Market Failures in Local Flexibility Market Proposals for Distribution Network Congestion Management

Hennig, R.J.; Tindemans, Simon H.; De Vries, Laurens

DOI
10.1109/EEM54602.2022.9920980

Publication date
2022

Document Version
Final published version

Published in
18th International Conference on the European Energy Market, EEM 2022

2022 18th International Conference on the European Energy Market (EEM) | 978-1-6654-0896-7/22/$31.00 ©2022 IEEE | DOI: 10.1109/EEM54602.2022.9920980

Market Failures in Local Flexibility Market Proposals for Distribution Network Congestion Management

Roman Hennig [a], Simon H. Tindemans [b], Laurens de Vries [a]

[a] Faculty of Technology, Policy and Management, Delft University of Technology, Jaffalaan 5, 2628 BX Delft, The Netherlands

[b] Department of Electrical Sustainable Energy, Delft University of Technology, The Netherlands

***Abstract*—The steady uptake of PV cells and high-power flexible loads such as electric vehicles (EVs) and heat pumps can lead to localized network congestion, if their power consumption or feed-in is not controlled well. One potential way that has been proposed to manage this congestion, are so-called Local Flexibility Markets. It is often argued, that these proposals are theoretically efficient, as they are market-based. However, some of these proposals may suffer from design flaws that allow market participants to obtain undue profits at the expense of the network operator. In this contribution, we discuss which kinds of market failures can occur based on theoretical reasoning and demonstrate them in a toy model. Based on this, we argue for a more careful consideration of congestion management options.**

***Index Terms*—distribution networks, congestion management, flexibility markets, electric vehicles, network operators**

I. Introduction

Congestion is becoming a problem in electric distribution networks. It is often due to high power flexible loads like electric vehicles and heat pumps, or due to excessive feed in of distributed generation. In order to deal with this problem, several families of solutions have been proposed: new forms of distribution tariffs [1], [2], direct load control schemes [3], [4] and different forms of Local Flexibility Markets (LFMs) [5]–[8] .

In this analysis, we focus on the LFM proposals for congestion management. Some of these market-based proposals are structured in a way that may lead to perverse incentives for profit-maximizing market parties. Rather than helping to relieve congestion, they may lead to more congestion and market participants may collect undue profits for their participation in the market. This is worrying, as LFMs are currently widely discussed as a potential remedy for distribution-level congestion and a careless application of the concept may lead to high inefficiencies and excessive costs for network operators.

The article is structured as follows: in section 2 we describe the versions of LFM proposals that we analyse here and give examples from the literature. In section 3 we give the theoretical argumentation for why these proposals can lead to market failures. In section 4 we demonstrate these failures in an illustrative toy model of an LV network. In section 5 we discuss the consequences of these findings and implications for policy making and academic work. Section 6 concludes.

II. LFM Proposals in the Literature

Many LFM proposals are based on the idea of *baseline schedules* [9]–[11]. In this implementation, aggregators of flexible loads submit schedules for the loads that they control, typically on a day-ahead time frame. The Distribution System Operator (DSO) collects these schedules and also forecasts the anticipated inflexible load at the congestion point. If the sum of the schedules and inflexible loads leads to congestion problems at any time step, it requests flexibility offers from the participating aggregators. The aggregators then submit bids for reducing load at the selected time steps. As [10] point out, it is important to also consider the time step at which this reduced load is then added again (called the “payback time” in [10]). Thus, a bid can include:

- The time step for reducing load
- The time step at which this load is added instead
- The maximal load reduction in *kW*
- A price, e.g. per kWh of shifted load.

Congestion can occur at many different points in the grid: e.g., at LV transformers, cables or HV/MV substations. Unfortunately, many proposals do not clearly state for which congestion point they are intended to be applied which makes it hard to assess how they would work in practice. The USEF framework [9] specifies: “A congestion point is a set of connections which (directly) relate to a part of the grid where grid capacity might be exceeded because it may be insufficient to distribute the requested amount of energy; e.g. the secondary side of an LV transformer.” Thus, it explicitly includes LV transformers, which we will use as a case study in the toy model in section IV.

An important aspect of the LFM is the clearing process by which bids are selected. [9]–[11] all propose to use pay-as-bid pricing. Other methods, such as pay-as-cleared [5] or Vickrey-Clarke-Groves auction types [12] have also been suggested.

In the discussion above and in the case study, we focus on proposals with products that are based on baseline schedule. Other “product types” that could be sold in LFMs have also been suggested:

This work is part of the research program STEP-UP with project number 438-18-404, which is partly financed by the Dutch Research Council (NWO).

978-1-6654-0896-7/22/$31.00 ©2022 IEEE

- An option-type trade where load reduction is only activated by the DSO if necessary [10], [11]
- An emergency load reduction that is activated in real time relative to the current consumption of the market participant, rather than relative to a baseline [11]
- A power limitation product, where load is limited to a maximum power rather than reduced relative to a baseline or current consumption [11], [12].

III. Pathways of Possible Market Failures in LFMs for Local Congestion MAngement

In theory, markets could lead to perfectly efficient solutions: this holds if all market participants are completely transparent about their costs and baseline schedules and submit flexibility bids in accordance with their true marginal cost of shifting loads and if the market operation has no unintended consequences. However, firms in competitive markets behave in profit-maximizing ways. This can become a problem in poorly designed markets. In the proposals described in the previous section, there are several potential avenues by which profit maximization would not result in managing congestion in the most cost-efficient way for the network operator. In this section, we describe the reasons behind two types of such market failures in managing congestion with LFMs. They are based on the ability of aggregators to control either the price or the volume of traded flexibility.

A. Price Control: Market Power

Distribution level congestion can be highly localized, e.g. at the single LV feeder level, as suggested by [9]. Thus, it may be possible for an aggregator of flexible loads to acquire market power in a local market: if it has control over a sufficient amount of flexible loads it can effectively hold a monopoly over the congestion on that feeder, in the sense that congestion cannot be removed without participation of this aggregator. Therefore, this aggregator can influence the clearing price of the market at this feeder, a concept known as strategic bidding. It can charge prices for adjusting its consumption that are significantly higher than its marginal cost of shifting loads. This problem is analogous to the pivotal supplier problem in wholesale electricity markets [13]. A similar case has been assessed in [14] for wind power operators in coupled local and central (wholesale) markets. Strategic bidding may also occur when it is not only a single aggregator holding market power, but rather an oligopoly of a few aggregators. They may collectively bid at higher-than-marginal prices. This can happen either due to explicit collusion between them, or due to implicitly learning that they can obtain higher profits in this way on average.

Market power can be diluted in highly liquid markets with many bidders. In the case of distribution level congestion management, this would be the more likely case at higher network levels in the distribution grid, e.g. at a substation which serves in the order of 10.000's of customers. This suggests that LFMs for local congestion management may be more suited to application at the substation level, rather than at LV feeder level.

B. Volume Control: Modified Schedules

The presence of an additional revenue stream due to an LFM for local congestion management may influence the schedules that aggregators submit in proposals such as [9]. The schedules may be modified relative to the case without the LFM because of expected payments in case of congestion. We can distinguish two subcases of schedule modification:

1) "Fake" schedules: These are schedules that the aggregator does not actually intend to fulfil. It submits them in the expectation that it will not have to do so because it will be paid under the LFM mechanism to reduce load anyways. This behaviour is more likely if the aggregator expects congestion with near certainty, or when there are low penalties for not sticking to the submitted baseline if no congestion occurs. In those circumstances, submitting fake schedules could become a low-risk winning strategy for a profit maximizing firm.

One could assume that the network operator should be able to detect falsified baseline schedule, but this is difficult because the network operator cannot know the true constraints of the aggregator. The aggregator could argue that the flexible loads, e.g. EVs charged at a home charger, are really only available during the submitted times. This might lead to a situation where there is a large burden on DSOs to proof that aggregators use manipulated schedules and aggregators try to find ever more elaborate methods of manipulation for which they can argue that these are their true baselines due to availability constraints.

2) "True" profit-maximizing modifications of schedules: In this case, the aggregator submits schedules that it actually intends to adhere to, even when congestion does not materialize and it is not paid for shifting loads. The presence of payments to shift loads during those time when flexibility is requested offsets the additional costs of modifying the schedule. Note that this is also a possible avenue of profit maximization in proposals that operate in near-real time, rather than relying on pre-submitted baseline schedules (e.g. in [11], [15]). In this case, aggregators could increase their consumption when they anticipate high-network load in real-time, in order to be paid to reduce it again.

Note also that this doesn't have to entail malign intent of employees of the aggregator themselves; it might simply be the outcome of an optimization algorithm that has been programmed to take all available revenue streams into account. Lastly, note that modified baseline schedules can have negative repercussions beyond network operation itself: they often imply that the aggregator has to charge more vehicles during high consumption hours, where market prices may be higher and added electricity demand is likely served by conventional electricity generators based on coal and gas.

Both cases of modified baseline schedules can aggravate the congestion problem relative to the case without LFM. They often occur in conjunction with market power: a market participant may modify its schedule in order to obtain a position where congestion cannot be removed without them and then charge inflated prices in an LFM. But even without market power, a market participant can add to an expected pre-existing congestion problem in order to get paid to subsequently reduce its consumption. While pure market power alone could be addressed by diversifying markets, modified baseline schedules cannot so easily be avoided.

The fundamental market failure that enables both types of scheduling problems is, that in many LFM proposals there is no cost associated with submitting a baseline schedule that targets times of likely congestion. Contrast this to energy markets: there, forward and day-ahead markets for energy at peak demand hours have already ‘priced in’ the high expected cost of marginal generation. It is therefore costly to take a position that allows one to profit from high intra-day prices. Although in some cases, the long-term costs of the energy market can reduce the benefits of submitting modified schedules in LFMs, the objectives of both markets do not always align. For example, an EV charging aggregator can optimize their *entire* fleet for the energy market, and can submit modified schedules for EVs in a congested area of the network, balancing those changes with modified schedules for EVs in other areas of the network.

IV. Demonstrations of Market Failures in a Toy Model

In this section we demonstrate the kinds of market failures described in the previous section in a toy model. The model consists of a simplified neighbourhood with a typical LV feeder that supplies 50 households and 24 charging EVs. The 24 EVs are controlled by 3 different aggregator companies, called A, B and C. We model 6 time steps of length 1 hour each, which exemplify a typical night from the evening peak through a drop of traditional loads during the night and then a rise in load in the morning. Table 1 gives the chosen values for inflexible loads and wholesale market prices (not to be confused with the LFM clearing prices). These are motivated by real world data, such as those used in our modelling in [16]. Each EV requires 12 *kWh* of energy overnight and has a maximal charging power of 11 *kW*. The LFM is modelled based on the proposals described in section II: aggregators submit a baseline schedule to the DSO and submit bids for reducing consumption. Each bid contains the possible amount of reduction, the payback time period and the price per *kWh*. The market is cleared pay-as-bid based on the lowest possible cost of resolving congestion.

TABLE I. Load and Price Time Series

	Time Step [hour]					
	0	***1***	***2***	***3***	***4***	***5***
Inflexible Load *kW*	70	50	35	25	30	40
Market Price *Euro*	0.15	0.1	0.05	0.01	0.025	0.032

We begin with a situation in which the 24 EVs are evenly split over the 3 aggregators and there is perfect competition. We assume that aggregators spread out charging schedules for EVs over all time steps, due to availability requirements. However, the lowest price time steps are preferred, leading to some expected congestion due to the scheduled EV charging at hour 3. This initial situation is shown in the top panel of Fig. 1.

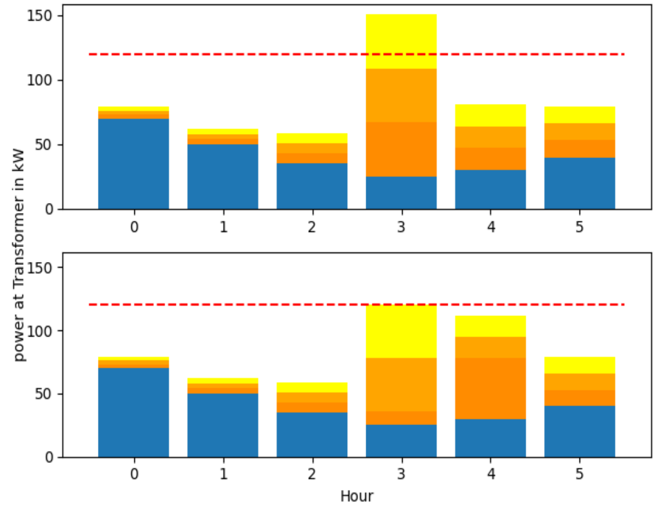

Figure 1: Initial and final schedules in a perfect competition market with efficient outcomes. Blue: inflexible load. Dark orange, orange, yellow: Aggregators A, B and C. Red dashed line: rated transformer capacity.

A. Scenario 1: Perfect competition, efficient outcomes

In the first scenario, we assume all aggregators submit their true original schedules (the same schedules as in the absence of the LFM) and submit truthful bids at their marginal cost of shifting from the congested time step to other time steps. The DSO selects these bids to remove congestion at the lowest possible cost, which lead to a cost-efficient outcome for the DSO. For aggregators the schedule change is revenue-neutral, as they are reimbursed at their marginal cost of shifting. See Fig. 1 bottom for the final schedules.

B. Scenario 2: Aggregator A is dominant and charges inflated prices

In the second scenario, we assume that aggregator A controls 20 of the 24 EVs. This means that congestion at hour 3 cannot be removed without participation of this aggregator and it can therefore charge higher prices. We assume that it charges the marginal prices of shifting to another time step (based on wholesale price differences), plus an additional mark-up of 1 Euro/*kWh*. Now the DSO first clears the cheaper bids of aggregators B and C and then also has to accept 11 *kWh* from aggregator A in order to fully remove congestion. See Fig. 3 for the initial and final situations in this scenario.

Note that the choice of the mark-up price that aggregator A charges is somewhat arbitrary here. Theoretically the upper limit for this mark-up is given by the DSO’s cost of the alternative of removing congestion with an LFM. This could be the Value of Lost Load (VoLL) in the short term or the cost of upgrading the transformer in the long-term. However, this cost would be many orders of magnitude higher than the marginal cost of shifting for the aggregator and in practice it would be quite easy to prove market power abuse at these values. Therefore, an aggregator would like choose a lower value which could be considered realistic and for which it would be hard to prove an abuse of market power.

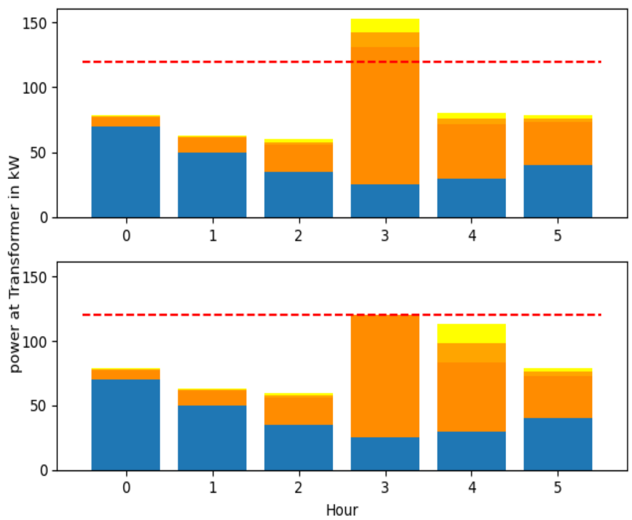

Figure 2: Network load schedules before and after LFM operation in case of market power and inflated prices by Aggregator A

C. Scenario 3: Aggregator A submits a fake schedule

In this scenario, each aggregator controls 8 EVs again. Aggregator A anticipated that the inflexible load in time step 1 will be high and that it can cause a congestion problem by scheduling all of its available loads in this time step. In reality, it may not even have all EVs available for charging at this time. But since it expects to be paid for shifting loads, it does not intend to actually fulfil this schedule anyways. We again assume it charges an additional mark-up price of 1 Euro/*kWh* for shifting. The DSO has to accept this bid to avoid overload at hour 0. The before- and after LFM operation schedules are depicted in Fig. 3. Note that due to aggregator A moving all of its load to hour 0, there is no longer any congestion at hour 3.

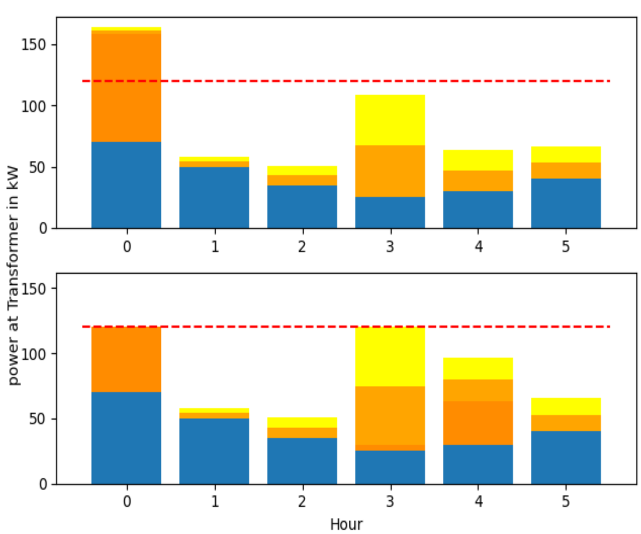

Figure 3: Network load schedules before and after LFM operation in case of fake schedule by Aggregator A

D. Scenario 4: Aggregator A submits a true modified profit-maximizing schedule

Similar to the previous scenario, Aggregator A anticipates that inflexible load is typically high in time step 1. However, rather than submitting a fictional schedule that it may not be able to fulfil, it submits a realistic schedule that can be fulfilled even if it is not called to shift loads.

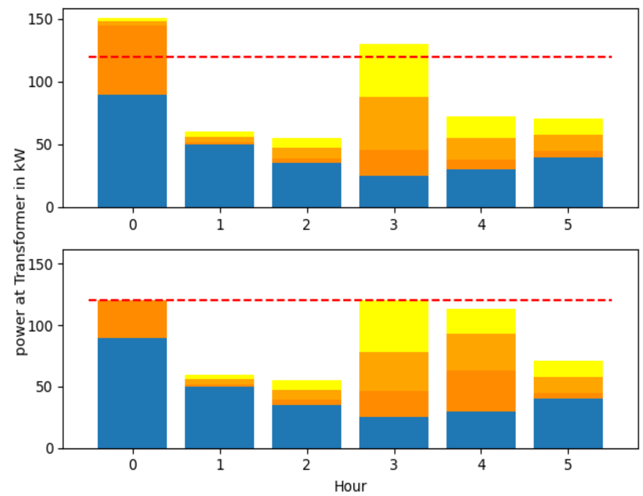

Figure 4: Modified schedule by Aggregator A, case of high inflexible load at hour 0: inflexible load is 90 (compared to 70 in other scenarios)

In some cases it will be called and paid according to its bids and in some cases not. As long as the expected value for the profit of this strategy is higher than submitting the truthful (without LFM payments) schedule, it has an incentive to follow this strategy. We assume that Aggregator A is certain that it has 5 out of its 8 vehicles available for charging at hour 0. Further it assumes that in 50 % of all days, the load at the LV feeder in hour 0 will be 90 *kW*, so that congestion occurs when all 5 of the available EVs are charged at full power and for simplicity we assume no congestion occurs in the remaining 50% of days. If these assumptions are correct and it charges a price higher than a certain threshold value, this is a profit maximizing strategy. It can be calculated that with the given values, the threshold value for the price it must charge is around 0.476 Euro (see Appendix A). We again assume that it charges an additional 1 *Euro/kWh* mark-up to the wholesale price difference for shifting. The initial and final schedules are shown in Fig. 4 for the situation with high load.

Table 2 gives an overview of the costs of removing congestion for the DSO in the different scenarios. In a perfect market with truthful schedules and bidding, the DSO can remove congestion for the low price of 0.47 Euro. In the scenarios with market failures on the other hand, the DSO ends up paying significantly higher prices. The aggregator who exerts market power and submits non-truthful schedules benefits from these payments. In scenario 2, where schedules are truthful but bidding is not, all of the extra costs of the DSO accrue as profits for the aggregator. In scenarios 3 and 4 the schedules themselves have been modified by the presence of LFM payments. In our toy model, we observe that the modified schedule actually helps in the cases where there is no congestion at hour 0, because aggregator A removed some load from the other congested hour (3) so that congestion is not as severe there anymore. However, as shown in Table 2, these savings are offset by the times when there is congestion at hour 0.

TABLE II. DSO COSTS OF LFM OPERATION IN EURO

	Scenario			
	Perfect Comp.	*Market Power*	*"Fake" Schedule*	*mod. Schedule*
DSO Costs [Euro]	0.47	11.47	32.34	10.64 (exp. value)

V. DISCUSSION AND POLICY RECOMMENDATIONS

As we have seen, LFM proposals that are based on baselines can lead to inefficient outcomes, where aggregators collect undue profits at the expense of the DSO. Since the DSO will typically pass on these costs through network tariffs, this also comes at the expense of other network customers. This creates strong fairness concerns: aggregators and their customers may benefit by behaving in network-burdening ways rather than network-serving ways. This is worrying from an income inequality perspective as well: owners of high-power flexible loads are typically more wealthy [17][18]. Some of the problems that we demonstrated can be partly remedied by regulation: fake schedules can be discouraged by imposing strong penalties and disqualification from further trades when they are detected. Market power can be diluted by imposing market diversification requirements (which is easier at a larger aggregation level). However, the exercise of market power may be difficult to detect. True profit maximizing schedule modifications may also be difficult to detect and are not technically illegal.

Therefore, we recommend to counteract these problems in one of the following ways:

1. Use of other congestion management methods that are more targeted to the problem.
2. In case the LFM method is still the preferred method by the DSO and other stakeholders, there should be a careful consideration of the possible market failures before implementing an LFM at large scale. LFM products based on baselines and pay-as-bid clearing should be avoided and other product types, such as those mentioned in section 2, should be considered. Strong regulatory oversight is necessary.

Academic and regulatory proposals that advocate for LFM type solutions should give a clear and precise problem definition for the kind of congestion that the solution is intended to resolve, and demonstrate how it does so without creating undue profit opportunities for aggregators. Other possible CM solutions should be taken into account as well. Network operators, regulators and other involved stakeholders should be aware of the possible unintended consequences of LFM type proposals and be informed about other possible solutions.

VI. CONCLUSION

We have reviewed a class of distribution congestion management proposals with Local Flexibility Markets and demonstrated how they may lead to inefficient solutions. These inefficiencies can occur because the private incentives of aggregators for profit maximizations are not aligned with the DSO objective of removing congestion at lowest societal cost. We identified market power and modified baseline schedules as two of the main market failures for LFMs that are based on baselines. The impact of market power can be reduced by increasing the number of market participants. Modified baseline schedules can pose a problem even in situations with many participants. We therefore recommend stakeholders to carefully consider possible market failures before attempting to implement LFM type solutions and to also consider using alternative congestion management mechanisms.

REFERENCES

[1] S. Bjarghov, H. Farahmand, and G. Doorman, "Capacity subscription grid tariff efficiency and the impact of uncertainty on the subscribed level," *Energy Policy*, vol. 165, p. 112972, Jun. 2022.

[2] N. Govaerts, K. Bruninx, H. Le Cadre, L. Meeus, and E. Delarue, "Forward-looking distribution network charges considering lumpy investments," *J. Regul. Econ.*, vol. 59, no. 3, pp. 280–302, Jun. 2021.

[3] Clean Energy Group, "ConnectedSolutions An Assessment for Massachusettes."

[4] Bundesnetzagentur, "Flexibility in the electricity system," 2017.

[5] X. Jin, Q. Wu, and H. Jia, "Local flexibility markets: Literature review on concepts, models and clearing methods," *Applied Energy*, vol. 261. Elsevier Ltd, p. 114387, 01-Mar-2020.

[6] S. Minniti, N. Haque, P. Nguyen, and G. Pemen, "Local Markets for Flexibility Trading: Key Stages and Enablers," *Energies 2018, Vol. 11, Page 3074*, vol. 11, no. 11, p. 3074, Nov. 2018.

[7] J. Radecke, J. Hefele, and L. Hirth, "Markets for Local Flexibility in Distribution Networks," Kiel, Hamburg, 2019.

[8] H. de Heer, M. van der Laan, and A. S. Armenteros, "USEF: THE FRAMEWORK EXPLAINED," 2021.

[9] USEF, "USEF Flexibility Trading Protocol Specifications."

[10] A. Esmat, J. Usaola, and M. Á. Moreno, "A Decentralized Local Flexibility Market Considering the Uncertainty of Demand," *Energies 2018, Vol. 11, Page 2078*, vol. 11, no. 8, p. 2078, Aug. 2018.

[11] Y. Ding *et al.*, "Development of a dso-market on flexibility services," 2013.

[12] C. Heinrich, C. Ziras, T. V. Jensen, H. W. Bindner,

and J. Kazempour, “A local flexibility market mechanism with capacity limitation services,” *Energy Policy*, vol. 156, p. 112335, Sep. 2021.

[13] M. Bataille, A. Steinmetz, and S. Thorwarth, “Screening Instruments for Monitoring Market Power in Wholesale Electricity Markets-Lessons from Applications in Germany,” Mannheim, ZEW Discussion Paper 14-048, 2014.

[14] M. Farrokhseresht, H. Slootweg, and M. Gibescu, “Strategic bidding of distributed energy resources in coupled local and central markets,” *Sustain. Energy, Grids Networks*, vol. 24, p. 100390, Dec. 2020.

[15] A. Esmat, J. Usaola, and M. Á. Moreno, “Distribution-Level Flexibility Market for Congestion Management,” *Energies 2018, Vol. 11, Page 1056*, vol. 11, no. 5, p. 1056, Apr. 2018.

[16] R. Hennig, M. Jonker, S. Tindemans, and L. De Vries, “Capacity Subscription Tariffs for Electricity Distribution Networks: Design Choices and Congestion Management,” in *International Conference on the European Energy Market, EEM*, 2020, vol. 2020-Septe.

[17] D. Ribó-Pérez, M. Heleno, and C. Álvarez-Bel, “The flexibility gap: Socioeconomic and geographical factors driving residential flexibility,” *Energy Policy*, vol. 153, p. 112282, Jun. 2021.

[18] M. Ansarin, Y. Ghiassi-farrokhfal, W. Ketter, and J. Collins, “The economic consequences of electricity tari ff design in a renewable energy,” *Appl. Energy*, vol. 275, no. February, p. 115317, 2020.

Appendix

In this appendix we derive the threshold value for the price of shifting load in scenario 4 in section 4. Beyond this threshold value, submitting a modified charging schedule, which worsens the original congestion problem, becomes a winning strategy for the aggregator.

In the following we assume that the aggregator originally would schedule load at time step t_2, which has a lower day-ahead wholesale market price π^{DA} than time step t_1. There is a Local Flexibility Market for congestion management, which pays a price $\pi^{LFM}(t_1, t_2)$ for shifting loads from t_1 to t_2. Congestion at t_1 happens with likelihood $p(cong)$. This may induce the aggregator to schedule some additional load $\tilde{q}$ at t_1. In the case of congestion, some of this load will be shifted due to the LFM, and the rest will not: $\tilde{q} = q^{shift} + q^{not-shift}$. With these quantities, we can write the expected profit of the modified schedule, S^{mod}, relative to the original schedule as:

$$\begin{aligned}\mathrm{EV}(S^{mod}) = p(cong) &* \left(q^{shift} * \pi^{LFM}(t_1, t_2) - q^{not-shift} * (\pi^{DA}(t_1) - \pi^{DA}(t_2)\right) \\ &- \overline{p(cong)} * \tilde{q} * \left(\pi^{DA}(t_1) - \pi^{DA}(t_2)\right)\end{aligned}$$

Where the first line gives the profit from congestion events: the payments from the LFM for the shifted load minus the cost of higher wholesale prices for the not-shifted load, and the second line gives the losses in case no congestion occurs: higher costs from the wholesale market for the load that has been added at t_1. The profitability condition $\mathrm{EV}(S^{mod}) > 0$ yields:

$$\pi^{LFM}(t_1, t_2) > \frac{1}{q^{shift}} * \left(\frac{\overline{p(cong)}}{p(cong)} * \tilde{q} + q^{not-shift}\right) * \left(\pi^{DA}(t_1) - \pi^{DA}(t_2)\right)$$

Using the values from scenario 4: $\tilde{q} = 55kW$, $q^{shift} = 25kW$ (we assume that some of the congestion can be removed with the help of the other aggregators who charge lower prices), $q^{not-shift} = 30kW$, $p(cong) = 0.5$, $\overline{p(cong)} = 0.5$, $\pi^{DA}(t_1) = 0.15\, Euro$, $\pi^{DA}(t_2) = 0.01\, Euro$. This yields a threshold price of $0.476\, Euro$ for $\pi^{LFM}(t_1, t_2)$.