only expenditure changes. Consumer surplus is computed assuming linear hourly demand using the trapezoidal approximation: $\Delta CS = -0.5 \times (Q_b + Q_t) \times (P_t - P_b)$, where Q_b and Q_t denote baseline and shifted quantities, and P_b and P_t denote baseline and treatment prices.

Under flat-rate pricing, aggregate annual expenditures total €10,226 million, where the flat rate equals the load-weighted average wholesale price. Real-time pricing reduces aggregate expenditures to €8,244 million, generating €1,979 million in consumer surplus. ToU2 and ToU3 produce nearly identical outcomes: expenditures of €8,246 million and surplus of €1,978 million, differing from real-time pricing by €2.6 million in expenditures and €1.4 million in surplus: less than 0.07 percentage points. ToU4 performs marginally worse but remains within 0.06 percentage points of real-time pricing.

The welfare equivalence between real-time pricing and time-of-use schemes follows mechanically from their similar impacts on system costs. Both induce comparable load-shifting patterns, generating similar reductions in wholesale clearing prices that pass through to retail expenditures. The slightly higher expenditure under ToU2 and ToU3 (€2.6 million) reflects two sources of inefficiency. First, consumers face block-average prices rather than exact hourly marginal costs, creating minor within-block misallocations. Second, time-of-use pricing cannot adapt to exceptional price events—such as renewable generation shortfalls (Dunkelflaute) or seasonal variations—forcing consumers to pay fixed block prices during periods when real-time prices would signal scarcity or abundance. Both effects are economically negligible. The near-equivalence in consumer welfare mirrors supply-side results: coarse pricing conveys sufficient information for consumers to capture essentially all gains from demand response.

Equilibrium efficiency and welfare cost of price simplification. These results quantify the pure informational loss from simplifying retail price signals while holding consumer response at perfect rationality. The finding that two-period or three-period time-of-use pricing captures 99.93% of real-time pricing's consumer surplus gains, while actually exceeding real-time pricing in system cost reductions, demonstrates that price granularity is not the binding constraint on demand flexibility, provided pricing structures preserve the rank ordering of costs across hours.[10]

This contrasts with prior literature documenting substantial efficiency losses from time-of-use pricing. Existing studies employ two measurement approaches. The first uses statistical fit

[10]ToU2 and ToU3 generate €1,977.5 million in consumer surplus compared to €1,978.9 million under real-time pricing, yielding 99.93% recovery (1,977.5/1,978.9). On system costs, ToU2 and ToU3 save €1,023 million versus €999 million under real-time pricing: a 2.4% improvement.

Table 5: Equilibrium consumer expenditure and surplus by pricing regime from simulation calibrated on Belgian electricity markets in 2023

	Consumer savings and welfare		
Pricing regime	Expend. (M€)	Surplus (M€)	Δ vs. baseline %
Baseline (flat)	10,225.6	—	—
Real-time pricing (RTP)	8,243.8	1,978.9	—
ToU2	8,246.4	1,977.5	−0.07
ToU3	8,246.4	1,977.5	−0.07
ToU4	8,245.1	1,977.8	−0.06
Efficiency comparisons:			
ToU2 or ToU3 − RTP	+2.6	−1.4	−0.07 pp
ToU4 − RTP	+1.3	−1.4	−0.06 pp

Notes: Baseline (flat) has no iteration. **Real-time pricing (RTP)** sets retail prices in each iteration equal to the wholesale prices from the previous iteration; only Stages 1–2 iterate to convergence (see Fig. 3). **Time-of-use (ToUF)** for $F = 2, 3, 4$ optimizes a F-block schedule each iteration to best match the *relative* hourly wholesale price profile from the previous iteration. Expenditure is total retail payments across all consumers. Surplus measures welfare gains relative to baseline, computed assuming linear demand assumption. Results shown at convergence after three iterations with 20% participation of estimated electric vehicle daily cycles. Percentage changes are relative to the baseline; comparison rows report percentage-point differences. ToU2 and ToU3 yield identical efficiency outcomes. Source: refer to Section 3.

measures: Jacobsen et al. (2020) estimate 13–43% efficiency recovery using R^2. This methodology mechanically penalizes time-of-use for rare extreme price events, even when rank orderings remain intact. The second uses rank correlation: Schittekatte et al. (2024) estimate 60–70% recovery using Spearman's correlation, while Cabot and Villavicencio (2024) find 25–50%. Although they better capture relative price signals, these designs confound information costs from discretization with attention-driven behavioral costs.

Except for Cabot and Villavicencio (2024) who model one feedback round, these studies ignore how retail pricing feeds back to wholesale markets. This paper addresses both gaps by isolating pure discretization costs—assuming equal consumer attention—while solving for *complete equilibrium*. The near-zero efficiency loss from discretization implies observed real-time pricing underperformance can reflect rational inattention to complexity rather than fundamental informational constraints.

5.2 Load profiles and peak reduction

Figure 10 shows how alternative retail pricing reshapes hourly electricity consumption across simulation iterations. The figure reveals convergence patterns and the timing of load shifts

under each pricing regime.

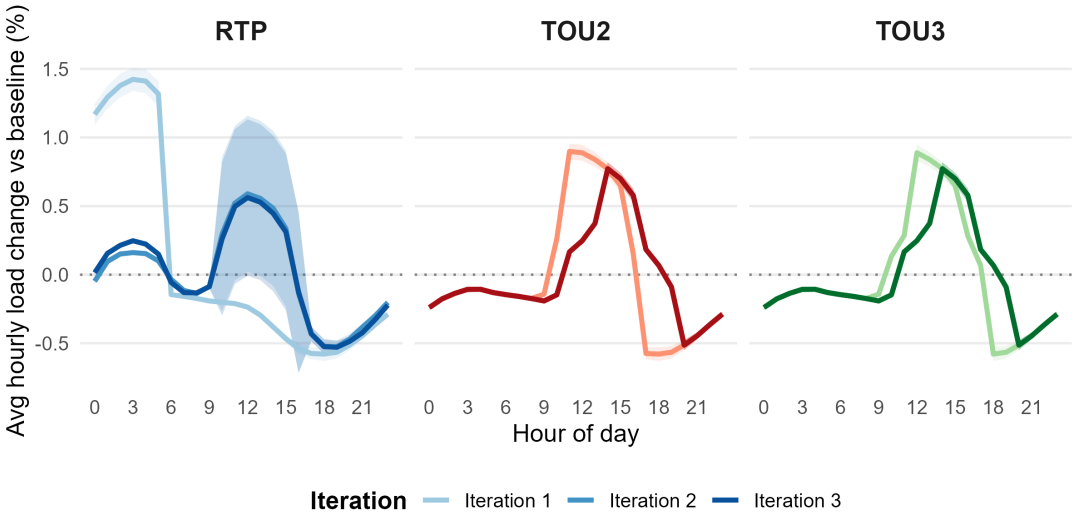


Figure 6: Percentage change in hourly load relative to baseline (simulation Belgian electricity markets in 2023)

Note: Average hourly load changes relative to a flat-pricing baseline are shown for each iteration under real-time pricing (RTP), two-period ToU (ToU2), and three-period ToU (ToU3). Iteration 1 is the initial consumer response; later iterations update wholesale prices (RTP) or the ToU partition (ToU2/ToU3) until convergence. Positive values mean load shifts toward that hour; negative values mean reductions. Results average all days with 20% of eligible flexible loads (EV) participating; shaded bands depict iteration-to-iteration adjustments.

Timing of load shifts. At equilibrium, all three retail pricing schemes shift electric vehicule charging away from evening peaks (18:00–22:00) toward lower-cost hours, though the timing of this shift differs between real-time and time-of-use pricing. Under real-time pricing, adjustments operate on three margins: (i) **peak reduction**: evening load declines by 0.4–0.6% as EV charging shifts away from post-work hours when wholesale prices average €125/MWh; (ii) **nighttime increase**: load rises by 0.6–1.2% during hours 1–6 when wholesale prices fall to €75/MWh; and (iii) **midday increase**: load rises by 0.4–0.6% during hours 11–15. Under ToU2 and ToU3, load shifts concentrate within their designated off-peak blocks, producing sharper increases during specific hours rather than the more dispersed pattern under real-time pricing.

Convergence of load profiles across pricing schemes. At equilibrium, all three schemes achieve similar peak reductions and off-peak increases, but real-time pricing exhibits substantially greater volatility. The shaded confidence intervals in Figure 10 reveal that real-time pricing induces wider day-to-day and hour-to-hour fluctuations in load adjustments, as consumers respond to granular price signals that vary with daily supply and demand conditions.

Time-of-use pricing generates more stable patterns by smoothing hourly variations into block-average prices.

The simulated convergence in average load shifts reflects the rank-preservation property of the time-of-use optimization. Because ToU2 and ToU3 are designed to correctly identify which hours are cheap versus expensive, the consumer load-shifting program generates similar consumption decisions under time-of-use as under real-time pricing.

Peak load reduction. Table 6 shows minimal peak reduction across all pricing schemes. The annual system peak declines by only 10 MW, and real-time pricing modestly reduces the top 100 hours by 29 MW, while time-of-use pricing leaves extreme hours essentially unchanged.

The limited peak effects reflect two constraints. First, with only 20% participation among EV owners, flexible loads are too small to materially affect system-wide peaks. Second, time-of-use pricing cannot adapt to seasonal or short-term supply shocks. Daily peak metrics illustrate this rigidity: summer weekday peaks increase under all schemes (+0.2% to +0.9%), as fixed time-of-use blocks shift load toward midday hours that coincide with summer air conditioning demand. Real-time pricing exhibits the largest summer peak increase (+0.8%), redirecting charging toward solar-abundant midday hours when instantaneous prices fall but aggregate demand remains elevated. Winter peaks remain flat across all schemes. These results clarify that the 16% system cost reduction documented in Table 10 operates through reallocating demand toward low-marginal-cost hours rather than through peak shaving or capacity deferral.

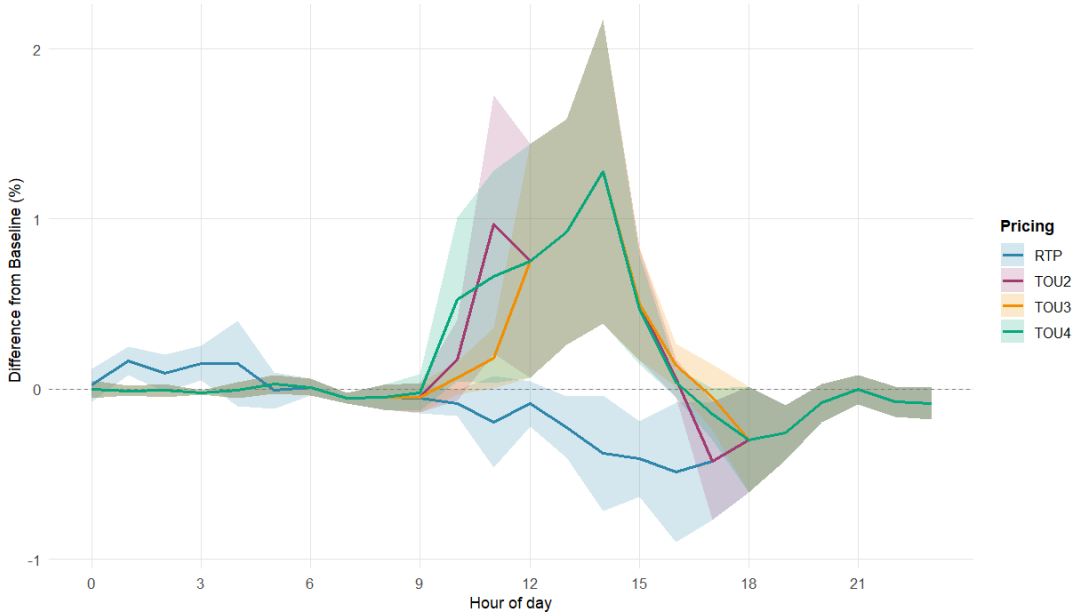
Table 6: Peak load metrics by pricing regime

		Change vs. baseline			
Metric	Baseline	RTP	ToU2	ToU3	ToU4
Annual peak metrics					
Annual peak load (MW)	13,932	-10 (-0.1%)	-10 (-0.1%)	-10 (-0.1%)	-10 (-0.1%)
Average of top 100 hours (MW)	12,193	-29 (-0.2%)	+10 (+0.1%)	+10 (+0.1%)	+13 (+0.1%)
Average of top 10% hours (MW)	11,425	-17 (-0.2%)	+12 (+0.1%)	+12 (+0.1%)	+14 (+0.1%)
Average daily peak metrics					
Daily peak - weekday summer (MW)	9,485	+75 (+0.8%)	+19 (+0.2%)	+19 (+0.2%)	+82 (+0.9%)
Daily peak - weekday winter (MW)	11,502	-24 (-0.2%)	+9 (+0.1%)	+9 (+0.1%)	+8 (+0.1%)

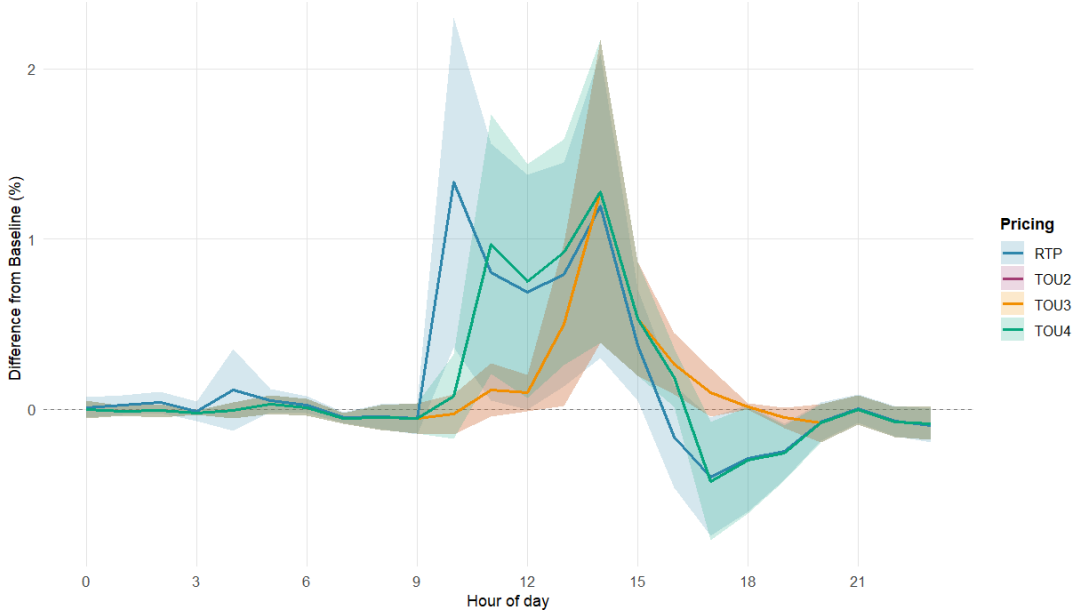
Notes: Annual peak is the maximum hourly load over the year. Top 100 hours approximate the 99th percentile. Daily peak metrics show average of daily maximum loads by day type and season. Winter: Dec-Feb; Summer: Jun-Aug. Changes are counterfactual minus baseline.

5.3 Wholesale price impacts

Equilibrium price effects of demand response. Figure 11 shows percentage changes in model-generated wholesale prices under the four retail pricing designs relative to a model-



(a) Percentage change in hourly wholesale prices relative to baseline - iteration 1



(b) Percentage change in hourly wholesale prices relative to baseline - iteration 3

Figure 7: Wholesale market price impacts of dynamic retail pricing for electric vehicule residential charging load shifting

Notes: Percentage changes in wholesale clearing prices under real-time pricing scenarios relative to a counterfactual baseline, averaged across 2023 weekdays. Panel (a): iteration 1; panel (b): iteration 3 (convergence). All wholesale prices are endogenously determined using the market clearing model in Section 2.2. The baseline represents a status quo flat retail pricing and no demand response, solved using the same wholesale market model to ensure comparability. Shaded regions show 95% confidence intervals across days. This is a controlled counterfactual comparison between model equilibria, not a comparison to observed historical prices.

generated flat-rate baseline. Both prices are outcomes of the wholesale market-clearing model (Section **??**), ensuring a controlled model-to-model comparison. Due to simplifications in the day-ahead wholesale electricity market model, model-based prices do not capture the full extent of power price volatility observed in historical data.

Load-shifting out of the evening peak (hours 17–20) is consistent across all retail pricing designs, reducing wholesale prices during these hours at convergence (panel b). Effects scale with pricing granularity: real-time pricing produces the largest impacts, followed by ToU2, ToU3, and ToU4. All four retail designs generate nearly identical equilibrium price impacts despite their different temporal structures. Wide confidence bands reflect price spikes during supply-constrained periods and seasonal variation in demand and generation patterns. The system converges to stable equilibrium patterns (panel b) as the wholesale-retail feedback stabilizes.

Implications for price volatility. Table 7 quantifies the effects of the four retail pricing designs on wholesale price variability. The rationale for time-varying retail pricing rests on exploiting wholesale price volatility: consumers shift demand from high-price to low-price hours, capturing arbitrage gains. However, this mechanism contains a fundamental tension: successful demand response reduces the very price variation it seeks to exploit. Our results confirm this "benefit cannibalization" effect through the compression of daily price variability.

Table 7: Wholesale price volatility metrics

		Change vs. baseline			
Metric	**Baseline**	**RTP**	**ToU2**	**ToU3**	**ToU4**
Within-day volatility (€/MWh):					
Mean daily std. deviation	19.6	−1.5%	−0.5%	−0.5%	−1.0%
Mean peak-to-trough	50.6	−1.2%	−0.6%	−0.6%	−1.2%
Overall volatility (€/MWh):					
Std. deviation (all hours)	28	−0.4%	−0.0%	0.0%	−0.4%

Notes: Within-day volatility metrics calculated separately for each day of 2023, then averaged across 365 days. Peak-to-trough measures maximum minus minimum price within each day. Overall volatility includes seasonal patterns and price spikes. All prices are model outcomes; comparisons are between model equilibria under alternative retail pricing designs, not between model and observed data. Changes are counterfactual minus baseline.

Real-time pricing reduces mean daily standard deviation by 1.5%, while time-of-use pricing achieve roughly half this compression (0.5–1.0%). Peak-to-trough spreads exhibit similar modest reductions of 0.6–1.2%. RTP's superior performance reflects two design features. First,

hourly prices allow consumers to respond to seasonal demand patterns and supply-constrained price spikes that time-of-use pricing aggregate away within their fixed period definitions. Second, the simulation assumes consumers perfectly understand and respond to price signals; in practice, the complexity of hourly price variation may reduce responsiveness, narrowing the advantage of RTP over simpler time-of-use structures. The limited overall impact—flexible load comprises only 3.2% of total energy under 20% EV participation—suggests that even substantially higher penetration rates will compress intraday spreads without eliminating the fundamental volatility driven by weather, outages, and fuel price shocks.

5.4 Limitations of results

The simulations provide useful estimates of efficiency and welfare gains from real-time and time-of-use pricing but abstract from several real-world complexities. First, I do not model supply and demand uncertainty, neither probabilistic generator outages nor stochastic demand shocks. In practice, system operators manage these uncertainties almost entirely through supply-side adjustments: ramping dispatchable plants or activating reserves. Second, for clarity of exposition, I present load-shifting results for electric vehicles and heat pumps separately, though joint optimization across both end uses would yield additional gains. Similarly, the model can optimize time-of-use pricing by season, but I omit these results for brevity; they are available upon request. Third, the analysis assumes regulators can implement optimal time-of-use structures directly, abstracting from retail market competition and supplier incentives. Retailers face different procurement costs, hedging strategies, and profit margins that may create resistance to regulatory guidance on period structures. A complete welfare assessment would require modeling how consumers respond to unexpected supply shortages or demand surges, as well as how retail market dynamics affect the translation of optimal pricing designs into actual consumer contract offerings, complexities my framework cannot fully capture.

6 Conclusion

This paper evaluates the efficiency of real-time versus time-of-use retail electricity pricing at partial equilibrium, accounting for feedback between consumer demand response and wholesale market clearing. Using calibrated simulations of the Belgian electricity market with electric vehicle charging as the primary flexible load, I find that simple two-period and three-period time-of-use pricing captures almost 100% of real-time pricing's consumer welfare gains. At

convergence, two-period time-of-use pricing reduces generation costs by 16.4% compared to 16.0% under real-time pricing—a difference of 0.2 percentage points. Consumer surplus gains differ by less than 0.07 percentage points across pricing designs.

These results challenge the conventional wisdom that pricing granularity determines demand response efficiency. The near-equivalence arises because consumers optimize based on relative price rankings rather than absolute levels: a two-period structure that correctly identifies which hours are cheap versus expensive conveys sufficient information to replicate real-time pricing's load-shifting incentives. Adding a third or fourth pricing period yields negligible additional gains. This finding contrasts sharply with prior literature documenting 30-75% efficiency losses from time-of-use pricing, estimates that confound pure discretization costs with behavioral inattention to complex signals or rely on statistical fit metrics that mechanically penalize time-of-use for rare price spikes.

The key methodological contribution lies in simulating partial equilibrium where retail prices affect wholesale clearing and vice versa. Previous studies either ignore this feedback or model only one iteration, potentially overstating real-time pricing's advantages. At equilibrium, demand response compresses wholesale price volatility, the very variation that dynamic pricing exploits, creating a cannibalization effect. Real-time pricing reduces mean daily price standard deviation by 1.5%, while time-of-use achieves roughly half this compression. Yet this difference translates to minimal efficiency losses because time-of-use block structures preserve hour rankings even as absolute spreads narrow.

The welfare analysis assumes perfect consumer rationality and full attention to price signals, isolating the pure informational value of granularity. This abstraction provides an upper bound on achievable gains. Incorporating behavioral costs—cognitive burden of tracking hourly prices, search frictions, or inattention documented in empirical studies—would likely favor simpler time-of-use structures, potentially reversing real-time pricing's marginal advantage. The policy implication is clear: regulators should prioritize simple, transparent pricing designs over complex real-time pricing, particularly when targeting mass-market adoption.

References

Astier, N. (2021). Second-best pricing for incomplete market segments: Application to electricity pricing. *Journal of Public Economic Theory*, 23(6):1287–1311.

Bailey, M., Brown, D. P., Shaffer, B., and Wolak, F. A. (2024). Centralized vs decentralized demand response: Evidence from a field experiment. Technical report, Working Paper.

Boiteux, M. (1949). La tarification des demandes en pointe: Application de la theorie de la vente au cout marginal. *Revue Generale de l'Electricite*, 58.

Borenstein, S. (2005). The long-run efficiency of real-time electricity pricing. *The Energy Journal*, 26(3):93–116.

Byrne, D. P. and Martin, L. A. (2025). Behavior, automation, and the energy transition. *Available at SSRN 5459934*.

Cabot, C. and Villavicencio, M. (2024). The demand-side flexibility in liberalised power market: A review of current market design and objectives. *Renewable and Sustainable Energy Reviews*, 201:114643.

Elia (2023). Electricity mix 2023: Belgian renewable generation reaches all-time high while consumption declines. Technical report, Elia Group. Accessed: 6 January 20255.

Elia (2024). Adequacy and flexibility study for belgium 2025-2034. Technical report, Elia Group. Accessed: 6 January 2025.

Enrich, J., Li, R., Mizrahi, A., and Reguant, M. (2024). Measuring the impact of time-of-use pricing on electricity consumption: Evidence from spain. *Journal of Environmental Economics and Management*, 123:102901.

Fabra, N., Rapson, D., Reguant, M., and Wang, J. (2021). Estimating the elasticity to real-time pricing: Evidence from the spanish electricity market. *AEA Papers and Proceedings*, 111:425–429.

Gils, H. C., Pfluger, B., and Reiter, U. (2024). Policy scenarios for the decarbonisation of residential buildings and heating: Update of the sectoral decarbonisation pathway for space heating in the eu. Technical report, Fraunhofer Institute for Systems and Innovation Research ISI. Accessed: 2025-06-09.

Hinchberger, A. J., Jacobsen, M. R., Knittel, C. R., Sallee, J. M., and van Benthem, A. A. (2024). The efficiency of dynamic electricity prices. Technical report, National Bureau of Economic Research.

Hogan, W. W. (2014). Time-of-use rates and real-time prices. *John F. Kennedy School of Government, Harvard University*.

Holland, S. P. and Mansur, E. T. (2006). The short-run effects of time-varying prices in competitive electricity markets. *The Energy Journal*, 27(4).

Holland, S. P. and Mansur, E. T. (2008). Is real-time pricing green? the environmental impacts of electricity demand variance. *The Review of Economics and Statistics*, 90(3):550–561.

IEA (2024a). Europe annual heat pump sales, 2013-2023. Charts – Data & Statistics. Accessed: October 30, 2025.

IEA (2024b). Global EV data explorer. Data Tools. Accessed: October 30, 2025.

International Energy Agency (2023). World energy outlook 2023. Accessed: 2025-06-09.

Ito, K. (2014). Do consumers respond to marginal or average price? evidence from nonlinear electricity pricing. *American Economic Review*, 104(2):537–563.

Jacobsen, M. R., Knittel, C. R., Sallee, J. M., and Van Benthem, A. A. (2020). The use of regression statistics to analyze imperfect pricing policies. *Journal of Political Economy*, 128(5):1826–1876.

Jessoe, K. and Rapson, D. (2014). Knowledge is (less) power: Experimental evidence from residential energy use. *American Economic Review*, 104(4):1417–1438.

Joint Research Centre (JRC) (2024). Belgium: Status of the heat pump market. Country Fiche JRC137131, European Commission. Accessed: 2025-06-09.

Joskow, P. L. and Wolfram, C. D. (2012). Dynamic pricing of electricity. *American Economic Review*, 102(3):381–385.

Nolting, L. and Praktiknjo, A. (2019). Techno-economic analysis of flexible heat pump controls. *Applied Energy*, 238:1417–1433.

Pébereau, C. and Remmy, K. (2023). Barriers to real-time electricity pricing: Evidence from new zealand. *International Journal of Industrial Organization*, 89:102979.

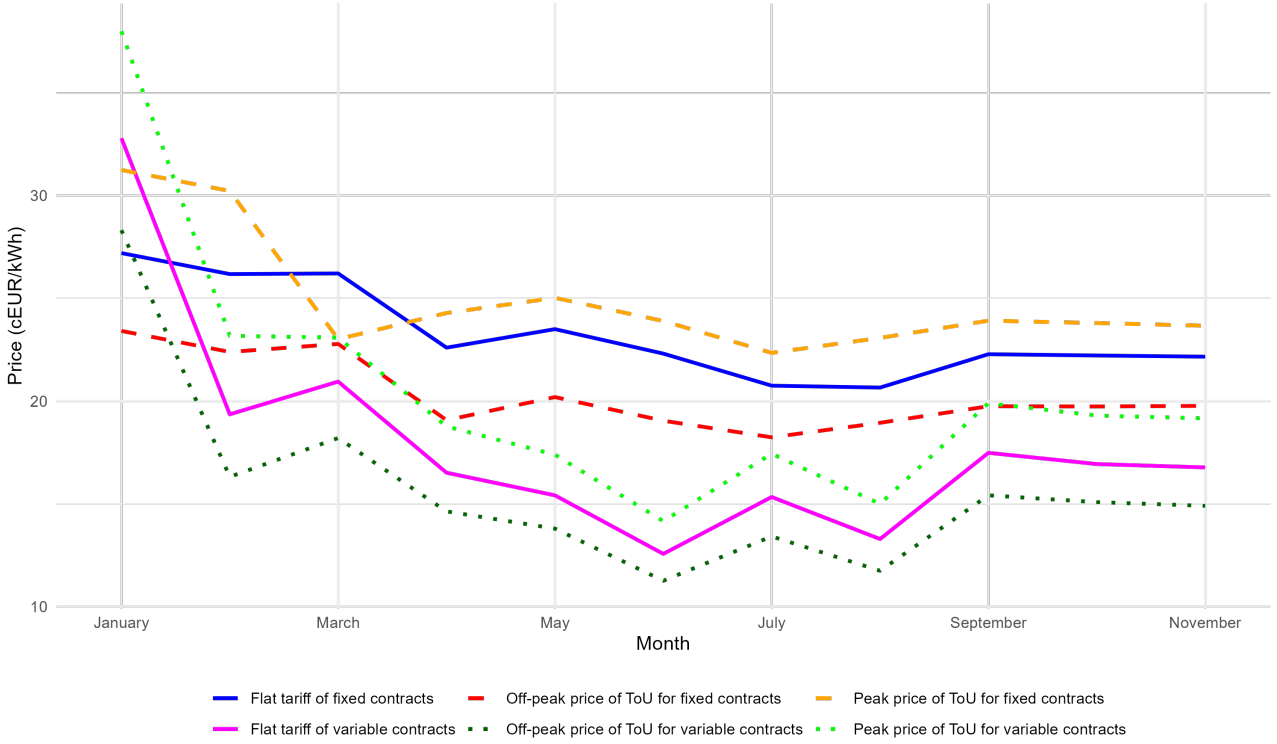
Poletti, S. and Wright, J. (2020). Real-time pricing and imperfect competition in electricity markets. *The Journal of Industrial Economics*, 68(1):93–135.

Rangaraju, S., De Vroey, L., Messagie, M., Mertens, J., and Van Mierlo, J. (2015). Impacts of electricity mix, charging profile, and driving behavior on the emissions performance of battery electric vehicles: A belgian case study. *Applied Energy*, 148:496–505.

Schittekatte, T., Mallapragada, D., Joskow, P. L., and Schmalensee, R. (2024). Electricity retail rate design in a decarbonizing economy: an analysis of time-of-use and critical peak pricing. *The Energy Journal*, 45(3):25–56.

A Appendix: Retail prices Belgium

Figure 8: Average retail electricity prices by contract type in 2023 in Belgium - prices displayed for contract entry



Note: This graph represents the weighted average retail electricity prices for residential contracts in Belgium in 2023. The data was collected from the monthly "cartes tarifaires" section of the MonEnergie comparator website, which provides historical electricity contract prices for all suppliers and contract types, consistent across all regions. For months where data was missing at the supplier level, the price from the previous month was used. If data was missing for more than two consecutive months, linear interpolation was applied to estimate the missing values. To ensure the representativeness of the prices, weighted averages for each contract type were computed using the yearly market shares of suppliers, as provided annually by the Belgian energy regulator (CREG). The complete table of the gathered data from 2019 to 2023 is available in the appendix for general knowledge, but it is not further used in the analysis to maintain consistent comparisons with the counterfactual time-of-use designs used.

B Appendix: Pearson and spearman correlation for time-of-use design (numerical illustration)

This example considers a 5-hour window and three 2-block time-of-use designs, with each block priced at the load-weighted average of wholesale prices. It shows that optimizing on pearson correlation (even on daily data) can favor a time-of-use that sends a poor timing signal, whereas spearman correlation correctly selects the design that preserves the cheap-to-expensive ordering.

Table 8: Numerical example: three candidate 2-block time-of-use designs

Hour	Load	Wholesale	R_W	ToU_A	R_A	ToU_B	R_B	ToU_C	R_C
1	2	1	1	4.3	1	3.4	1	7.5	3
2	3	3	2	4.3	1	3.4	1	7.5	3
3	4	7	4	4.3	1	8.0	2	5.3	1
4	5	4	3	8.4	2	8.0	2	5.3	1
5	6	12	5	8.4	2	8.0	2	7.5	3
			Pearson	0.555		**0.726**		0.667	
			Spearman	0.577		**0.866**		−0.289	

Note: Note: Load is hourly demand; Wholesale is the hourly wholesale price. ToU_A, ToU_B and ToU_C are alternative two-block time-of-use designs: each assigns two constant price levels to disjoint hour sets; the level in each block equals the load-weighted average wholesale price over that block. R_A, R_B and $R_C$are hourly ranks (1 = cheapest) computed within this five-hour example for Wholesale and for each time-of-use design. Pearson and Spearman are the period correlations between Wholesale and each time-of-use schedule; bold marks the highest alignment among the candidates.

Design ToU_B cleanly separates cheap and expensive hours and has the highest spearman (0.866) and should be preferred for price signal timing clarity. Design ToU_C has a moderately high Pearson (0.667), however for the design the ordering is reversed: the spearman correlation is negative (-0.289). ToU_C design is a misleading signal. This illustrates that optimizing on time-of-use design with Pearson correlation can favor designs that do not preserve the wholesale prices' order. Pearson correlation measures whether prices move together in levels, but this is ill-suited to step-wise time-of-use design where within each block, the time-of-ise price is fixed, while wholesale prices still fluctuate. Therefore, optimizing on spearman correlation directly selects the partition that provides clear signal on *when* to consume.

C Appendix: Time-of-use optimisation

For each candidate number of daily periods $F \in \{2, \ldots, 5\}$, the regulator solves the following optimization problem. The algorithm takes as input a distribution of wholesale prices: either historical data for the initial iteration, or simulated prices from the updated model in subsequent iterations of the full estimation procedure.

Algorithm 1 Optimal design of a time-of-use design for a fixed number of blocks F

for each $F \in \{3, 4, 5\}$ **do**

Input: Historical wholesale prices and load $\{p^w_{t,d}, Q_{t,d}\}_{t=1..24, d\in\mathcal{D}}$. The regulator treats this as the prior distribution of wholesale prices (updated in later iterations of the full equilibrium simulation).

Generate the set of feasible partitions $\mathcal{P} = \{P_1, \ldots, P_F\}$ of $\{1, \ldots, 24\}$ into F contiguous blocks such that $|P_f| \geq 4$ for all f.

for each partition $\mathcal{P}$ in this set **do**

For each block P_f, compute the time-of-use price

$$p^r_f = \frac{\sum_{d\in\mathcal{D}} \sum_{t\in P_f} p^w_{t,d} Q_{t,d}}{\sum_{d\in\mathcal{D}} \sum_{t\in P_f} Q_{t,d}}.$$

Construct the induced hourly retail price vector $\mathbf{p}^r(\mathcal{P}) = (p^r_1, \ldots, p^r_{24})$ with $p^r_t = p^r_f$ for all $t \in P_f$.

Evaluate expected Spearman rank correlation $\mathbb{E}_d[\rho_S(\mathbf{p}^r(\mathcal{P}), \mathbf{p}^w_d)]$.

end for

Select the optimal partition

$$\mathcal{P}^*(F) = \arg\max_{\mathcal{P}} \mathbb{E}_d[\rho_S(\mathbf{p}^r(\mathcal{P}), \mathbf{p}^w_d)]\,.$$

end for

D Appendix: Strategic parameter calibration

For each BRP i and net load bin b, the strategic conduct parameter $\hat{\theta}_{i,b}$ is estimated through the following five-step algorithm.

Step 1: Identify hours in bin b. I begin by selecting all hours t in 2023 where the hourly net load falls within bin b: $\mathcal{T}_b = \{t : Q_t \in \mathcal{B}_b\}$. Each bin groups hours with comparable system conditions. With 20 bins covering the full range of observed net load, the average bin contains roughly $|\mathcal{T}_b| \approx 365 \times 24/20 \approx 438$ hours.

Step 2: Construct unit-level cost and capacity data. For every generating unit k and hour t, the marginal cost is calculated as:

$$MC_{k,t} = \frac{\text{FuelPrice}_{f(k),t}}{\text{Efficiency}_k} + \text{CarbonPrice}_t \times \text{EmissionsRate}_k,$$

where $f(k)$ denotes the fuel type (e.g., natural gas, coal, oil). Fuel prices vary monthly, carbon prices daily, and unit efficiencies are fixed technical parameters. The installed capacity $\text{Cap}_{k,t}$ is assumed to be available for thermal units (adjusted for planned outages where data permits).

Step 3: Construct residual demand curves. For each hour $t \in \mathcal{T}_b$ and focal BRP i, I compute the residual demand faced by that BRP.

(a) Competitor supply. The aggregate supply of all competitors (all BRPs except i) at price p is:

$$S_{-i}(p,t) = \sum_{j \neq i} \sum_{k \in j} \text{Cap}_{k,t} \times \mathbb{1}\{MC_{k,t} \leq p\}.$$

(b) Residual demand. The residual demand for firm i is: $RD_i(p,t) = Q_t - S_{-i}(p,t)$, that is, the quantity firm i must supply at price p for the market to clear.

(c) Local slope estimation at the observed price.

For each hour in 2023 and for each BRP, I evaluate the slope of the *inverse* residual demand curve around the observed wholesale clearing price P_t. Specifically, I create a grid of hypothetical prices within a €10/MWh window around the actual market price:

$$\mathcal{P}_t = \{P_t - 10, P_t - 9, \ldots, P_t, \ldots, P_t + 10\}$$

For each hypothetical price $p \in \mathcal{P}_t$, I compute how much competitors would supply by summing the available capacity of all competing units with marginal costs at or below that

price level:

$$S_{-i}(p,t) = \sum_{u \in U_{-i}} \text{Cap}_u \cdot \mathbb{1}\{\text{MC}_u \leq p\}$$

where U_{-i} denotes all generation units not owned by firm i.

The residual demand facing firm i at each hypothetical price is then:

$$RD_i(p,t) = Q_t - S_{-i}(p,t)$$

where Q_t is the total net load that must be served. This residual demand represents what remains for firm i to supply after competitors have offered their economically rational capacity at price p.

To estimate the local slope at the observed clearing price, I identify the price points immediately below and above P_t:

$$i_1 = \max\{j : p_j \leq P_t\}$$
$$i_2 = \min\{j : p_j > P_t\}$$

The slope of the inverse residual demand curve is computed as:

$$\theta_{it} = \frac{\Delta p}{\Delta RD_i} = \frac{p_{i_2} - p_{i_1}}{RD_i(p_{i_2}, t) - RD_i(p_{i_1}, t)}$$

This slope is typically *negative*, reflecting the downward-sloping nature of residual demand: as price increases, competitors supply more, leaving less residual demand for firm i. Since $\Delta p > 0$ (price increases) and $\Delta RD_i < 0$ (residual demand decreases), the ratio $\theta_{it} < 0$.

The economic interpretation of $|\theta_{it}|$ is the price responsiveness of firm i's residual demand: it quantifies how much the clearing price must rise for firm i to be able to supply one additional MW (or equivalently, how much price would fall if firm i withheld 1 MW). A large $|\theta_{it}|$ indicates steep residual demand—competitors have limited spare capacity, so small quantity changes by firm i lead to large price movements, signaling substantial market power.

Step 4: Aggregate local slopes within bin b. From Step 3, I collect the hour-specific slopes $\{\theta_{it}\}_{t \in \mathcal{T}_b}$ for all hours in bin b and compute the bin average:

$$\overline{\theta}_{i,b} = \frac{1}{|\mathcal{T}_b|} \sum_{t \in \mathcal{T}_b} \theta_{it}$$

This captures the typical inverse residual demand slope faced by BRP i under net load condi-

tions in bin b.

Step 5: Map the average slope to the conduct parameter. The strategic markup parameter used in the bid function is obtained by negating the average slope (since $\overline{\theta}_{i,b} < 0$):

$$\hat{\theta}_{i,b} = -\overline{\theta}_{i,b} > 0$$

This ensures $\hat{\theta}_{i,b}$ is positive and can be directly used in the bid markup formula:

$$B_i(q) = MC_i(q) - \hat{\theta}_{i,b} \cdot \text{CumCap}_i$$

where the negative sign in front of $\hat{\theta}_{i,b}$ combined with $\hat{\theta}_{i,b} > 0$ produces a positive markup above marginal cost.

A more negative (steeper) $\overline{\theta}_{i,b}$ implies larger $|\overline{\theta}_{i,b}|$, yielding higher $\hat{\theta}_{i,b}$ and stronger strategic markups. Conversely, flatter residual demand (smaller $|\overline{\theta}_{i,b}|$) implies weaker market power and smaller markups.

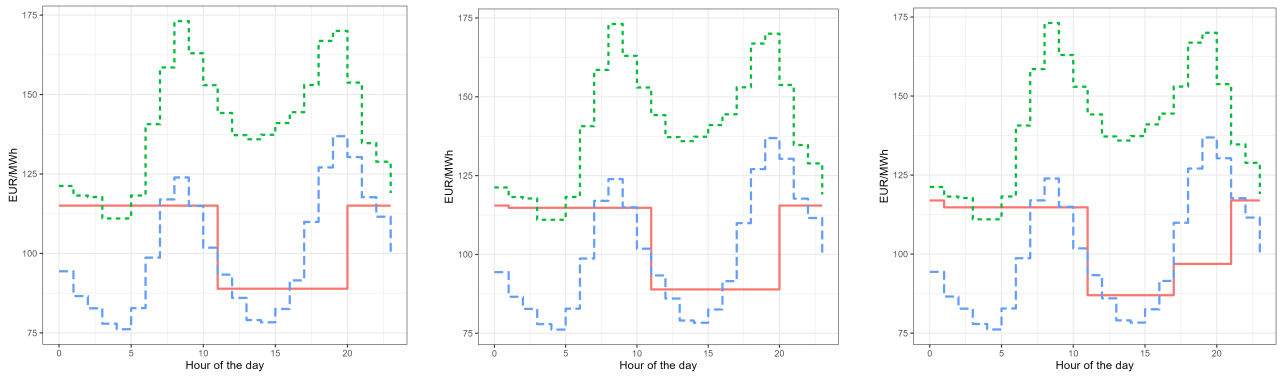
Table 9: Calibrated strategic conduct parameters

Statistic	Value
Total firm-bin combinations	340
Non-zero parameters	187 (55%)
Median $\theta_{j,b}$ (non-zero)	0.00363 €/MWh
Mean (non-zero)	0.00650 €/MWh
Standard deviation (non-zero)	0.01626
By net load range:	
Low bins (1-7)	0.000
Mid bins (8-14)	0.00668
High bins (15-20)	0.00118

Note: Estimated from 2023 Belgian day-ahead data. Full methodology in Appendix D.

E Appendix: Results simulation for heat-pumps load shifting 2023

Time-of-use optimal designs at equilibrium



(a) Two-block ToU (b) Three-block ToU (c) Four-block ToU

Figure 9: Optimized time-of-use pricing structures after convergence

Notes: Red solid line: optimized time-of-use designs and prices after three iterations of 20% of heat pumps cycle shifting calibrated on 2023 Belgian data. Green dotted line: strategic-bidding wholesale price for March 8, 2023 (illustrative day). Blue dotted line: baseline wholesale hourly average over 2023.*Source:* ENTSO-E Transparency Platform; authors' calculations.

Table 10: Equilibrium system costs by pricing regime from simulation calibrated on Belgian electricity markets in 2023 (heat pump load shifting

Pricing regime	**Supply-side efficiency**		
	Average cost (€/MWh)	System cost (M€)	Δ vs. baseline %
Baseline (flat)	143	6,247.6	—
Real-time pricing (RTP)	122.64	5,222.7	−16.4
ToU2	122.70	5,223.6	−16.4
ToU3	122.70	5,223.6	−16.4
ToU4	122.69	5,222.8	−16.4
Efficiency comparisons:			
ToU2 or ToU3 − RTP	≈ 0	+0.9	*O*pp
ToU4 − RTP	≈ 0	+1.1	*O*pp

Notes: Baseline (flat) has no iteration. **Real-time pricing (RTP)** sets retail prices in each iteration equal to the wholesale prices from the previous iteration; only Stages 1–2 iterate to convergence (see Fig. 3). **Time-of-use (ToUF)** for $F = 2, 3, 4$ optimizes a F-block schedule each iteration to best match the *relative* hourly wholesale price profile from the previous iteration. *System cost* is total costs of serving the demand at market clearing price; *Average cost* is system cost divided by total energy served. Results shown at convergence after three iterations with 20% participation of estimated heat pumps daily cycles. Percentage changes are relative to the baseline; comparison rows report percentage-point differences. ToU2 and ToU3 yield identical efficiency outcomes. Source: refer to Section 3.

Table 11: Equilibrium consumer expenditure and surplus by pricing regime from simulation calibrated on Belgian electricity markets in 2023 (heat pump load shifting)

Pricing regime	**Consumer savings and welfare**		
	Expend. (M€)	Surplus (M€)	Δ vs. baseline %
Baseline (flat)	10,225.6	—	—
Real-time pricing (RTP)	8,245.2	1,977.9	—
ToU2	8,247.1	1,976.9	−0.07
ToU3	8,247.1	1,976.9	−0.07
ToU4	8,246.1	1,977.3	−0.03
Efficiency comparisons:			
ToU2 or ToU3 − RTP	+1.9	−1.	−0.07 pp
ToU4 − RTP	+0.9	−0.6	−0.03 pp

Notes: Baseline (flat) has no iteration. **Real-time pricing (RTP)** sets retail prices in each iteration equal to the wholesale prices from the previous iteration; only Stages 1–2 iterate to convergence (see Fig. 3). **Time-of-use (ToUF)** for $F = 2, 3, 4$ optimizes a F-block schedule each iteration to best match the *relative* hourly wholesale price profile from the previous iteration. Expenditure is total retail payments across all consumers. Surplus measures welfare gains relative to baseline, computed assuming linear demand assumption. Results shown at convergence after three iterations with 20% participation of estimated heat pumps daily cycles. Percentage changes are relative to the baseline; comparison rows report percentage-point differences. ToU2 and ToU3 yield identical efficiency outcomes. Source: refer to Section 3.

Welfare and system costs

Load profiles and peak reduction

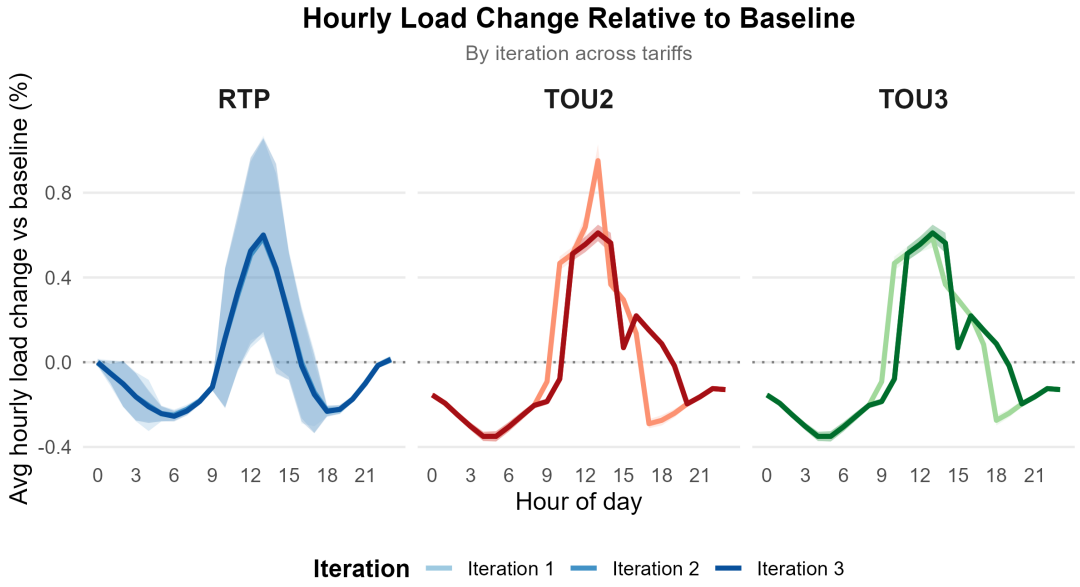
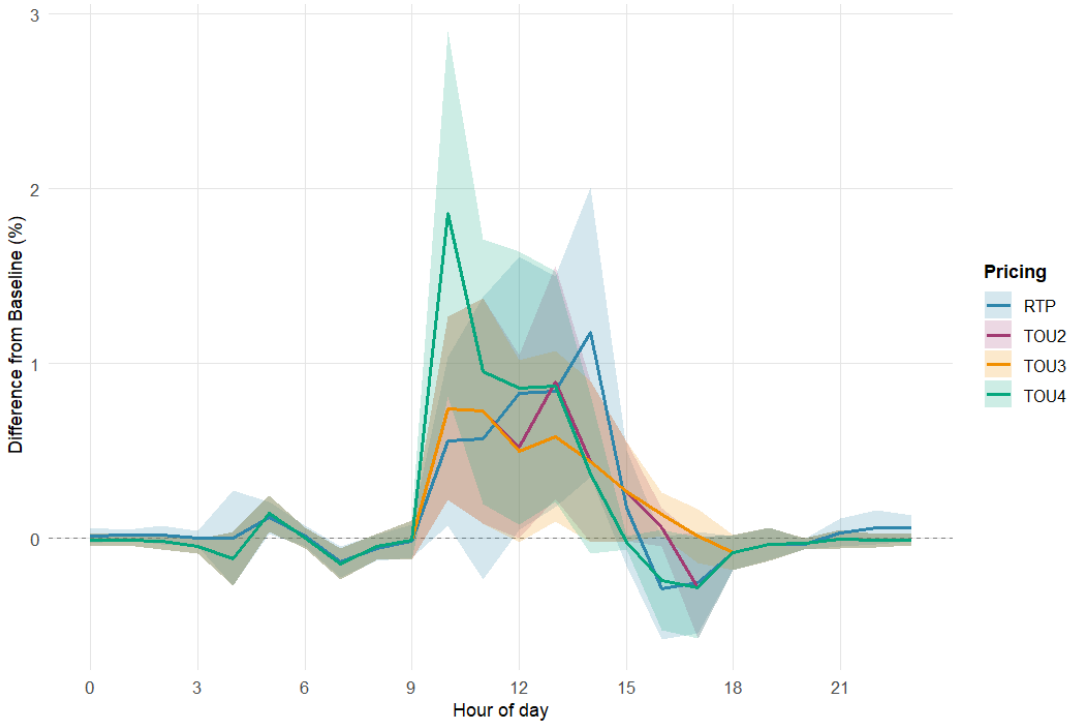


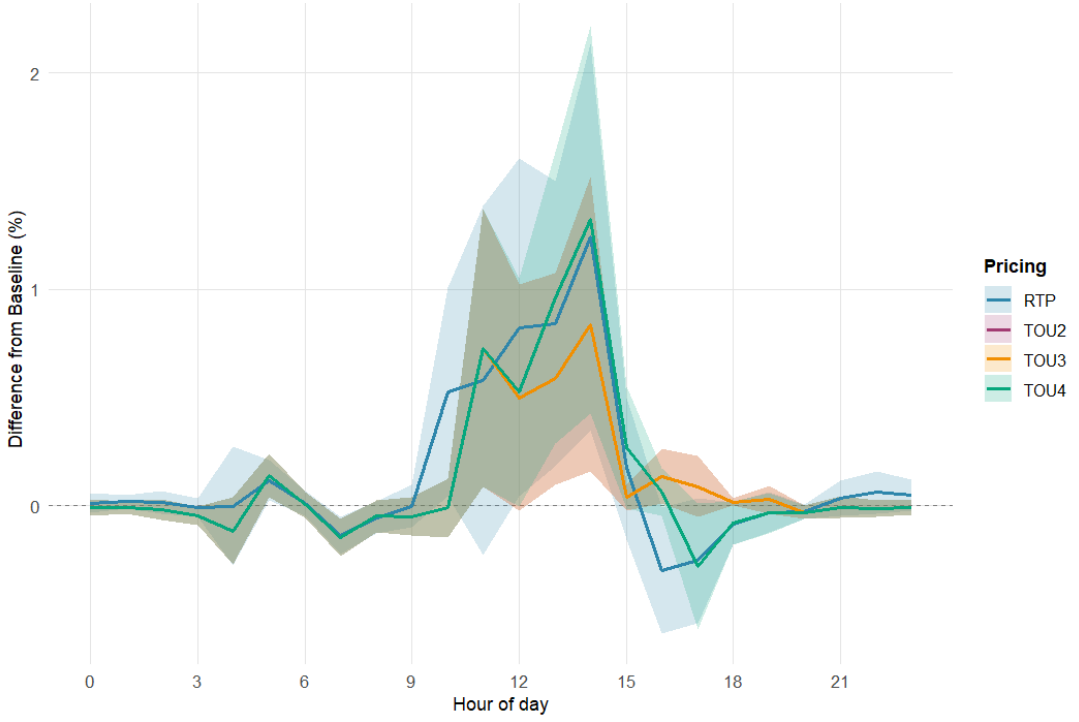
Figure 10: Percentage change in hourly load relative to baseline (simulation Belgian electricity markets in 2023)

Note: Average hourly load changes relative to a flat-pricing baseline are shown for each iteration under real-time pricing (RTP), two-period ToU (ToU2), and three-period ToU (ToU3). Iteration 1 is the initial consumer response; later iterations update wholesale prices (RTP) or the ToU partition (ToU2/ToU3) until convergence. Positive values mean load shifts toward that hour; negative values mean reductions. Results average all days with 20% of eligible flexible loads (HP) participating; shaded bands depict iteration-to-iteration adjustments.

Wholesale prices impact



(a) Percentage change in hourly wholesale prices relative to baseline - iteration 1



(b) Percentage change in hourly wholesale prices relative to baseline - iteration 3

Figure 11: Wholesale market price impacts of dynamic retail pricing for heat pumps load shifting

Notes: Percentage changes in wholesale clearing prices under real-time pricing scenarios relative to a counterfactual baseline, averaged across 2023 weekdays. Panel (a): iteration 1; panel (b): iteration 3 (convergence). All wholesale prices are endogenously determined using the market clearing model in Section 2.2. The baseline represents a status quo flat retail pricing and no demand response, solved using the same wholesale market model to ensure comparability. Shaded regions show 95% confidence intervals across days. This is a controlled counterfactual comparison between model equilibria, not a comparison to observed historical prices.