Competition for flexible distribution resources in a 'smart' electricity distribution network*

Thomas P. Tangerås†

May 1, 2022

Abstract

In a 'smart' electricity distribution network, flexible distribution resources (FDRs) can be aggregated to improve efficiency in the power market. But aggregation enables whoever controls resources to exercise market power. This paper establishes a ranking of market structures: Independent aggregators competing for FDRs are more efficient than an integrated distribution system operator (DSO)/aggregator, which is more efficient than an integrated generator/aggregator, which is more efficient than no FDR market. A no-market solution is more efficient than an FDR market with independent aggregators and an integrated DSO/aggregator. The paper also characterizes a regulation that implements the efficient outcome.

Key words: Aggregator, balancing market, distribution system operator, market power, regulation, smart grid

JEL codes: H41; L12; L51; L94

*Many thanks to David Andrés-Cerezo, Therése Hindman Persson, Lennart Söder, seminar participants at the Swedish Competition Authority and the Swedish Energy Markets' Inspectorate and conference participants at the 9th Mannheim Conference on Energy and the Environment for comments and discussion. This work was conducted within the "Sustainable Energy Transition" research program at IFN. Financial support from the Swedish Energy Agency (2015-002474) is gratefully acknowledged. Data sharing is not applicable to this article as no datasets were generated or analysed during the current study.

†Research Institute of Industrial Economics (IFN), PO. Box 55665, SE-10215 Stockholm, Sweden. Telephone: +46(0)86654526. E-mail: thomas.tangeras@ifn.se. Personal website: www.ifn.se/thomast. Associated researcher with the Energy Policy Research Group (EPRG), University of Cambridge. Faculty affiliate at the Program on Energy and Sustainable Development (PESD), Stanford University.

1 Introduction

Over the last two decades, many jurisdictions have implemented support schemes to substantially increase electricity production from renewable energy sources. The trend towards more solar and wind power has made it increasingly challenging to maintain a balance between the electricity injected into the network and the electricity withdrawn from it. Keeping the balance between consumption and production is fundamental because electrical units and network components will otherwise disconnect, which can develop into a serious system interruption unless contained.[1] The difficulty of predicting the availability of variable renewable energy far in advance and the presence of network capacity constraints have increased the importance of local short-term balancing of electricity supply to offset the variability in renewable production. The transformation of the electricity system has also started to affect the nature of the system in the sense that imbalances have become more frequent and severe in the lower-voltage parts of the network instead of occurring mainly at the transmission level. A major explanation for this development is connection of small-scale renewable energy resources to the distribution grid. Phase-in of electric vehicles can be expected to add to future system complexity.

In a restructured electricity market, the responsibility for maintaining system balance is in the hands of a *system operator*. The system operator performs its task by procuring reserve capacity through long-term mechanisms or in a real-time market. Some also possess reserve capacity of their own. This capacity is activated in the degree necessary to offset imbalances between consumption and production relative to the scheduled dispatch.[2] The traditional approach to balancing the electricity system by expanding network capacity at all voltage levels and then leave it to a centralized system operator to balance the entire system, is challenged when local variability in demand and supply is large. This is particularly evident in those regions where network owners experience severe difficulties related to expansion of network capacity in the urban areas that need this capacity the most. Instead of continuing to rely entirely on centralized optimization at the transmission network level, some local imbalance problems might be solved more efficiently at the distribution network level (EDSO, 2015). In a more decentralized system, the responsibility for maintaining system balance within the different local distribution areas is delegated to a distribution system operator (DSO), for instance the owner of the distribution network that covers the local distribution area.

The advent of new technical solutions and improvements in information and communications technology have increased the potential for handling imbalance problems more at the distribu-

[1] The instantaneous balance is measured by the frequency with which the electric current oscillates through the grid. Network frequency increases (decreases) if the amount of electricity produced increases (decreases) relative to the amount that is consumed. All electrical units and machinery connected to the grid have the same nominal operating frequency. They will automatically disconnect if the actual frequency deviates too much from the nominal. The strain on the system caused by failed equipment can lead nearby units or network connections to become overloaded and also disconnect. In the worst case scenario, a domino effect can ripple through the network and cause system collapse. A famous example is the Northeast blackout of 2003 that started with a power plant in Ohio shutting down. The failure spread through the system as transmission lines sequentially tripped offline. The ensuing outage affected 10 million people in Ontario and 45 million people in eight U.S. states.

[2] As a last resort, the system operator has a mandate to physically ration demand or supply. A recent example of such curtailment occurred in August 2020 when the California Independent System Operator implemented rotating power outages to prevent electricity supply from breaking down during a heat wave (CAISO, 2021).

tion level. Households increasingly install digital thermostats that collect data on heating and cooling. Digital heating systems can be optimized to minimize heating costs by adapting consumption to the prices of electricity. The batteries of plug-in electric vehicles can be charged in such as way as to minimize user cost. Micro generation can be combined with distributed battery technologies to minimize household electricity costs. These and other *flexible distribution resources* (FDRs) can be used for a broader purpose than to optimize the energy use for individual households. The 'smart' electricity distribution network measures consumption and production at granular level and communicates with the different units in the network. Instead of having a strictly domestic application, distribution resources can then be aggregated and coordinated in an effort to optimize the whole electricity system.[3] A firm that coordinates distribution resources is known as an *aggregator*.[4]

This paper analyzes the emerging market for flexible distribution resources and the role of aggregation in relation to the utilization of those resources for system balancing purposes. A main issue is that control of substantial amounts of FDRs yields market power in the local balancing market. Market power raises a number of questions: How does the structure of the FDR market affect the efficiency with which FDRs are deployed? Is there a fundamental difference between a DSO handling those resources compared to one or more independent aggregators doing it? What if a generation company also participates in the FDR market? Is there a feasible way to regulate competition, and if so, what are the properties of such regulation?

There is increasing policy awareness of the benefits of aggregating distribution resources for system purposes. The common rules concerning the internal market for electricity in the European Union contain stipulations that aim to facilitate implementation of such solutions (EU, 2019). In the U.S., the Federal Energy Regulatory Commission recently adopted a reform with the stated purpose of removing barriers to entry of distributed energy resource aggregation. Participation of such resources, for instance has the potential to help "alleviate congestion and congestion costs during peak load conditions and to reduce costs related to transmitting energy into persistently high-priced load pockets" according to FERC (2020, para. 7). Answers to the questions addressed in this paper can be helpful in making informed policy choices regarding competition and regulation of the market for flexible distribution resources.

To analyze the performance of different market arrangements, I build in Section 2 a two-period model of an electricity system where demand fluctuates exogenously across periods. Consumption and production must be cleared in each period to maintain system balance. The standard way of balancing demand is by dispatching flexible generation. I assume that there is also a second way to accomplish system balance. Some of the electricity withdrawn from the grid goes into producing an energy service consumed by households. A prime example is indoor temperature control (heating and cooling). A constant indoor temperature can be achieved by different

[3] The term distributed energy resource (DER) is commonly used for describing a unit connected to the lower-voltage part of the network. Depending on the context, DERs may include also intermittent micro generation such as roof-top solar power and small-scale wind power. However, a DER must be *flexible* for it to be used for balancing purposes. To emphasize this distinction, this paper introduces the concept of flexible distribution resources (FDRs), which are those DERs that can be dispatched in a controllable and flexible manner.

[4] See, for instance, Burger et al. (2017) for a conceptual discussion of aggregators and their potential role in the electricity system.

combinations of electricity consumption in the two periods. Household utility depends on keeping an ideal indoor temperature, not on how this temperature is produced. By these properties, the heating system is a flexible distribution resource because it can substitute electricity consumption across periods. Intertemporal substitution of household consumption reduces system cost because it offsets the variability in demand and thus stabilizes the production of balancing power. Section 2 shows that the system can be perfectly stabilized if the variation in demand is sufficiently small relative to the capacity for intertemporal substitution. In this case, the efficient electricity consumption equalizes the marginal production cost across periods. If the variation in demand is large, then it is efficient to consume all electricity used in the provision of the energy service in the off-peak demand period. In this case, the marginal cost of flexible generation is higher in the peak than the off-peak period. Section 2 also demonstrates how to implement the efficient consumption and production of electricity in a decentralized market if the supply of electricity is perfectly competitive and all market participants are exposed to real-term prices.

There are plausible reasons why efficient consumption would be infeasible even in a competitive power market. First, consumers are at most exposed to hourly, or perhaps half-hourly, price changes. Because of the real-time variability of renewable electricity generation, minimization of system costs should have consumers reacting to even more frequent price changes. Second, consumers with the ability or technology to respond to price changes even within a shorter time horizon may nevertheless dislike price variability and therefore be unwilling to take the price risk associated with implementing the efficient solution. Third, even if short-time prices were allowed to vary depending on instantaneous changes in underlying system conditions and consumers were indeed risk neutral, they could still have insufficient incentives to respond to them. This happens if the perceived household cost savings associated with real-time optimization are small relative to the costs, for instance because of incremental investments households would have to make to acquire the necessary technology. To account for such consumption inefficiency, I assume that households are *passive consumers* in the sense that they would consume the same amount of electricity in both periods to produce the household energy service. Then there are potential efficiency gains associated with third parties assuming control over the production of the household energy service and coordinating the activation of flexible distribution resources across households. Section 3 evaluates and compares different market solutions for such aggregation and coordination.

I start the analysis of markets for flexible distribution resources by assuming in Section 3.1 that the regulated DSO and an unregulated aggregator belong to the same parent company. EU regulation allows such vertically integrated undertakings, provided the businesses operate independently. The aggregator is the only firm that participates in the market for FDRs. Each household pays the aggregator a fixed fee for supplying the energy service (for instance heat), but nothing for the electricity used in the production of the energy service. The aggregator's total profit consists of the fees collected from household customers minus the aggregator's cost of electricity, which it purchases from local generation owners. I assume that the aggregator pays a uniform price per unit of electricity consumed in each period, equal to that period's marginal cost of generation. Allocating more electricity consumption to one period drives up the price of electricity in that period and reduces the price of electricity in the other period. The integrated DSO/aggregator accomplishes a reduction in total electricity costs by increasing consumption in the peak and decreasing it in the off-peak demand period, compared to the efficient consumption. Exercise of market power thus distorts allocations by driving a wedge between the marginal cost of electricity production across periods.

An increasing number of firms provide systems for household resource management. Examples include Comverge, Energy Pool, Enbala, Enel X and Ngenic. Instead of minimizing household energy costs, such companies could equally well manage these resources by supplying balancing power. An example is the German company Sonnen, which combines roof-top solar power with battery technology. Under the sonnenFlat plan, customers pay a fixed fee for an annual consumption allowance. Sonnen then takes control over the system and uses it to supply the household and provide balancing power by optimizing storage in the system.[5] Formal *flexibility markets* that enable trade of balancing power at the distribution level are under development.[6] An aggregator whose only business is to coordinate such reserves is *independent*. Section 3.2 shows that independent aggregators competing in the market for flexible distribution resources increase efficiency by reducing market power in the flexibility market.

Section 3.3 considers the case where a DSO/aggregator competes with independent aggregators for customers in the market for FDRs. If the DSO is in charge of the local dispatch of electricity, then it will allocate the entire consumption of the independent aggregators to the peak demand period to minimize the electricity costs of its own aggregator. It also makes a difference whether aggregators are independent or integrated with generation units that supply balancing power. Section 3.4 shows that integration between generation and aggregation reinforces the incentives to increase the price of electricity in the peak demand period because the firm makes additional profit on its generation capacity, compared to the case of an independent aggregator. This is an instance in which vertical integration *reduces* market performance.

The different market structures are compared in terms of cost efficiency in Section 3.5. An integrated DSO/aggregator allocates FDRs less efficiently than multiple independent aggregators because of softer competition. Yet, it is more efficient for the DSO/aggregator to manage FDRs than for an integrated generator/aggregator to do so. This is still better than not activating them at all, since the firm allocates more electricity to the off-peak demand period than in the benchmark case with passive electricity consumption and no FDR market. The distortions associated with a mixed market structure are so severe that market efficiency can be *lower* than in the case without any market for flexible distribution resources.

Section 3.6 illustrates how competition can be regulated by way of compensation payments related to how aggregators allocate flexible distribution resources to the balancing market. This regulation implements the efficient allocations by making each firm residual party to any effi-

[5] https://sonnen.de/stromtarife/sonnenflat-home/, October 15, 2019.

[6] Most of these trading platforms are pilot projects. The Nordic power exchange, Nord Pool, has developed the platform *Nodes* for trading flexibility and energy at the distribution level. One application of this platform is *sthlmflex*, a flexibility market in the Stockholm region of Sweden. Examples of other flexibility market platforms include *Coordinet* (applied in Greece, Spain and Sweden), *Enera* (Germany), *Gopacs* (The Netherlands) and *Piclo flex* (Great Britain).

ciency loss it causes by its actions in the balancing market.

Section 4 extends the model to include imperfect competition in the production of electricity.[7] Most of the above results hold under this alternative assumption. An exception occurs when a producer with market power also supplies aggregation. Such vertical integration can increase efficiency. The integrated generator/aggregator now has an incentive to increase production in both periods to reduce the aggregator's cost of electricity. The integrated generator/aggregator has an additional incentive to allocate consumption to the off-peak period in order to minimize production cost.

Section 5 concludes the paper. Some proofs of formal statements are in the Appendix.

Contribution This paper examines how the incentives to deploy resources efficiently in the real-time balancing market depend on the structure of the market for flexible distribution resources. Burger et al. (2019a,b) provide a comprehensive description of the core activities performed in an electricity distribution system that incorporates flexible distribution resources, and they characterize the actors as well as their potential roles in that system. Their policy analysis emphasizes governance structure in relation to competition and coordination in the distribution system, and in particular the incentives of a network owner that also controls FDR capacity to distort competition in the FDR market by restricting access to the network or to system services. The present paper shows that short-term distortions are likely to persist because of imperfect competition, even if the regulatory framework can guarantee third-party equal access to the network.

Previous theoretical analysis has mainly focused on the incentives for network owners to procure such resources and how these incentives depend on regulatory policy; see for instance Brown and Sappington (2018, 2019). Kim et al. (2017) consider a model where multiple DSOs own local balancing capacity. Uncoordinated balancing is inefficient because of an assumed externality on the other DSOs. The paper devises an optimal cost-sharing mechanism of total balancing costs to ensure efficient real-time balancing of the market. In the present context, the single DSO does not own any capacity of its own and there are no externalities across DSOs. Instead, inefficiencies associated with local real-time balancing of the market stem from the exercise of market power in the deployment of these services.

The modeling framework in this paper bears resemblance to the classical Hotelling model of resource extraction which considers the problem of how much of a finite resource to extract today and how much to save for the future. Under imperfect competition, and if there are no production constraints, then the equilibrium in the Hotelling model is found at the point at which the marginal revenue is equalized across periods; see Tangerås and Mauritzen (2018) and the references therein.[8] However, this literature does not compare different market structures.

[7] Pollitt and Anaya (2021) discuss the potential for competition in real-term relative to the day-ahead market.

[8] A related literature analyzes strategic interaction between sequential markets by firms with market power, but without considering storage technologies; see, for instance Ito and Reguant (2016) and Rintamäki et al. (2020).

2 The efficient benchmark

This section first presents the theoretical model and then characterizes the efficient allocations associated with centralized dispatch that will be used to evaluate the different market structures. I demonstrate that a real-time market design can implement the efficient allocation if the market is perfectly competitive and all market participants face marginal real-time prices. I identify a fundamental inefficiency associated with this market design if consumers are passive.

The model Consider a two-period model of consumption and production within a local distribution area. In each period $i = 1, 2$, there is exogenous demand for $x_i > 0$ megawatt hours (MWh) electricity that is entirely unresponsive to price changes. This demand consists of industry and residential consumption minus the supply of variable renewable energy such as solar and wind power. This demand is also net of the transmission import capacity into the area. This means that (x_1, x_2) must be covered with local production to achieve a physical balance of electricity supply within the distribution area. Period 1 is the *peak demand period* and 2 the *off-peak demand period* if $x_1 > x_2$. The peak and off-peak definitions are reversed if $x_2 > x_1$. Although the production of variable renewable energy cannot be forecasted with any great precision far in advance, I let the time frame between period 1 and 2 be so short that both x_1 and x_2 are known entities at the start of period 1. This innocuous assumption avoids clouding the analysis with unnecessary stochastic expressions.

There is also a continuum of households of measure one that consumes an energy service in amount $\bar{s} > 0$. The energy service is produced by withdrawing $s_1 = s \in [0, \bar{s}]$ MWh electricity from the grid in period 1 and $s_2 = \bar{s} - s$ MWh in period 2. One can think of the energy service as the production of heat (or cooling) to generate indoor temperature $\bar{s}$. Different combinations of s and $\bar{s} - s$ produce by the laws of thermodynamics a constant indoor temperature $\bar{s}$. The energy service effectively allows substitution of electricity consumption across periods. Linearity and perfect substitutability across periods are only to keep things simple.[9] The local distribution area is *resource unconstrained* if the variation in exogenous demand across periods is small in the sense that $|x_1 - x_2| \leq \bar{s}$. Conversely, the area is *resource constrained* if the variation in exogenous demand is so large that $|x_1 - x_2| > \bar{s}$.

A flexible and dispatchable technology is available to cover total demand $q_i = x_i + s_i$ in both periods $i = 1, 2$. It is most relevant to think of this technology as local thermal electricity generation capacity, but it can also be other flexible consumption. Let $C(q)$ be the cost of supplying q MWh electricity. The continuous cost function $C(q)$ is the same in both periods and strictly increasing for all $q > 0$. The continuous marginal cost function $C'(q)$ is strictly increasing for $q > 0$, smooth and convex.

[9] Energy storage (e.g. batteries, plug-in electric vehicles) is another relevant example that enables intertemporal substitution. Storage technologies can be incorporated through a slight reformulation of the framework. Consider a battery with $2\bar{s}$ MWcharging capacity. Assume that this battery enters period 1 with a charge of $\bar{s}$ MW. The battery can then withdraw $s \in [-\bar{s}, \bar{s}]$ from the grid in period 1. Under the assumption that the battery must maintain charging capacity $\bar{s}$ after the end of period 2, it follows that $s_2 = -s$. Such modifications to the model would have no bearing on the results of the paper

Motivating example This paper is concerned with imbalances between supply and demand that must be resolved at the very local level. An example from Sweden illustrates the type of problem the paper is meant to address. The city of Stockholm, with one million inhabitants, relies mostly on surrounding regions for electricity supply. Import capacity into the Stockholm city distribution area is abundant on average. While the transmission network has an import capacity of 1 525 MW, the average annual hourly electricity consumption in the city is approximately 800 MWh (in 2018). However, hourly peak demand is more than double of the average consumption, and amounted to 1 721 MWh in 2018. In such instances, it is necessary to dispatch production within the local distribution area to cover excess demand in the city. This local generation capacity consists mainly of 320 MW combined heat and power (CHP). The owner of this generation capacity, *Stockholm Exergi*, does not always have enough financial incentive to dispatch the required capacity. Electricity production in the spot (day-ahead) market is remunerated on the basis of the much larger mid-Sweden bidding zone. This bidding zone contains all Swedish nuclear power, which means that the mid-Sweden spot price can clear at levels below the marginal cost of supplying electricity within Stockholm city. Tax increases have made local generation capacity even less profitable in an extent to which Stockholm Exergi has threatened to retire CHP units in Stockholm. Facing an increased risk of a power shortage, *Ellevio*, the owner of the Stockholm city distribution network, and *Svenska Kraftnät*, the owner of the Swedish transmission network, have signed a procurement contract with Stockholm Exergi to maintain the local generation capacity on the grid. These procurement costs are in turn passed onto consumers.

Given the small security margins in the network, 93% of total capacity was utilized to cover Stockholm peak demand in 2018, network owners want to increase the robustness and efficiency of electricity supply in Stockholm. The main instrument has been to create a flexibility market for the Stockholm region, called *sthlmflex*. The goal for the winter of 2021/22, was to activate at least 53 MW flexible generation and consumption capacity to smooth out peak demand.

The main problem that *sthlmflex* is meant to mitigate is the internal congestion problem within the mid-Sweden bidding zone, which creates excess demand in Stockholm at the mid-Sweden spot price. Such bottlenecks can be resolved by increasing the number of bidding zones. These are difficult to implement in practice because they create regional price differences that are politically controversial.[10] However, local supply imbalances can occur even under fully-fledged nodal pricing that accounts for *all* transmission bottlenecks by allowing unique prices at every single node in the network. Network congestion in the lower-voltage distribution network cannot be resolved locational prices at the transmission level. In those situations, the market solutions discussed in this paper can improve efficiency even under nodal pricing.

The efficient allocation The central planner chooses $s \in [0, \bar{s}]$ to minimize the total cost

$$C(x_1 + s) + C(x_2 + \bar{s} - s) \tag{1}$$

[10] In most European countries there is just one national bidding zone, which implies that consumers pay the same price for electricity regardless of their location in the network.

of clearing consumption and production in the two periods. Let s^{fb} be the efficient amount of electricity used in the first period to produce the energy service. Let $q_1^{fb} = x_1 + s^{fb}$ and $q_2^{fb} = x_2 + \bar{s} - s^{fb}$ denote efficient production within the local distribution area in each of the two periods. Then (the proof is in Appendix A.1):

Lemma 1 *Under the efficient allocation $(s^{fb}, q_1^{fb}, q_2^{fb})$:*
(i) The electricity used in the production of the energy service is withdrawn from the grid in such a way as to smooth out all variations in marginal production costs across periods if the local distribution area is resource unconstrained [$s^{fb} = s^* = \frac{1}{2}(\bar{s} + x_2 - x_1)$ and $q_1^{fb} = q_2^{fb} = q^* = \frac{1}{2}(\bar{s} + x_1 + x_2)$ if $|x_1 - x_2| \leq \bar{s}$].
(ii) The energy service is produced entirely by withdrawing electricity from the grid in the off-peak demand period if the local distribution area is resource constrained. If so, then local production is larger in the peak relative to the off-peak demand period [$s^{fb} = 0$ if $x_1 - x_2 > \bar{s}$ and $s^{fb} = \bar{s}$ if $x_2 - x_1 > \bar{s}$. Moreover, $(q_1^{fb} - q_2^{fb})(x_1 - x_2) > 0$ if $|x_1 - x_2| > \bar{s}$].

The intertemporal aspect of the energy service production function allows to smooth out exogenous fluctuations in demand across periods and thus achieve full efficiency by eliminating the variability in the marginal production costs, if the variability in renewable production is sufficiently small relative to the capacity for intertemporal substitution of electricity consumption, i.e. $|x_1 - x_2| \leq \bar{s}$. In more extreme cases, the central planner adapts to resource constraints by withdrawing all electricity that goes into producing the energy service in the off-peak demand period and clears excess demand by dispatching relatively more production resources in the period with the highest demand.

Decentralized market implementation I now establish conditions under which a decentralized market can implement the efficient allocation. Assume that all production and consumption is cleared in a real-time market operating each period $i = 1, 2$. All production in period i is remunerated at the marginal real-time price $p_i = C'(q_i)$ and all consumption in that period also pays the marginal real-time price p_i.

Assume that all market participants are price-takers. This implies that all thermal electricity is bid in at marginal cost $C'(q_i)$. If all market participants expect the market to clear at the efficient real-time price $p_i^{fb} = C'(q_i^{fb})$ in each period i, then the production of thermal electricity equals q_i^{fb} in each period. Consumption of the energy service costs $p_1^{fb}s + p_2^{fb}(\bar{s} - s)$. If the area is resource constrained because of peak demand in period 1, $x_1 - x_2 > \bar{s}$, then $p_1^{fb} = C'(q_1^{fb}) > C'(q_2^{fb}) = p_2^{fb}$ by Lemma 1. In this case, the household minimizes expenditures by consuming all electricity in the second period: $s = 0 = s^{fb}$. Conversely, the household consumes all electricity in the first period, $s = \bar{s} = s^{fb}$, if $x_2 - x_1 > \bar{s}$ because then $p_2^{fb} > p_1^{fb}$. Finally, the real-time price is the same in both periods, $p_1^{fb} = p_2^{fb} = p^* = C'(q^*)$, if the control area is resource unconstrained, $|x_1 - x_2| \leq \bar{s}$. In that case, the household's total expenditure $p^*\bar{s}$ on the energy service is independent of s, so it is individually rational to set $s = s^* = s^{fb}$. Hence, a competitive real-time market can implement the efficient allocation as a decentralized equilibrium. The following implications are immediate:

Corollary 1 *In a competitive real-time market that implements the efficient allocation* $(s^{fb}, q_1^{fb}, q_2^{fb})$*, the real-time price is weakly higher in the peak than the off-peak period, and the representative household consumes relatively more electricity in the low-price compared to the high-price period [*$(p_1^{fb} - p_2^{fb})(x_1 - x_2) \geq 0$ *and* $(p_1^{fb} - p_2^{fb})(\frac{1}{2}\bar{s} - s^{fb}) \geq 0$ *with strict inequalities if and only if* $|x_1 - x_2| > \bar{s}$*].*

The cost of passive consumption As argued in the introduction, there are plausible reasons why the efficient allocation of electricity would be infeasible even in a competitive market. In particular, households typically are not exposed to marginal real-time prices. Instead, consumption decisions are most often based on average prices. To account for this lack of contractual flexibility, I henceforth assume that households pay the average price $\frac{1}{2}(p_1 + p_2)$ for consumption of the energy service. Households then cannot strictly benefit from varying s across the two periods because their cost of electricity is $\frac{1}{2}(p_1 + p_2)\bar{s}$ regardless of s. They might as well remain *passive* and withdraw the same amount of electricity from the grid in period 1 as in period 2 to produce the energy service, in which case $s = \frac{1}{2}\bar{s}$. Under such passive household consumption, the thermal electricity production in the local distribution area that solves the balancing problem in period $i = 1, 2$ equals $\bar{q}_i = x_i + \frac{1}{2}\bar{s}$. This solution generally is inefficient:

Proposition 1 *Passive households withdraw too much electricity in the peak relative to the off-peak demand period, compared to the efficient withdrawal. Real-time system balancing then requires excessive local production in the peak demand period to offset this inefficiency [*$(\frac{1}{2}\bar{s} - s^{fb})(x_1 - x_2) > 0$*,* $(\bar{q}_1 - q_1^{fb})(x_1 - x_2) > 0$ *and* $(\bar{q}_2 - q_2^{fb})(x_2 - x_1) > 0$ *for all* $x_1 \neq x_2$*]*

Proof: By way of Lemma 1, $(\frac{1}{2}\bar{s} - s^{fb})(x_1 - x_2) = \frac{1}{2}(x_1 - x_2)^2$ for $|x_1 - x_2| \leq \bar{s}$ and $(\frac{1}{2}\bar{s} - s^{fb})(x_1 - x_2) = \frac{1}{2}\bar{s}|x_1 - x_2|$ for $|x_1 - x_2| > \bar{s}$. To see the second part of the proposition, plug in (q_1^{fb}, q_2^{fb}) and $(\bar{q}_1, \bar{q}_2)$ to get $(\bar{q}_1 - q_1^{fb})(x_1 - x_2) = (\bar{q}_2 - q_2^{fb})(x_2 - x_1) = \frac{1}{2}(x_1 - x_2)^2$ for $|x_1 - x_2| \leq \bar{s}$ and $(\bar{q}_1 - q_1^{fb})(x_1 - x_2) = (\bar{q}_2 - q_2^{fb})(x_2 - x_1) = \frac{1}{2}\bar{s}|x_1 - x_2|$ for $|x_1 - x_2| > \bar{s}$.■

The $(\frac{1}{2}\bar{s}, \bar{q}_1, \bar{q}_2)$ allocation associated with passive household energy consumption is inefficient relative to the efficient allocation $(s^{fb}, q_1^{fb}, q_2^{fb})$ except in the knife-edge case where exogenous demand is the same in both periods, $x_1 = x_2$, because the difference in system marginal costs across periods is no longer minimized. This inefficiency opens up for solutions that aim to achieve *flexible* household electricity consumption that reacts to marginal real-time price signals. This paper considers a market for flexible distribution resources (FDR).

3 A market for flexible distribution resources

Assume from now on that neither centralized allocation nor the fully decentralized market with marginal real-time price exposure for all consumers is feasible. All household consumers are passive. A distribution system operator (DSO) is responsible for maintaining system balance within the distribution area in each period. This DSO owns the local distribution network, but has no production capacity of its own. Therefore, it has to ensure that generation owners supply the required balancing power to match production with consumption in every period. The DSO can either sign contracts with generation owners or organize trade of balancing power through a local real-time market, which I refer to as a *flexibility market*. Regardless, I assume that dispatchable generation is supplied at marginal cost and that all activated balancing power receives the same compensation, equal to the marginal cost of the most expensive unit that is activated. Hence, $P(q_i) = C'(q_i)$ represents the inverse supply function of dispatchable electricity generation q_i in period i.[11] The assumption that thermal electricity is supplied by generation owners at marginal cost allows to isolate the effects on efficiency of imperfect competition in the market for flexible distribution resources. Yet, most of the results carry over to a setting with market power in the supply of thermal electricity; see Section 4.

I assume that smart grid solutions can also be utilized in order to accomplish system balance. Specifically, the network infrastructure is 'smart' in the sense that one or more external parties can assume the task of supplying the energy service $\bar{s}$ to households by remote control of s. This remote control is achieved through online digital thermostats that adjust indoor temperatures. The economic incentive to assume this responsibility comes from the ability to supply the flexible distribution resource to balance electricity supply within the local distribution area. I refer to $\bar{s}$ as the size of the market for flexible distribution resources (FDRs). The key policy question is how this market should be structured to maximize efficiency. To address this question, I examine four market structures.

3.1 Integrated DSO/aggregator

The DSO procures balancing power through contracts with generation owners. DSOs are regulated because of their monopoly ownership of the local distribution network. The constraint on network tariffs often is in the form of a cap on the amount of revenues the DSO can collect from customers. However, the cost $P(q_1)q_1 + P(q_2)q_2$ of procuring balancing power is often considered unavoidable and therefore added in full to the revenue cap. Under this assumption, the total revenue equals $F + P(q_1)q_1 + P(q_2)q_2$, where F is the revenue cap on all services other than balancing power delivered by the DSO. The DSO has net profit equal to F under such complete pass-through, which is independent of q_1 and q_2. Consequently, the DSO has no incentive to ensure efficient supply of balancing power under this regulatory regime. On the contrary, the DSO has a strict incentive *not* to implement efficient solutions if such implementation creates non-monetary costs that the DSO cannot pass onto consumers (Kim et al., 2017).

EU regulation generally permits parent companies to own business entities that operate distribution networks as well as entities that participate in the deregulated parts of the electricity market, provided these entities are independent in terms of "legal form, organisation and decision-making" (Article 35 of EU, 2019). Under such *vertical unbundling*, a parent company

[11] An alternative to marginal pricing would be pay-as-bid pricing. Under this market design, consumers pay and producers earn based on their individual demand and supply bids, much like in a first-price auction. Pay-as-bid pricing would substantially complicate the analysis, not least because market participants have an incentive to shade their bids even in a competitive framework. Current flexibility markets can be organized both as pay-as-bid and marginal pricing markets. An example of the latter is Cornwall Local Electricity Market.

that owns a DSO can establish another firm called an *aggregator* that can enter the market for energy services.[12] I assume in this section that this is the only aggregator in the market.

The markets for flexible distribution resources and balancing power operate as follows in this particular setting. The aggregator first approaches households and offers to supply the energy service $\bar{s}$ in return for a fixed fee t. There is an upper bound $\bar{t} > 0$ to the fixed fee that is acceptable to the customer, for instance because households always can consume the energy service passively at a cost per MWh equal to the average real-time price. The aggregator then purchases from the generation owners within the local distribution area the amount of electricity s and $\bar{s}-s$ that is necessary to supply the energy service, at total cost $P(q_1)s+P(q_2)(\bar{s}-s)$. The DSO purchases residual demand x_1 and x_2 from generation owners at cost $P(q_1)x_1+P(q_2)x_2$.[13] This cost is passed onto consumers through the revenue cap. The joint profit of the monopoly DSO/aggregator thus equals:

$$\Pi^{DSO}(s) = F + \bar{t} - P(x_1 + s)s - P(x_2 + \bar{s} - s)(\bar{s} - s). \tag{2}$$

Two things are worth noticing about the profit expression $\Pi^{DSO}(s)$. First, electricity withdrawal s matters for profit even if the DSO is subject to a revenue cap because the aggregator is residual claimant to any of its associated cost savings. Second, the aggregator faces the *inverse supply function* $p_i = P(q_i)$ in the real-time market in period i. In other words, monopoly control over flexible distribution resources implies monopsony power. This is unlike in the benchmark case of a fully decentralized market where s is decided by a representative household that takes prices as exogenously given.[14]

The DSO/aggregator has an incentive to allocate consumption to the first period (increase s) if $p_1 < p_2$:

$$\Pi^{DSO\prime}(s) = p_2 - p_1 - P'(q_1)s + P'(q_2)(\bar{s} - s). \tag{3}$$

It also takes into account the increase in the first period price and the reduction in the second period price resulting from the reallocation. This market power effect is identified by the two last terms in (3). Let $s^{DSO} \in [0, \bar{s}]$ be the equilibrium amount of electricity withdrawn by the DSO/aggregator in period 1 to supply the energy service to households, so that $\bar{s} - s^{DSO}$ is the amount withdrawn in period 2. The proof of the following result is in Appendix A.2:

Proposition 2 *Consider the electricity withdrawal s^{DSO} of an integrated monopoly DSO/aggregator:*
(i) Electricity withdrawal in the peak demand period exceeds the corresponding efficient with-

[12] The rules concerning vertical unbundling of DSOs are not mandatory for vertically integrated undertakings that serve fewer than 100 000 connected customers (Article 35 of EU, 2019).

[13] Households typically not only consume s and $\bar{s} - s$ in the two periods, but also x_1 and x_2. Aggregators are only viable if the electricity that goes into producing the energy service can be metered separately from the other household consumption. Such granular metering is a prerequisite for flexible consumption through third-parties.

[14] There is an alternative way to derive the profit expression (2). Assume that only the DSO contracts with the owners of dispatchable generation capacity for balancing power, whereas the aggregator purchases electricity from the DSO at cost $P(q_1)s + P(q_2)(\bar{s} - s)$. The DSO's net cost of purchasing balancing power then equals

$$P(q_1)q_1 + P(q_2)q_2 - P(q_1)s - P(q_2)(\bar{s} - s) = P(q_1)x_1 + P(q_2)x_2,$$

which is passed onto consumers. Summing up aggregator and DSO profit yields (2).

drawal $[(s^{DSO} - s^{fb})(x_1 - x_2) \geq 0$ *with strict inequality if* $|x_1 - x_2| \in (0, \bar{s}]]$.
(ii) The DSO withdraws more electricity from the grid in the off-peak relative to the peak demand period to supply the household energy service $[(\frac{1}{2}\bar{s} - s^{DSO})(x_1 - x_2) > 0$ *for all* $x_1 \neq x_2]$.

The DSO/aggregator wields market power by its control of $\bar{s}$ and uses it to minimize the total cost of supplying the energy service. The firm does not internalize the price effect on the cost of consuming x_1 and x_2 by full pass-through of $p_1x_1 + p_2x_2$ to consumers. The consumption externality plus market power cause the inefficiency in this market.[15]

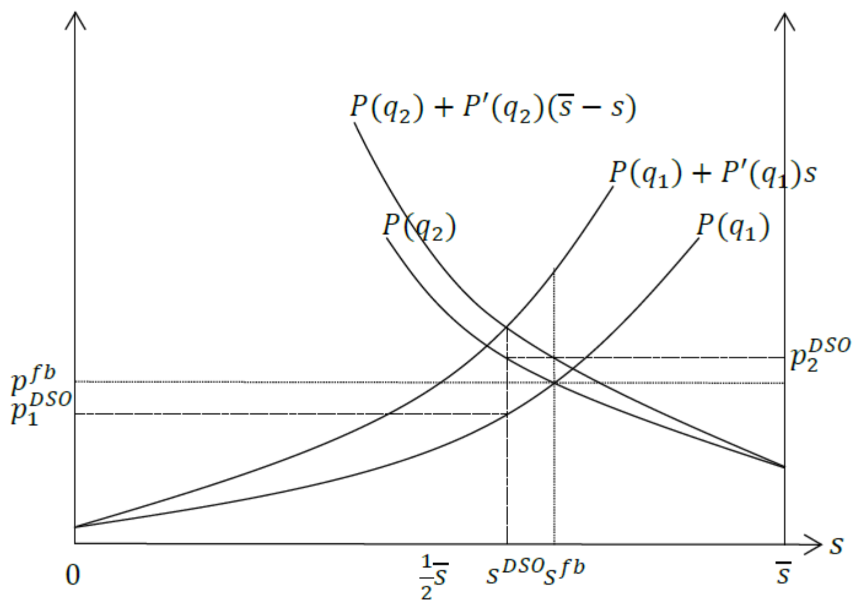

Figure 1: Monopoly DSO

To see how market power affects the equilibrium allocation, consider a classical "bathtub" diagram depicted in Figure 1. The horizontal axis measures the withdrawal of electricity in the first period from left to right and in the second period from right to left. At $s = 0$, all electricity used in the production of the energy service is consumed in period 2, and at $s = \bar{s}$ it is consumed in its entirety in period 1. The left-most vertical axis measures marginal system costs and marginal expenditures in period 1 and the right-most axis the marginal system costs and marginal expenditures in period 2. The two curves $P(q_1)$ and $P(q_2)$ are the inverse supply functions of electricity in each period. At $s = \frac{1}{2}\bar{s}$, the system marginal cost is smaller in period 1 than period 2, $P(x_1 + \frac{1}{2}\bar{s}) < P(x_2 + \frac{1}{2}\bar{s})$, thus establishing period 2 as the peak-demand period.

An increase in electricity withdrawal s in period 1 drives up the system marginal cost in period 1, but reduces it in period 2. The control area is resource unconstrained by assumption, so the efficient consumption is found at the point $s = s^{fb} \in (\frac{1}{2}\bar{s}, \bar{s})$ at which the system marginal

[15] Market power alone is not necessarily a source of inefficiency. Let the DSO/aggregator minimize the total system cost $P(x_1 + s)(x_1 + s) + P(x_2 + \bar{s} - s)(x_2 + \bar{s} - s)$. Then, s solves $P(q_1) + P'(q_1)q_1 = P(q_2) + P'(q_2)q_2$ in interior optimum. The profit maximizing choice equals $s = s^* = s^{fb}$ if $|x_1 - x_2| \leq \bar{s}$ because then $q_1 = q_2 = q^*$.

costs are equated across the two periods. However, the marginal benefit to the DSO/aggregator of increasing s is not measured in terms of the difference $P(q_2) - P(q_1)$ in system marginal costs. Instead, the DSO/aggregator accounts also for the price increase in period 1 and the price decrease in period 2 because these changes affect the total cost of the aggregator. The marginal expenditure of increasing electricity withdrawal in period 1 is illustrated in Figure 1 from left to right by $P(q_1) + P'(q_1)s$, whereas the marginal expenditure of increasing electricity withdrawal in period 2 is illustrated the figure from right to left to right by $P(q_2)+P'(q_2)(\bar{s}-s)$. Starting at the efficient consumption s^{fb}, a marginal reduction in s below s^{fb} reduces the period 1 expenditure by $p^{fb}+P'(q^{fb})s^{fb}$ and increases it by $p^{fb}+P'(q^{fb})(\bar{s}-s^{fb})$ in period 2 because the DSO/aggregator reallocates electricity consumption from the off-peak demand period 1 to the peak demand period 2. This manipulation is strictly profitable to the DSO/aggregator because electricity withdrawal is strictly larger in period 1 than period 2, $s^{fb} > \frac{1}{2}\bar{s}$. The allocation that minimizes the total cost is found at the point $s^{DSO} \in (\frac{1}{2}\bar{s}, s^{fb})$ at which the marginal expenditure is the same in both periods. The corresponding equilibrium prices are $p_1^{DSO} = P(x_1 + s^{DSO})$ and $p_2^{DSO} = P(x_2 + \bar{s} - s^{DSO})$.

Exploitation of market power causes the DSO/aggregator to withdraw too much electricity from the grid in the peak period and too little in the off-peak period, compared to the efficient consumption, $s^{DSO} < s^{fb}$. However, the DSO/aggregator still withdraws more electricity from the grid in the off-peak than the peak demand period, $s^{DSO} > \frac{1}{2}\bar{s}$, because the difference in demand implies that the marginal cost is lower in the off-peak compared to the peak demand period, and the DSO/aggregator therefore can reduce spending by allocating relatively more consumption to the off-peak demand period.

The exercise of market power in Figure 1 drives up the price of electricity in period 2 relative to period 1 in equilibrium, $p_2^{DSO} > p_1^{DSO}$, even if there is no real scarcity of resources, $|x_1 - x_2| \in (0, \bar{s}]$, so that full price equalization would be feasible and efficient. The following corollary to Proposition 2 therefore arises (the proof is in Appendix A.3):

Corollary 2 *Exercise of market power leads to excessive price fluctuations in the market for balancing power, compared to the efficient outcome [*$(p_1^{DSO} - p_2^{DSO})(x_1 - x_2) > 0 = (p_1^{fb} - p_2^{fb})(x_1 - x_2)$ *for* $|x_1 - x_2| \in (0, \bar{s}]$*].*

3.2 Independent aggregators

The DSO operates a flexibility market to clear demand x_1 and x_2. The electricity $\bar{s}$ used for producing the energy service is purchased by $A \geq 1$ *independent aggregators* whose only business is to supply the energy service to households. The DSO does not have any stake in the flexibility market. It therefore has no incentive to behave strategically because it passes all electricity costs onto consumers. However, aggregators exercise market power.

I solve in an online appendix (Tangerås, 2022) a model in which aggregators are horizontally differentiated and compete for household consumers in period 0 before (x_1, x_2) is known. Each aggregator $a \in \{1, ..., A\}$ offers to supply the energy service $\bar{s}$ in return for a fixed fee t_a. The fees chosen by the aggregators yield a distribution $(L_1, ..., L_a, ..., L_A)$ of market shares. Aggregator a

then purchases $s_a \in [0, L_a\bar{s}]$ in the flexibility market in period 1 and the remaining $L_a\bar{s} - s_a$ in the period 2 to supply the energy service to its customers. The measure $s = \sum_{a=1}^{A} s_a$ represents the aggregators' total demand for electricity in period 1. Their total demand in period 2 equals $\bar{s} - s$. Tangerås (2022) establishes a sufficient condition for when all aggregators charge a symmetric fee t^A in equilibrium, all aggregators have the same market share $\frac{1}{A}$, and the market is fully covered (all households purchase the energy service from some aggregator). The equilibrium fee affects the distribution of rent between households and aggregators, but not efficiency.

Aggregators compete in quantities in the flexibility market, so the profit of aggregator a is

$$\Pi^A(t^A, s_a) = \frac{1}{A}t^A - P(x_1 + s_a + \frac{A-1}{A}s^A)s_a - P(x_2 + \bar{s} - s_a - \frac{A-1}{A}s^A)(\frac{1}{A}\bar{s} - s_a). \tag{4}$$

if all aggregators have the same market share $\frac{1}{A}$, and all of them, except possibly a, withdraw the equilibrium amount of electricity $\frac{1}{A}s^A$ from the grid in the first period. Aggregator a's marginal incentive to increase consumption in the first period is

$$\frac{\partial \Pi^A(t^A, s_a)}{\partial s_a}|_{s_a = \frac{1}{A}s^A} = p_2^A - p_1^A - \frac{1}{A}P'(q_1^A)s^A + \frac{1}{A}P'(q_2^A)(\bar{s} - s^A) \tag{5}$$

evaluated at $s_a = \frac{1}{A}s^A$. In the above expression, $q_1^A = x_1 + s^A$ and $q_2^A = x_2 + \bar{s} - s^A$ measure the equilibrium production in period 1 and 2, respectively, in the market with aggregators, whereas $p_1^A = P(q_1^A)$ and $p_2^A = P(q_2^A)$ are the equilibrium prices. Aggregators have an incentive to allocate electricity purchases to the period with the lowest real-time price. However, market power causes them also to take into account how their demand in the flexibility market affects prices. The marginal price effects in (5) are smaller in magnitude when A is larger because a larger share of the marginal price effects then spill over to the other aggregators in the market (the proof is in Appendix A.4):

Proposition 3 *Consider the electricity withdrawal s^A of $A \geq 1$ independent aggregators with symmetric market shares in the market for flexible distribution resources:*
(i) Electricity withdrawal in the peak demand period exceeds the corresponding efficient withdrawal, but less so when there are more aggregators [$(s^A - s^{fb})(x_1 - x_2) \geq 0$ with strict inequality if $|x_1 - x_2| \in (0, \bar{s}]$. Moreover, $\frac{d}{dA}(s^A - s^{fb})(x_1 - x_2) < 0$ if $s^A \in (0, \bar{s})$ and $x_1 \neq x_2$].
(ii) The equilibrium converges to the efficient allocation in the limit when the number of aggregators becomes sufficiently large [$s^A \to s^{fb}$ for $A \to \infty$].
(iii) Aggregators withdraw more electricity from the grid in the off-peak and less electricity in the peak demand period than a monopoly DSO/aggregator [$(s^{DSO} - s^A)(x_1 - x_2) \geq 0$ with strict inequality if and only if $A \geq 2$, $x_1 \neq x_2$ and $s^{DSO} \in (0, \bar{s})$].

Aggregators exercise market power by consuming too much electricity in the peak demand period, compared to the efficient allocation, but a more fragmented market structure mutes the incentive to exercise market power. By implication, real-time prices are (at least weakly) more stable in a market with independent aggregators compared to the monopoly DSO/aggregator solution. Competition among independent aggregators both improves efficiency in the upstream flexibility market for electricity and reduces rent in the downstream energy services market.

3.3 Mixed market structure

This section extends the analysis of Section 3.1 to include $A-1 \geq 1$ independent aggregators in addition to the one aggregator owned by the same parent company that owns the distribution network. A classical concern is that a vertically integrated firm can benefit from limiting other firms' access to an essential facility (Rey and Tirole, 2007), the distribution network in this case. Regulatory policies that mandate network owners to grant competitors access to the network on non-discriminatory terms are standard in many electricity markets to prevent such direct foreclosure.[16] Yet, a non-discrimination policy may be insufficient to generate a viable market if the network owner can squeeze competitors in other ways than through direct foreclosure. This risk is obvious if aggregators pay the cost of their electricity consumption directly to the DSO who in turn contracts with owners of flexible generation capacity for balancing power. In this case, the DSO can drive competing aggregators out of the FDR market by overcharging all aggregators by the same amount for the cost of electricity. A solution to the problem of foreclosure through inflated system costs is to establish a real-time (flexibility) market in which aggregators purchase their electricity on market-based terms directly from generation owners. The costs of the aggregators' electricity consumption then depend on the market-clearing prices. Still, market-based electricity prices may be insufficient to create a level playing field for aggregators.

Let the DSO operate a flexibility market as described in Section 3.2. Assume that the integrated DSO/aggregator has market share L^{DSO} in the FDR market. Let $\bar{s}^{DSO} = L^{DSO}\bar{s}$ be the total electricity demand of this aggregator, and denote by $\bar{s}^A = \bar{s} - \bar{s}^{DSO}$ the total electricity demand of the independent aggregators. The integrated DSO/aggregator has profit

$$F + L^{DSO}t^{DSO} - P(x_1 + s^{DSO} + s^A)s^{DSO} - P(x_2 + \bar{s} - s^{DSO} - s^A)(L^{DSO}\bar{s} - s^{DSO}) \quad (6)$$

depending on the amount of electricity $s_1^{DSO} = s^{DSO}$ the subsidiary withdraws from the grid in the first period [and $s_2^{DSO} = \bar{s}^{DSO} - s^{DSO}$ in period 2] to supply the energy service to its households, and on the amount $s_1^A = s^A$ of electricity consumed by the $A-1$ aggregators in the first period [and $s_2^A = \bar{s}^A - s^A$ in period 2] to supply the energy service to households that purchase this service from an independent aggregator. Then, $q_i = x_i + s_i^{DSO} + s_i^A$ is the production of electricity in period $i = 1, 2$. The price of electricity is given by $p_i = P(q_i)$.

Under fully decentralized dispatch, the outcome under a mixed market structure is the same as the one with A independent aggregators. Allowing a DSO/aggregator to participate in the flexibility market represents a competitive benefit in this case. In reality, participants in a balancing market commit their capacity for a fixed time interval. The system operator then dispatches electricity to accomplish a constant balance of production and consumption within the duration of this time interval. A simple way to incorporate such DSO dispatch into the model is to assume that generation owners commit to the inverse supply function $C'(q)$ for each

[16] For instance, Article 31 of EU (2019) states that "the distribution system operator shall not discriminate between system users or classes of system users, particularly in favour of its related undertakings."

period before the start of period 1, the independent aggregators jointly commit to purchasing $\bar{s}^A$ MWh electricity over the two periods as a whole, and the subsidiary aggregator likewise commits to purchasing $\bar{s}^{DSO}$. The allocation of electricity consumption across periods is left to the DSO. Under complete economic unbundling between the DSO and the subsidiary aggregator, the DSO does not have any incentive to distort the dispatch of electricity. Here, I explore polar opposite assumption of full economic integration between the DSO and the aggregator. Specifically, the DSO allocates $(\bar{s}^{DSO}, \bar{s}^A)$ to maximize the joint profit (6). This profit function is non-concave. To simplify the analysis, I therefore impose a bounded elasticity assumption

$$\frac{C'''(q_i)s_i^{DSO}}{C''(q_i)} < 1 + \frac{C'''(q_1)s_1^{DSO} + C'''(q_2)s_2^{DSO}}{C''(q_1) + C''(q_2)}, \forall s_i^A \in [0, \bar{s}^A],\ s_i^{DSO} \in [0, \bar{s}^{DSO}],\ i = 1, 2 \quad (7)$$

on the cost function.[17]

Let period 1 be the peak demand period. Independent aggregators then can save on electricity costs by reducing the off-peak price p_2 and increasing the peak price p_1 since they consume most of their electricity in the off-peak period. Independent aggregators achieve this price adjustment by increasing their own consumption in period 1. A DSO/aggregator that controls dispatch can accomplish this manipulation in a much more profitable way by maximizing the consumption of the *other* aggregators in peak period 1 and then allocate most of its own consumption to the off-peak period 2 (the proof is in Appendix A.5):

Proposition 4 *Consider a mixed market structure in which a DSO/aggregator and $A-1$ independent aggregators compete in the market for flexible distribution resources. Assume that the DSO allocates $(\bar{s}^A, \bar{s}^{DSO})$ to maximize the joint DSO/aggregator profit.*
(i) The DSO allocates most of its own electricity consumption to the off-peak demand period and all of the consumption of the independent aggregators to the peak demand period [$s^{DSO} < \frac{1}{2}\bar{s}^{DSO}$ and $s^A = \bar{s}^A$ if $x_1 > x_2$, $s^{DSO} > \frac{1}{2}\bar{s}^{DSO}$ and $s^A = 0$ if $x_1 < x_2$].
(ii) More electricity is withdrawn from the grid in the peak relative to the off-peak demand period to produce the household energy service if the DSO controls half or less of the market for flexible distribution resources [There exists $\hat{L}^{DSO} \in [\frac{1}{2}, 1)$ such that $(s^{DSO} + s^A - \frac{1}{2}\bar{s})(x_1 - x_2) > 0$ for all $x_1 \neq x_2$ if and only if $L^{DSO} < \hat{L}^{DSO}$].

The above analysis has treated the shares of the aggregators in the FDR market as exogenous. Tangerås (2022) develops a model that determines these market shares in equilibrium. An interesting outcome is that an equilibrium in which independent aggregators have a positive market share may fail to exist if the DSO dispatches electricity in such a way as to maximize the joint DSO/aggregator profit. The DSO's incentive to allocate the electricity consumption of independent aggregators to the peak demand period may drive up these aggregators' cost of electricity to such an extent as to make it impossible for them to compete on profitable terms in the FDR market. In that case, the only viable market structure is a DSO/aggregator monopoly.

[17] The cost function $C(q) = bq^{2+\sigma}$, $b > 0$, $\sigma \in [0, 1]$ satisfies the conditions of this assumption as well as the others imposed on the model. The left-hand side of (7) can be written as $\sigma\frac{s_i^{DSO}}{q_i}$ under this parametrization of the cost function. This expression is below 1 by the assumption that $\sigma \leq 1$, and since $q_i = x_i + s_i^{DSO} + s_i^A > s_i^{DSO}$.

3.4 Integrated generator/aggregator

It is likely that firms already present in the balancing market will compete in the market for flexible distribution resources. To assess such vertical integration, assume that only one firm participates in the market for FDRs and that this firm also owns a thermal generation facility with production capacity $y > 0$ within the control area. I also maintain the assumption that thermal electricity is competitively supplied. For instance, the integrated generator's production capacity can be so small that it is a profit-maximizing choice to supply the entire capacity into the flexibility market in both periods. Specifically, $y < \frac{1}{2}\min\{x_1, x_2\}$. The assumption of perfect competition delivers some robust results and facilitates the comparison with previous results. I discuss the consequences of market power in generation in Section 4. Perfect competition on the production side gives the period i electricity price $p_i = P(q_i) = C'(q_i)$, where $q_i = x_i + s_i - y$ measures residual production.

The integrated generator/aggregator maximizes

$$\Pi^I(s) = P(x_1 + s - y)(y - s) + P(x_2 + \bar{s} - s - y)(y - \bar{s} + s) \tag{8}$$

over s (the production cost of the large generator is uninteresting here because output is fixed at y in both periods). To ensure concavity of $\Pi^I(s)$, I impose the property

$$\frac{C'''(x_i + s_i - y)(y - s_i)}{C''(x_i + s_i - y)} < 1 \ \forall s_i \in [0, \bar{s}],\ i = 1, 2 \tag{9}$$

on the cost function $C(q)$.[18]

It is straightforward to verify that the firm's marginal revenue is smaller in the peak than the off-peak demand period, evaluated at $s = \frac{1}{2}\bar{s}$. It is therefore profitable to consume more in the off-peak than the peak demand period. The marginal profit of the integrated generator/aggregator can be written as

$$\Pi^{I\prime}(s) = \Pi^{DSO\prime}(s) + [P'(q_1) - P'(q_2)]y,$$

where I have substituted in the marginal profit of an integrated DSO/aggregator; see eq. (3). The second part of this expression measures the marginal effect of vertical integration. The marginal price effect of increasing demand is larger in the peak price than the off-peak price period. The integrated generator/aggregator therefore has an additional incentive to allocate consumption to the peak price period because this reallocation increases the profit on generation (the proof is in Appendix A.6):

Proposition 5 *An integrated generator/aggregator that supplies its thermal generation competitively into the balancing market and has a monopoly in the market for flexible distribution resources, withdraws more electricity s^I from the grid in the off-peak than the peak demand period $[(\frac{1}{2}\bar{s} - s^I)(x_1 - x_2) > 0$ for all $x_1 \neq x_2]$. This firm withdraws more electricity in the peak demand period than an integrated DSO/aggregator $[(s^I - s^{DSO})(x_1 - x_2) \geq 0]$.*

[18] The cost function in the previous footnote satisfies this assumption. The left-hand side of (9) can then be written as $\sigma\frac{y-s_i}{q_i}$. This expression is below 1 by $\sigma \leq 1$, and since $q_i = x_i + s_i - y > y - s_i$ by $y < \frac{1}{2}x_i$.

Distortions arising from the exercise of market power in the supply of the energy service $\bar{s}$ are exacerbated if the service is provided by a firm that also supplies thermal electricity to the balancing market, compared to the case of a single aggregator that does not own any generation capacity. Independent aggregators have excessive incentives to allocate electricity consumption to the peak demand period. The integrated generator/aggregator has an even stronger incentive to allocate electricity consumption to the peak demand period because it can then increase the income on the generation capacity supplied to the balancing market.

3.5 Comparison of market structures

The efficiency with which flexible distribution resources are used for balancing electricity supply within the local distribution area depends fundamentally on the structure of the market:

Proposition 6 *The different structures of the market for flexible distribution resources can be ranked as follows in decreasing order of cost efficiency:*
(i) $A \geq 2$ *independent aggregators.*
(ii) Integrated DSO/aggregator.
(iii) Integrated generator/aggregator.
(iv) No FDR market (passive consumers).
(v) Mixed market structure (and L^{DSO} *sufficiently small).*

Proof: By way of propositions 2, 3, 4 and 5, it follows that $s^{fb} \leq s^A \leq s^{DSO} \leq s^I < \frac{1}{2}\bar{s} \leq s^{DSO} + s^A$ if $x_1 > x_2$ and $s^{DSO} + s^A \leq \frac{1}{2}\bar{s} < s^I \leq s^{DSO} \leq s^A \leq s^{fb}$ if $x_2 > x_1$. This ranking of the equilibrium consumption under the different market structures, and the strict convexity of the total system cost $C(x_1 + s) + C(x_2 + \bar{s} - s)$, deliver the result.■

The extent to which the power market is susceptible to firms' exploitation of market power, varies with the structure of the market for flexible distribution resources. The most efficient market structure is the one in which multiple independent aggregators compete for flexible distribution resources. Competition in the downstream FDR market reduces market power in the upstream real-time market. The equilibrium converges to the efficient allocation in the limit when the number of aggregators becomes large. A firm that is active on both sides of the balancing market by producing electricity as well as consuming it through an aggregator, can exercise market power by manipulating consumption even if this firm is unable to exploit market power by withholding production.[19] Compared to the case with independent aggregators, this firm will allocate excessive consumption to the peak demand period in order to increase profit on its generation capacity. Contrary to what one might expect, vertical integration between the supply side (generation) and demand side (aggregation) is less efficient than independent

[19] Regulatory scrutiny may be one reason why a generation owner would restrict its exercise of market power. The analysis in this paper shows that the generation owner can circumvent this problem through vertical integration. It can then distort competition through strategic bidding on the demand side. Detecting exercise of buyer market power is more challenging than seller market power insofar as measuring the marginal value of consumption is more difficult than assessing the marginal cost of production.

operation of generation and aggregation. However, this inefficiency is never so pervasive as to render the FDR market inefficient compared to not having any market at all.

A central policy question is whether the DSO itself should be involved in the market for flexible distribution resources. I consider first the case where a parent company owns the DSO and a monopoly aggregator. Because of the substantial market power of this ownership constellation, the market structure is less efficient than one with multiple independent aggregators. Still, the monopoly aggregator withdraws more electricity from the network in the off-peak period to supply the household energy service than what is the case in the benchmark without any FDR market and passive consumers. Notwithstanding market power issues, a monopoly DSO/aggregator therefore is better than not utilizing the flexible distribution resources at all. This conclusion may be overturned if there are also independent aggregators in the market. In the polar extreme case where the DSO has full discretion over electricity dispatch, the DSO might allocate so much consumption to the peak-period that it would in fact be better not to have any FDR market at all.

The risk of creating a power shortage in the peak demand period would soften the DSO incentive to deliberately aggravate resource constraints, but the analysis nevertheless demonstrates important concerns that may arise under insufficient unbundling of DSO and aggregator activities. Preventing the DSO from exercising market power associated with DSO dispatch of FDRs is challenging because identifying efficient dispatch requires precise information about all relevant aspects of historical operating conditions. A multitude of unit-specific properties, such as ramping constraints and location in the network that determine the efficient dispatch of resources, contribute to the difficulty of the problem. More generally, the cost of DSO presence in the FDR market is likely to exceed the benefits if multiple independent aggregators are already present in the market.

3.6 Efficient regulation

Competition for flexible distribution resources generally leads to inefficient market allocations if this market is imperfectly competitive. An option then is to introduce some type of regulation if one cannot establish well-functioning competition, for example because of entry barriers.

To illustrate the type of regulatory scheme that will implement efficient allocations in the setting with competitive generation, consider an independent aggregator that holds a monopoly position in the market for flexible distribution resources. Assume that this aggregator sells the energy service to the representative household at a fixed fee t^R. The maximal fee the aggregator can charge from the consumer equals $\frac{1}{2}(p_1^{fb} + p_2^{fb})\bar{s}$ if the outside option of the household is to purchase the electricity needed for production of the energy service at expected cost and the household expects efficient real-time prices. If so, the total profit of the aggregator is

$$\Pi^R(s) = \frac{1}{2}(p_1^{fb} + p_2^{fb})\bar{s} - P(x_1 + s)s - P(x_2 + \bar{s} - s)(\bar{s} - s) + R(s)$$

if it withdraws s from the grid in the first period. The final term in the above profit expression is a compensation payment defined in (10) below, levied on consumers in a lump-sum fashion.

The marginal incentive to increase period 1 electricity consumption equals

$$\Pi^{R\prime}(s) = P(x_2 + \bar{s} - s) - P(x_1 + s)$$

by the construction of $R(s)$. The marginal payment $R'(s)$ neutralizes the aggregator's market power in the flexibility market, so that the optimal choice by the aggregator is to consume $s = s^{fb}$ in the first period. Moreover, the compensation payment is designed in such a way that the actual payment is zero (balanced-budget) at the efficient allocation, i.e. $R(s^{fb}) = 0$. The aggregator's profit in equilibrium equals $\Pi^R(s^{fb}) = (p_1^{fb} - p_2^{fb})(\frac{1}{2}\bar{s} - s^{fb})$, which is non-negative; see Corollary 1. I state the following immediate result without further proof:

Proposition 7 *Assume that flexible generation supplied at marginal cost. The compensation function*

$$R(s) = \int_s^{s^{fb}} [P'(x_2 + \bar{s} - \tilde{s})(\bar{s} - \tilde{s}) - P'(x_1 + \tilde{s})\tilde{s}]d\tilde{s} \gtreqless 0 \tag{10}$$

incites an independent aggregator with a monopoly in the market for flexible distribution resources to withdraw the efficient amount s^{fb} of electricity from the grid. The compensation payment is zero in equilibrium, $R(s^{fb}) = 0$, and satisfies the aggregator's participation constraint, $\Pi^R(s^{fb}) \geq 0$.

Figure 2 essentially reproduces Figure 1 to illustrate how the compensation payment $R(s)$ operates in this framework.

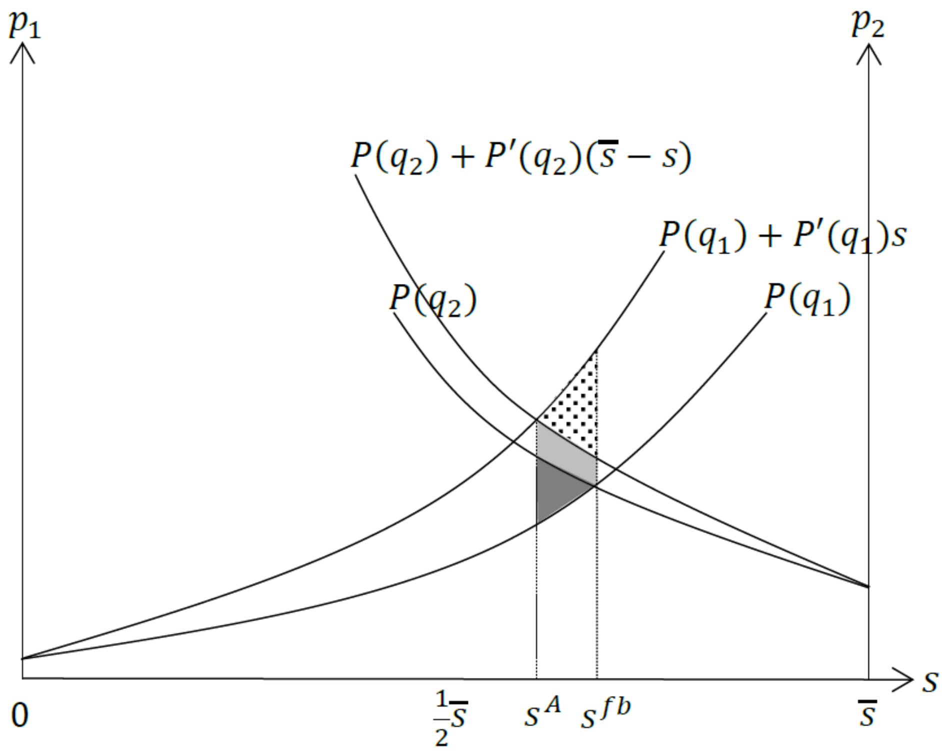

Figure 2: Regulation

As in Figure 1, period 2 is the peak demand period by construction, so that an unregulated aggregator in equilibrium withdraws relatively more electricity from the grid in period 1 than in period 2, $s^A > \frac{1}{2}\bar{s}$. Yet, exploitation of market power implies that electricity consumption is downward distorted in the off-peak period, $s^A < s^{fb}$. Increasing electricity withdrawal in the first period from s^A to the efficient level s^{fb} increases the aggregator's total expenditures in the real-time market by the dotted area in Figure 2 and thus is unprofitable without any compensation. An increase in electricity consumption in period 1 from s^A to s^{fb} under $R(s)$ increases the first period compensation by an area equal to the dotted plus the medium grey and the dark grey area in Figure 2. Period 2 compensation falls by the medium grey area. By summing up all incremental effects, it follows that an increase from s^A to s^{fb} increases the aggregator's profit by the dark grey area in the figure. This is equal to the total efficiency gain of increasing electricity consumption in the first period up to the efficient level from the profit maximizing level. Turning this argument around, the construction of $R(s)$ implies that the monopoly aggregator is residual party to all efficiency losses associated with a deviation from s^{fb}. The regulation is thus an application of the "polluter pays" principle to the context of imperfect competition.

Implementation of $R(s)$ requires that the supply functions of all participants in the flexibility market are observable by the regulatory authority. The regulator can often collect the necessary data to compute these functions from the DSO operating the real-time market. Implementing $R(s)$ can be more complicated in other settings. A relevant example is when a subsidiary aggregator of a parent company that owns the DSO participates in the market for flexible distribution resources. In this case, the DSO might have weak incentives to truthfully disclose supply functions because doing so might enable the regulator to extract surplus from the subsidiary.

4 Market power in generation

The analysis has been conducted under the assumption that generators bid all their flexible production into the local real-time (flexibility) market at marginal cost. This section extends the model to include imperfect competition in the production of electricity. The detailed analysis in contained in an online appendix (Tangerås, 2022).

The model One large generation owner strategically bids production $y_1 \geq 0$ into the market in period 1 and $y_2 \geq 0$ in period 2. The continuous cost function $\Psi(y)$ of this producer is the same in both periods, increasing and weakly convex. The residual demand $q_i = x_i + s_i - y_i$ is supplied at marginal cost by a competitive fringe in both periods $i = 1, 2$. Hence, the price for electricity in period i is $p_i = P(q_i) = C'(q_i)$.

The large generator maximizes profit $P(x_i + s_i - y_i)y_i - \Psi(y_i)$, $i = 1, 2$. Optimizing over y_i returns the production $Y(x_i + s_i) > 0$ of the large firm in period i as a solution to the first-order condition

$$P(Q(x_i + s_i)) - P'(Q(x_i + s_i))Y(x_i + s_i) = \Psi'(Y(x_i + s_i)).$$

In this expression, $Q(x_i+s_i) = x_i+s_i-Y(x_i+s_i)$ measures the supply of the competitive fringe in period i as a function of total demand $x_i + s_i$. Production by the large generation owner and the competitive fringe are both increasing in total demand by appropriate assumptions on $C(q)$.

System costs The total system cost equals

$$Sys(s) = \Psi(Y(x_1 + s)) + C(Q(x_1 + s)) + \Psi(Y(x_2 + \bar{s} - s)) + C(Q(x_2 + \bar{s} - s))$$

as a function of the amount s of electricity withdrawn from the grid in the first period to supply the household energy service. Assume for the sake of the argument that total demand in period 1 is smaller than total demand in period 2, so that $x_1 + s < x_2 + \bar{s} - s$. Production levels $Y_1 = Y(x_1 + s)$ and $Q_1 = Q(x_1 + s)$ in period 1 then are smaller than $Y_2 = Y(x_2 + \bar{s} - s)$ and $Q_2 = Q(x_2 + \bar{s} - s)$ in period 2. A marginal increase in electricity consumption s then tends to reduce the total system cost by allocating more consumption to the low cost period. However, such a reallocation of demand also affects the production of the large generator, such that the marginal effect on system cost becomes

$$Sys'(s) = \Psi'(Y_1)Y_1' + C'(Q_1)Q_1' - \Psi'(Y_2)Y_2' - C'(Q_2)Q_2',$$

where $Y_i' = Y'(x_i + s_i)$ and $Q_i' = Q(x_i + s_i)$. In principle, the marginal effects Q_1' and Y_2' could be close to one, and $C'(Q_1) = P(Q_1) > \Psi'(Y_2)$, so that the total system cost increases with an increase in s, despite all generators producing less in period 1 than 2. Appropriate regularity assumptions on the elasticity $C'''(q)q/C''(q)$ of the cost function ensure that the direct effect dominates; see Tangerås (2022). From a system perspective, the most efficient way to allocate consumption s is then to equalize total demand across periods if possible. Hence, price equalization is efficient even if market power in generation is an issue. If such price equalization is not possible, then all electricity consumed to produce the energy service should be withdrawn in the off-peak demand period. Hence, the efficient consumption satisfies $s = s^{fb}$ even under generator market power.

Strategic behavior by aggregators I show in Tangerås (2022), that market power in generation does not matter for the incentives for an integrated DSO/aggregator or A independent aggregators to allocate consumption across periods. They consume relatively more of their electricity in the off-peak demand period, although too little from an efficiency viewpoint. The inefficiency diminishes when there are more aggregators. The results of sections 3.1 and 3.2 still hold. Even the results of Section 3.3 are robust to market power in generation. Under centralized DSO dispatch, an integrated DSO/aggregator allocates most of its own consumption to the off-peak demand period and all the electricity consumption by the independent aggregators to the peak demand period. An additional distortion may arise under this market structure because the DSO has an incentive to dispatch the capacity of the large generator in the off-peak demand period. This inefficiency only reinforces the result in Proposition 6 that passive consumption (no

FDR market) may be preferable from an efficiency view point compared to centralized dispatch by an integrated DSO/aggregator when other aggregators also participate in the market. The other market structures are more efficient than not having any FDR market.

Interesting qualitative differences arise under the market structure in Section 3.4 when the large generator with market power also has a monopoly in the market for flexible distribution resources. This integrated generator/aggregator maximizes profit

$$P(x_1 + s - y_1)(y_1 - s) + P(x_2 + \bar{s} - s - y_2)(y_2 - \bar{s} + s) - \Psi(y_1) - \Psi(y_2)$$

over (y_1, y_2, s). The profit-maximizing production $Y^I(x_i, s_i)$ of this firm in period i solves

$$P(Q^I(x_i, s_i)) - P'(Q^I(x_i, s_i))(Y^I(x_i, s_i) - s_i) = \Psi'(Y^I(x_i, s_i)). \tag{11}$$

In this expression, $Q^I(x_i, s_i) = x_i + s_i - Y^I(x_i, s_i)$ measures the production of the competitive fringe. For given demand, an integrated generator/aggregator has a stronger incentive to supply generation to the market than an independent generator, $Y^I(x_i, s_i) \geq Y(x_i + s_i)$ with strict inequality if $s_i > 0$, because the associated reduction in the price of electricity reduces the aggregator's cost of electricity. This mechanism is formally equivalent to the well-known pro-competitive effect of producers selling forward contracts (e.g. Wolak, 2007) or when there is vertical integration between generation and retail (e.g. Bushnell et al., 2008).

The firm allocates s to equate marginal net expenditures across periods,

$$P(q_1^I) - P'(q_1^I)(y_1^I - s^I) = P(q_2^I) - P'(q_2^I)(y_1^I - \bar{s} + s^I)$$

in interior optimum $s^I \in (0, \bar{s})$. In the above expression, $q_i^I = Q^I(x_i, s_i^I)$ and $y_i^I = Y^I(x_i, s_i^I)$, where $s_1^I = s^I$ and $s_2^I = \bar{s} - s^I$. Using the first-order condition (11) for optimal production, I can rewrite this equilibrium condition simply as $\Psi'(y_1^I) = \Psi'(y_2^I)$. The integrated generator/aggregator uses production y_1 and y_2 to exercise market power and then allocates s across periods to minimize production cost.

The integrated generator/aggregator allocates most of its consumption to the off-peak demand period to minimize production costs, but still consumes too much electricity in the peak demand period. Fundamental results of Section 3.4 therefore continue to hold. The comparison with other market structures is less straightforward because the integrated generator/aggregator utilizes consumption for different purposes than other aggregators. Under robust conditions, however, $(s^A - s^I)(x_1 - x_2) \geq 0$ for $A = 1$. Generator and aggregator integration has a double efficiency benefit in this case. First, vertical integration increases production efficiency by replacing high-cost generation of the competitive fringe with low cost production of the generator with market power. Second, the household energy service is produced more efficiently because the aggregator allocates relatively more consumption to the off-peak demand period. The sum of these two effects jointly turns the efficiency ranking (*ii*) versus (*iii*) in Proposition 6 around.

I conclude that results do not depend fundamentally on the assumption of competitive supply of generation capacity. Most of them go through also under the alternative assumption that one large generation owner behaves strategically by withholding production from the real-time market. A main difference is that vertical integration between generation and aggregation can be more efficient than a market structure in which generators and aggregators behave independently. In general, vertical integration between generation and aggregation can be more or less efficient than independent management of flexible distribution resources depending on the competitiveness of the real-time market.

5 Conclusion

This paper has built a simple two-period model of an electricity market in which smart grid solutions enable third parties to supply flexible distribution resources (FDRs), in the form of intertemporal substitution of electricity consumption, for the purpose of facilitating local balancing of production and consumption. Examples of such consumption include indoor temperature control (heating or cooling) or battery charge and discharge. The main purpose was to examine how assumptions about the structure of the evolving market for FDRs affect the efficiency with which these resources are deployed in the balancing market. Despite the simplicity of the model, it has generated results and policy implications that are likely to be robust and carry over to more general settings.

Control over substantial FDR capacity enables whoever controls this resource to wield market power in the balancing market, which causes excessive price differences between peak and off-peak periods. The distortions are smaller if multiple independent firms, *aggregators*, compete for FDRs. I show how an aggregator that also supplies generation to the balancing market deploys resources less efficiently than an independent aggregator. This inefficiency is reversed if the integrated firm also exercises market power on the production side. I characterize a regulatory policy that can implement the efficient allocation on the basis of data that can be obtained from the balancing market.

Reaping the competitive benefits of an FDR market requires a formal local balancing market. Such a market is most often run by the distribution system operator (DSO) that owns the local network. A key policy issue concerns the interaction between system operation and the market for FDR resources. I show that the DSO has an incentive to dispatch electricity inefficiently if the DSO itself or a parent company holds a controlling stake in an aggregator. Efficiency concerns then suggest that system operation should be completely separated from the FDR market, even if direct foreclosure might not be a problem.

A market design featuring a formal local balancing market is not necessarily economically viable, for instance if running such a market features scale returns that cannot be achieved in a small control area. Absent a formal market, perhaps the only option is to delegate the supply of flexible distribution resources to the DSO itself or a subsidiary. A DSO/aggregator monopoly structure is not an argument against employing flexible distribution resources because efficiency is still higher than in the benchmark case where that potential remains unused.

References

Brown, David P. and David E.M. Sappington (2018): Optimal Procurement of Distributed Energy Resources. *The Energy Journal* 39(5), 131–155.

Brown, David P. and David E.M. Sappington (2019): Employing Cost Sharing to Motivate the Efficient Implementation of Distributed Energy Resources. *Energy Economics* 81(June), 974–1001.

Burger, Scott, Jose Pablo Chaves-Ávila, Carlos Batlle and Ignacio J. Pérez-Arriaga (2017): A Review of the Value of Aggregators in Electricity Systems. *Renewable and Sustainable Energy Reviews* 77(September), 395–405.

Burger, Scott P., Jesse D. Jenkins, Carlos Batlle and Ignacio J. Pérez-Arriaga (2019a): Restructuring Revisited Part 1: Competition in Electricity Distribution Systems. *The Energy Journal* 40(3), 31–54.

Burger, Scott P., Jesse D. Jenkins, Carlos Batlle and Ignacio J. Pérez-Arriaga (2019b): Restructuring Revisited Part 2: Coordination in Electricity Distribution Systems. *The Energy Journal* 40(3), 55–76.

Bushnell, James B., Erin T. Mansur and Celeste Saravia (2008): Vertical Arrangements, Market Structure, and Competition: An Analysis of Restructured US Electricity Markets. *American Economic Review* 98(1), 237–66.

CAISO (2021): Final Root Cause Analysis: Mid-August 2020 Extreme Heat Wave. January 13.

EDSO (2015): Coordination of Transmission and Distribution System Operators: A Key Step for the Energy Union, https://www.edsoforsmartgrids.eu/wp-content/uploads/public/Coordination-of-transmission-and-distribution-system-operators-May-2015.pdf, October 15, 2019.

EU (2019): Directive 2019/944 on common rules for the internal market for electricity. *Official Journal of the European Union* L158, 125–195.

FERC (2020): Participation of Distributed Energy Resource Aggregations in Markets Operated by Regional Transmission Organizations and Independent System Operators. FERC Order No. 2222, September 17.

Ito, Koichiro and Mar Reguant (2016): Sequential Markets, Market Power, and Arbitrage. *American Economic Review* 106(7), 1921–1957.

Kim, Seung Wan, Michael G. Pollitt, Young Gyu Jin, Jip Kim and Yong Tae Yoon (2017): Contractual Framework for the Devolution of System Balancing Responsibility from the Transmission System Operator to Distribution System Operators. University of Cambridge, EPRG Working Paper 1715.

Pollitt, Michael G. and Karim L. Anaya (2019): Competition in Markets for Ancillary Services? The Implications of Rising Distributed Generation. *The Energy Journal* 41(SI1), 5–31.

Rey, Patrick and Jean Tirole (2007): A Primer on Foreclosure. In Mark Armstrong and Robert Porter (eds): *Handbook of Industrial Organization Volume 3.* Amsterdam, The Netherlands: North-Holland, 2145–2220.

Rintamäki, Tuomas, Afzal S. Siddiqui and Ahti Saloa (2020): Strategic Offering of a Flexible Producer in Day-ahead and Intraday Power Markets. *European Journal of Operational Research* 284, 1136–1153.

Tangerås, Thomas (2022): Competition for Flexible Distribution Resources in a 'Smart' Electricity Distribution Network: Online Appendix. Available at www.ifn.se/thomast.

Tangerås, Thomas and Johannes Mauritzen (2018): Real-time versus Day-ahead Market Power in a Hydro-based Electricity Market. *Journal of Industrial Economics* 66(4), 904–941.

Wolak, Frank A. (2007): Quantifying the Supply-side Benefits from Forward Contracting in Wholesale Electricity Markets. *Journal of Applied Econometrics* 22(7), 1179–1209.

Appendix

A.1 Proof of Lemma 1

The Lagrangian $-C(x_1 + s) - C(x_2 + \bar{s} - s) + \underline{\lambda}s + \bar{\lambda}(\bar{s} - s)$ of the central planner's problem is strictly concave by strict convexity $C(q)$, where $\underline{\lambda}$ is the Kuhn-Tucker (KT) multiplier associated with $s \geq 0$, and $\bar{\lambda}$ is the KT multiplier associated with $s \leq \bar{s}$. The first-order condition

$$-C'(x_1 + s^{fb}) + C'(x_2 + \bar{s} - s^{fb}) + \underline{\lambda}^{fb} - \bar{\lambda}^{fb} = 0 \tag{12}$$

and complementary slackness conditions

$$s^{fb} \in [0, \bar{s}],\ \underline{\lambda}^{fb} \geq 0,\ \bar{\lambda}^{fb} \geq 0,\ \underline{\lambda}^{fb} s^{fb} = \bar{\lambda}^{fb}(\bar{s} - s^{fb}) = 0 \tag{13}$$

are therefore necessary and sufficient to characterize the efficient consumption s^{fb} and the efficient KT multipliers $(\underline{\lambda}^{fb}, \bar{\lambda}^{fb})$. It is straightforward to verify that the following are solutions to (12) and (13): If $|x_1 - x_2| \leq \bar{s}$, then $s^{fb} = s^*$ and $\underline{\lambda}^{fb} = \bar{\lambda}^{fb} = 0$; if $x_1 - x_2 > \bar{s}$, then $s^{fb} = 0$, $\underline{\lambda}^{fb} = C'(x_1) - C'(x_2 + \bar{s}) > 0$ and $\bar{\lambda}^{fb} = 0$; if $x_2 - x_1 > \bar{s}$, then $s^{fb} = \bar{s}$, $\underline{\lambda}^{fb} = 0$ and $\bar{\lambda}^{fb} = C'(x_2) - C'(x_1 + \bar{s}) > 0$.

Let $q_1^{fb} = x_1 + s^{fb}$ and $q_2^{fb} = x_2 + \bar{s} - s^{fb}$. If $|x_1 - x_2| \leq \bar{s}$, then $q_1^{fb} = q_2^{fb} = q^*$. If $x_1 - x_2 > \bar{s}$, then $q_1^{fb} - q_2^{fb} = x_1 - x_2 - \bar{s} > 0$ and therefore $(q_1^{fb} - q_2^{fb})(x_1 - x_2) > 0$. If $x_2 - x_1 > \bar{s}$, then $q_1^{fb} - q_2^{fb} = x_1 - x_2 + \bar{s} < 0$ and again $(q_1^{fb} - q_2^{fb})(x_1 - x_2) > 0$.■

A.2 Proof of Proposition 2

The Lagrangian $\Pi^{DSO}(s) + \underline{\lambda}s + \bar{\lambda}(\bar{s} - s)$ of the DSO is strictly concave by the assumption of competitive supply of thermal electricity in the balancing market and the properties of $C'(\cdot)$. Hence, the first-order condition

$$p_2^{DSO} - p_1^{DSO} - P'(x_1 + s^{DSO})s^{DSO} + P'(x_2 + \bar{s} - s^{DSO})(\bar{s} - s^{DSO}) + \underline{\lambda}^{DSO} - \bar{\lambda}^{DSO} = 0 \tag{14}$$

and complementary slackness conditions

$$s^{DSO} \in [0, \bar{s}],\ \underline{\lambda}^{DSO} \geq 0,\ \bar{\lambda}^{DSO} \geq 0,\ \underline{\lambda}^{DSO} s^{DSO} = \bar{\lambda}^{DSO}(\bar{s} - s^{DSO}) = 0 \tag{15}$$

are necessary and sufficient optimality conditions for s^{DSO} and the equilibrium KT multipliers $(\underline{\lambda}^{DSO}, \bar{\lambda}^{DSO})$. In the first-order condition (14), $p_1^{DSO} = P(x_1 + s^{DSO})$ and $p_2^{DSO} = P(x_2 + \bar{s} - s^{DSO})$ denote the equilibrium real-time prices of electricity.

Part (i) of the proposition: If $x_1 - x_2 > \bar{s}$, then $(s^{DSO} - s^{fb})(x_1 - x_2) = s^{DSO}(x_1 - x_2) \geq 0$ because $s^{fb} = 0$ by Lemma 1. By that same Lemma, $(s^{DSO} - s^{fb})(x_1 - x_2) = (\bar{s} - s^{DSO})(x_2 - x_1) \geq 0$ for $x_2 - x_1 > \bar{s}$ because then $s^{fb} = \bar{s}$. If $x_1 = x_2$, then $(s^{DSO} - s^{fb})(x_1 - x_2) \geq 0$ trivially holds. The final case is where $|x_1 - x_2| \in (0, \bar{s}]$. By full price equalization for $s^{fb} = s^* \in [0, \bar{s}]$:

$$\begin{aligned} \Pi^{DSO\prime}(s^{fb}) &= -P(x_1 + s^{fb}) - P'(x_1 + s^{fb})s^{fb} + P(x_2 + \bar{s} - s^{fb}) + P'(x_2 + \bar{s} - s^{fb})(\bar{s} - s^{fb}) \\ &= P'(q^*)(\bar{s} - 2s^*) = P'(q^*)(x_1 - x_2). \end{aligned}$$

As $P' > 0$, it follows that $\Pi^{DSO\prime}(s^{fb}) > 0$ for $x_2 + \bar{s} \geq x_1 > x_2$, in which case $s^{DSO} > s^{fb}$ by strict concavity of $\Pi^{DSO}(s)$ and because $s^{fb} < \bar{s}$ in this range of (x_1, x_2). In an analogous manner, $x_1 + \bar{s} \geq x_2 > x_1$ implies $s^{DSO} < s^{fb}$. Hence, $(s^{DSO} - s^{fb})(x_1 - x_2) > 0$ for $|x_1 - x_2| \in (0, \bar{s}]$.

Part (ii) of the proposition: Rearrange the first-order condition (14) and multiply through by $(p_1^{DSO} - p_2^{DSO})$ to get

$$\begin{aligned} &[2P'(q_1^{DSO})(\tfrac{1}{2}\bar{s} - s^{DSO}) + \underline{\lambda}^{DSO} - \bar{\lambda}^{DSO}](p_1^{DSO} - p_2^{DSO}) \\ &\quad = [P'(q_1^{DSO}) - P'(q_2^{DSO})](p_1^{DSO} - p_2^{DSO})(\bar{s} - s^{DSO}) + (p_1^{DSO} - p_2^{DSO})^2. \end{aligned} \tag{16}$$

In this expression $q_1^{DSO} = x_1 + s^{DSO}$ and $q_2^{DSO} = x_2 + \bar{s} - s^{DSO}$ measure equilibrium production in each period. The first term on the right-hand side (RHS) of (16) is non-negative because convexity of the marginal production cost $C'(q)$ implies

$$[P'(x_1 + s) - P'(x_2 + \bar{s} - s)][P(x_1 + s) - P(x_2 + \bar{s} - s)] \geq 0 \text{ for all } s \in [0, \bar{s}], \tag{17}$$

The second term on the right-hand side (RHS) of (16) is strictly positive for all $p_1^{DSO} \neq p_2^{DSO}$. The left-hand side (LHS) of (16) is zero if $s^{DSO} = \frac{1}{2}\bar{s}$ by $\underline{\lambda}^{DSO} s^{DSO} = \bar{\lambda}^{DSO}(\bar{s} - s^{DSO}) = 0$. Hence, $p_1^{DSO} \neq p_2^{DSO}$ implies $s^{DSO} \neq \frac{1}{2}\bar{s}$. If $p_1^{DSO} > p_2^{DSO}$, then LHS of (16) is strictly negative if $s^{DSO} > \frac{1}{2}\bar{s}$ by $P'(q_1^{DSO}) > 0$, $\bar{\lambda}^{DSO} \geq 0$ and $\underline{\lambda}^{DSO} = 0$. Hence, $p_1^{DSO} > p_2^{DSO}$ implies $s^{DSO} < \frac{1}{2}\bar{s}$. Similarly, $p_2^{DSO} > p_1^{DSO}$ implies $s^{DSO} > \frac{1}{2}\bar{s}$. Next, let $x_1 > x_2$ and suppose $s^{DSO} \geq \frac{1}{2}\bar{s}$. In this case,

$$q_1^{DSO} - q_2^{DSO} = x_1 - x_2 + 2s^{DSO} - \bar{s} > 0,$$

and therefore $p_1^{DSO} = P(q_1^{DSO}) > P(q_2^{DSO}) = p_2^{DSO}$ by $P' > 0$. But then $s^{DSO} < \frac{1}{2}\bar{s}$ from the previous argument, which is a contradiction. Hence, $x_1 > x_2$ implies $s^{DSO} < \frac{1}{2}\bar{s}$. By a similar argument, $x_2 > x_1$ implies $s^{DSO} > \frac{1}{2}\bar{s}$. This concludes the proof that $(\frac{1}{2}\bar{s} - s^{DSO})(x_1 - x_2) > 0$

for all $x_1 \neq x_2$.■

A.3 Proof of Corollary 2

Rearrange the first-order condition (14) as

$$2P'(q_1^{DSO})(\tfrac{1}{2}\bar{s}-s^{DSO})+\underline{\lambda}^{DSO}-\bar{\lambda}^{DSO} = P(q_1^{DSO})-P(q_2^{DSO})+[P'(q_1^{DSO})-P'(q_2^{DSO})](\bar{s}-s^{DSO}). \tag{18}$$

Assume that $x_1 > x_2$. Proposition 2 has demonstrated that $s^{DSO} < \frac{1}{2}\bar{s}$ in this case. By implication, LHS of (18) is strictly positive since $P'(q_1^{DSO}) > 0$, $\underline{\lambda}^{DSO} \geq 0$ and $\bar{\lambda}^{DSO}(\bar{s}-s^{DSO}) = 0$. RHS of (18) is non-positive if $q_1^{DSO} \leq q_2^{DSO}$ because then $P(q_1^{DSO}) \leq P(q_2^{DSO})$ by $P' > 0$ and $P'(q_1^{DSO}) \leq P'(q_2^{DSO})$ by (17). Hence, $x_1 > x_2$ implies $q_1^{DSO} > q_2^{DSO}$, and therefore $p_1^{DSO} = P(q_1^{DSO}) > P(q_2^{DSO}) = p_2^{DSO}$ by $P' > 0$. By analogous arguments, $x_2 > x_1$ implies $p_2^{DSO} > p_1^{DSO}$.■

A.4 Proof of Proposition 3

The Lagrangian $\Pi^A(s_a) + \underline{\lambda}_a s_a + \bar{\lambda}_a(L_a\bar{s} - s_a)$ of aggregator a is strictly concave by the assumptions that thermal capacity is competitively supplied and the properties of $C'(\cdot)$. Hence, the first order-condition

$$p_2^A - p_1^A - \tfrac{1}{A}P'(q_1^A)s^A + \tfrac{1}{A}P'(q_2^A)(\bar{s} - s^A) + \underline{\lambda}^A - \bar{\lambda}^A = 0 \tag{19}$$

and complementary slackness conditions

$$s^A \in [0, \bar{s}],\ \underline{\lambda}^A \geq 0,\ \bar{\lambda}^A \geq 0,\ \underline{\lambda}^A s^A = \bar{\lambda}^A(\bar{s} - s^A) = 0 \tag{20}$$

characterize the unique symmetric second-stage equilibrium $s_a^A = s^A/A$, $L_a^A = 1/A$, $\bar{\lambda}_a^A = \bar{\lambda}^A$, $\underline{\lambda}_a^A = \underline{\lambda}^A$ for all a, and where $q_1^A = x_1 + s^A$, $q_2^A = x_1 + \bar{s} - s^A$ and $p_i^A = P(q_i^A)$, $i = 1, 2$.

Part (i) of the proposition: I omit the proof that $(s^A - s^{fb})(x_1 - x_2) \geq 0$ with strict inequality if $|x_1 - x_2| \in (0, \bar{s}]$ because it is identical to the proof of the first part of Proposition 2. As for the comparative statics result, differentiate the equilibrium condition $\Pi^{A\prime}(\frac{1}{A}s^A) = 0$ for $s^A \in (0, \bar{s})$:

$$\frac{ds^A}{dA}(x_1 - x_2) = \frac{-(p_1^A - p_2^A)(x_1 - x_2)}{(A+1)(C''(q_1^A) + C''(q_2^A)) + C'''(q_1^A)s^A + C'''(q_2^A)(\bar{s} - s^A)}.$$

The denominator is strictly positive. I can then follow the same steps as in the proof of Corollary 2 to establish $(p_1^A - p_2^A)(x_1 - x_2) > 0$ for all $x_1 \neq x_2$.

Part (ii) of the proposition: If $x_1 > x_2$, then Lemma 1 implies $q_1^{fb} \geq q_2^{fb}$ and Part (i) of this proposition implies $s^A \geq s^{fb}$. Assume that $s^A > s^{fb} \geq 0$. I can then write (19) as

$$\tfrac{1}{A}[P'(q_2^A)(\bar{s} - s^A) - P'(q_1^A)s^A] = P(q_1^{fb} + s^A - s^{fb}) - P(q_2^{fb} + s^{fb} - s^A) + \bar{\lambda}^A > 0.$$

The term in square brackets on LHS of this expression is bounded, so LHS converges to zero when A goes to infinity. If $\lim_{A\to\infty} s^A > s^{fb}$, then RHS of this expression is strictly positive in the limit, which is a contradiction. By a similar argument $\lim_{A\to\infty} s^A < s^{fb}$ is a contradiction if $x_1 < x_2$. It is straightforward to verify that $s^A = s^* = s^{fb}$ if $x_1 = x_2$.

Part (iii) of the proposition, necessity: Obviously, $s^A = s^{DSO}$, $\underline{\lambda}^A = \underline{\lambda}^{DSO}$ and $\bar{\lambda}^A = \bar{\lambda}^{DSO}$ satisfy (19) and (20) for $A = 1$. The necessity of $x_1 \neq x_2$ is trivial. For $s^{DSO} = \bar{s}$, it is straightforward to verify that $s^A = \bar{s}$, $\underline{\lambda}^A = 0$ and

$$\bar{\lambda}^A = \tfrac{A-1}{A}(P(x_2) - P(x_1 + \bar{s})) + \tfrac{1}{A}\bar{\lambda}^{DSO} \geq 0$$

solve (19) and (20). To see why $\bar{\lambda}^A \geq 0$ in this case, recall that $(p_1^{DSO} - p_2^{DSO})(\frac{1}{2}\bar{s} - s^{DSO}) \geq 0$ from the proof of Proposition 2. Hence, $s^{DSO} = \bar{s}$ implies $p_2^{DSO} - p_1^{DSO} = P(x_2) - P(x_1 + \bar{s}) \geq 0$. Similarly, $s^{DSO} = 0$ implies that $s^A = 0$, $\bar{\lambda}^A = 0$ and

$$\underline{\lambda}^A = \tfrac{A-1}{A}(P(x_1) - P(x_2 + \bar{s})) + \tfrac{1}{A}\underline{\lambda}^{DSO} \geq 0$$

solve (19) and (20).

Part (iii) of the proposition, sufficiency: Assume that $A > 1$, $x_1 \neq x_2$ and $s^{DSO} \in (0, \bar{s})$. Evaluated at $s_{a'} = \frac{1}{A}s^{DSO}$ for all a', the marginal profit of aggregator a simplifies to $\Pi^{A\prime}(\frac{1}{A}s^{DSO})(x_1 - x_2) = -\frac{A-1}{A}(p_1^{DSO} - p_2^{DSO})(x_1 - x_2) < 0$. The strict negativity follows from $(p_1^{DSO} - p_2^{DSO})(x_1 - x_2) > 0$ for all $x_1 \neq x_2$, see the proof of Corollary 2. It follows that $s^A < s^{DSO}$ if $x_1 > x_2$, whereas $s^A > s^{DSO}$ if $x_1 < x_2$.■

A.5 Proof of Proposition 4

Part (i) of the proposition: Maximizing the objective function

$$\Omega(\tilde{s}^{DSO}, \tilde{s}^A, \mathbf{x}) = -P(x_1 + \tilde{s}^{DSO} + \tilde{s}^A)\tilde{s}^{DSO} - P(x_2 + \bar{s} - \tilde{s}^{DSO} - \tilde{s}^A)(\bar{s}^{DSO} - \tilde{s}^{DSO}) \quad (21)$$

over $(\tilde{s}^{DSO}, \tilde{s}^A) \in [0, \bar{s}^{DSO}] \times [0, \bar{s}^A]$ is complicated by non-concavity. Although (21) is strictly concave in $\tilde{s}^{DSO}$ holding $\tilde{s}^A$ constant, and it is concave in $\tilde{s}^A$ holding $\tilde{s}^{DSO}$ constant, its Hessian matrix has determinant $-[C''(q_1) + C''(q_2)]^2 < 0$. This means that all interior solutions to the first-order conditions are saddle points. By implication, the profit-maximizing electricity withdrawal (s^{DSO}, s^A) features corner solutions. Assume for the rest of the proof that $x_1 > x_2$. The $x_1 < x_2$ case is symmetric and thus omitted.

Consider now the optimal $s^A \in \{0, \bar{s}^A\}$. Let

$$S^{DSO}(\tilde{s}^A, \mathbf{x}) = \arg \max_{\tilde{s}^{DSO} \in [0, \bar{s}^{DSO}]} \Omega(\tilde{s}^{DSO}, \tilde{s}^A, \mathbf{x})$$

be the optimal consumption of the aggregator controlled by the DSO as a function of the consumption $\tilde{s}^A$ by all other aggregators and demand $\mathbf{x}$. Let $\omega(\tilde{s}^A, \mathbf{x}) = \Omega(S^{DSO}(\tilde{s}^A, \mathbf{x}), \tilde{s}^A, \mathbf{x})$ be the value function of the integrated DSO/aggregator. Then, $s^A = \bar{s}^A$ for demand configuration $\mathbf{x}$ if $\omega(\bar{s}^A, \mathbf{x}) > \omega(0, \mathbf{x})$, and $s^A = 0$ if the strict inequality is reversed. I establish in three claims that $s^A = \bar{s}^A$ for all $x_1 > x_2$. The claims establish the optimal choice for the different possible signs of $\partial\Omega(\tilde{s}^{DSO}, 0, \mathbf{x})/\partial\tilde{s}^{DSO}$ evaluated at $\tilde{s}^{DSO} = \bar{s}^{DSO}$ and $x_1 \geq x_2$.

Claim 1 *If* $\frac{\partial\Omega(\tilde{s}^{DSO},0,\mathbf{x})}{\partial\tilde{s}^{DSO}}|_{\tilde{s}^{DSO}=\bar{s}^{DSO}} \geq 0$ *for all* $x_1 \geq x_2$*, then* $s^A = \bar{s}^A$ *for all* $x_1 > x_2$*.*

Proof: By this assumption, $S^{DSO}(0,\mathbf{x}) = \bar{s}^{DSO}$ and therefore $\omega(0,\mathbf{x}) = -P(x_1 + \bar{s}^{DSO})\bar{s}^{DSO}$ for all $x_1 \geq x_2$. Then

$$\omega(\bar{s}^A,\mathbf{x}) - \omega(0,\mathbf{x}) \geq \Omega(0,\bar{s}^A,\mathbf{x}) - \omega(0,\mathbf{x}) = [P(x_1 + \bar{s}^{DSO}) - P(x_2 + \bar{s}^{DSO})]\bar{s}^{DSO} > 0$$

for all $x_1 > x_2$ by $P' > 0$.■

Claim 2 *If* $\frac{\partial\Omega(\tilde{s}^{DSO},0,x_2,x_2)}{\partial\tilde{s}^{DSO}}|_{\tilde{s}^{DSO}=\bar{s}^{DSO}} \leq 0$*, then* $s^A = \bar{s}^A$ *for all* $x_1 > x_2$*.*

Proof: I first show that the right-hand side of

$$\omega(\bar{s}^A,\mathbf{x}) - \omega(0,\mathbf{x}) - [\omega(\bar{s}^A,x_2,x_2) - \omega(0,x_2,x_2)] = \int_0^{\bar{s}^A}\int_{x_2}^{x_1} \frac{\partial^2\omega(\tilde{s}^A,\tilde{x}_1,x_2)}{\partial x_1\partial\tilde{s}^A} d\tilde{x}_1 d\tilde{s}^A \tag{22}$$

is strictly positive for all $x_1 > x_2$ by evaluating the sign of $\frac{\partial^2\omega(\tilde{s}^A,\mathbf{x})}{\partial x_1\partial\tilde{s}^A}$. Seeing as $\frac{\partial^2\Omega}{\partial\tilde{s}^{DSO}\partial\tilde{s}^A} < 0$ and $\frac{\partial^2\Omega}{\partial\tilde{s}^{DSO}\partial x_1} < 0$, it follows that $\frac{\partial\Omega}{\partial\tilde{s}^{DSO}}|_{\tilde{s}^{DSO}=\bar{s}^{DSO}} < 0$ under the assumptions of this claim, and therefore $S^{DSO}(\tilde{s}^A,\mathbf{x}) < \bar{s}^{DSO}$ for all $\tilde{s}^A \in [0,\bar{s}^A]$ and $x_1 > x_2$. Moreover,

$$\frac{\partial\Omega(\tilde{s}^{DSO},0,x_2,x_2)}{\partial\tilde{s}^{DSO}}|_{\tilde{s}^{DSO}=0} = P(x_2+\bar{s}) - P(x_2) + P'(x_2+\bar{s})\bar{s}^{DSO} > 0$$

implies $S^{DSO}(0,x_2,x_2) > 0$. By continuity, $S^{DSO}(\tilde{s}^A,\mathbf{x}) > 0$ also for all $(\tilde{s}^A,x_1) \in [0,\varepsilon]\times[x_2,\delta]$ and some $\varepsilon > 0$ and $\delta > x_2$. If $S^{DSO}(\tilde{s}^A,\mathbf{x}) = 0$, then $\omega(\tilde{s}^A,\mathbf{x}) = -P(x_2+\bar{s}-\tilde{s}^A)\bar{s}^{DSO}$, which is independent of x_1. Hence, $\frac{\partial^2\omega(\tilde{s}^A,\mathbf{x})}{\partial\tilde{s}^A\partial x_1} = 0$ in this case. If $S^{DSO}(\tilde{s}^A,\mathbf{x}) > 0$, then

$$\frac{\partial\omega(\tilde{s}^A,\mathbf{x})}{\partial x_1} = -C''(x_1 + S^{DSO}(\tilde{s}^A,\mathbf{x}) + \tilde{s}^A)S^{DSO}(\tilde{s}^A,\mathbf{x})$$

by the envelope theorem, and I have used $P'(q) = C''(q)$. Hence,

$$\frac{\partial^2\omega(\tilde{s}^A,\mathbf{x})}{\partial x_1\partial\tilde{s}^A} = -C''(Q_1^{DSO})\frac{\partial S^{DSO}}{\partial\tilde{s}^A} - C'''(Q_1^{DSO})S^{DSO}(1+\frac{\partial S^{DSO}}{\partial\tilde{s}^A})$$

where $Q_1^{DSO} = x_1 + S^{DSO} + \tilde{s}^A$. Let $Q_2^{DSO} = x_2 + \bar{s} - S^{DSO} - \tilde{s}^A$. The cross-partial derivative $\frac{\partial^2\omega}{\partial x_1\partial\tilde{s}^A}$ is strictly positive if

$$\frac{C'''(Q_1^{DSO})S^{DSO}}{C''(Q_1^{DSO})} < \frac{-\frac{\partial S^{DSO}}{\partial\tilde{s}^A}}{1+\frac{\partial S^{DSO}}{\partial\tilde{s}^A}} \tag{23}$$

Using

$$\frac{\partial S^{DSO}}{\partial\tilde{s}^A} = -\frac{C''(Q_1^{DSO}) + C'''(Q_1^{DSO})S^{DSO} + C''(Q_2^{DSO}) + C'''(Q_2^{DSO})(\bar{s}^{DSO} - S^{DSO})}{2C''(Q_1^{DSO}) + C'''(Q_1^{DSO})S^{DSO} + 2C''(Q_2^{DSO}) + C'''(Q_2^{DSO})(\bar{s}^{DSO} - S^{DSO})},$$

I obtain

$$\frac{-\frac{\partial S^{DSO}}{\partial\tilde{s}^A}}{1+\frac{\partial S^{DSO}}{\partial\tilde{s}^A}} = 1 + \frac{C'''(Q_1^{DSO})S^{DSO} + C'''(Q_2^{DSO})(\bar{s}^{DSO} - S^{DSO})}{C''(Q_1^{DSO}) + C''(Q_2^{DSO})}.$$

Inequality (23) then holds by assumption (7). It follows that the right-hand side of (22) is strictly positive for all $x_1 > x_2$. Finally,

$$\omega(\bar{s}^A, x_2, x_2) - \omega(0, x_2, x_2) \geq \Omega(\bar{s}^{DSO} - S^{DSO}(0, x_2, x_2), \bar{s}^A, x_2, x_2) - \omega(0, x_2, x_2) = 0$$

completes the proof that $\omega(\bar{s}^A, \mathbf{x}) > \omega(0, \mathbf{x})$.■

Claim 3 *If* $\frac{\partial\Omega(\tilde{s}^{DSO},0,x_2,x_2)}{\partial\tilde{s}^{DSO}}|_{\tilde{s}^{DSO}=\bar{s}^{DSO}} > 0$ *and* $\frac{\partial\Omega(\tilde{s}^{DSO},0,\mathbf{x})}{\partial\tilde{s}^{DSO}}|_{\tilde{s}^{DSO}=\bar{s}^{DSO}} < 0$ *for some* $x_1 > x_2$*, then* $s^A = \bar{s}^A$ *for all* $x_1 > x_2$*.*

Proof: By $\frac{\partial\Omega(\tilde{s}^{DSO},0,\mathbf{x})}{\partial\tilde{s}^{DSO}\partial x_1} < 0$ and continuity, there exists a unique $x_1^c > x_2$ such that

$$\frac{\partial\Omega(\tilde{s}^{DSO}, 0, x_1^c, x_2)}{\partial\tilde{s}^{DSO}}|_{\tilde{s}^{DSO}=\bar{s}^{DSO}} = 0.$$

Hence $S^{DSO}(0, \mathbf{x}) = \bar{s}^{DSO}$ for all $x_1 \in (x_2, x_1^c]$. A line of argument similar to the one used to prove Claim 1, can then be applied to establish $\omega(\bar{s}^A, \mathbf{x}) > \omega(0, \mathbf{x})$ for all $x_1 \in (x_2, x_1^c]$. If $x_1 > x_1^c$, then

$$\omega(\bar{s}^A, \mathbf{x}) - \omega(0, \mathbf{x}) - [\omega(\bar{s}^A, x_1^c, x_2) - \omega(0, x_1^c, x_2)] = \int_0^{\bar{s}^A} \int_{x_1^c}^{x_1} \frac{\partial^2\omega(\tilde{s}^A, \tilde{x}_1, x_2)}{\partial x_1 \partial\tilde{s}^A} d\tilde{x}_1 d\tilde{s}^A \geq 0$$

because $S^{DSO}(\tilde{s}^A, \mathbf{x}) < \bar{s}^{DSO}$ for all $\tilde{s}^A \in [0, \bar{s}^A]$ and $x_1 > x_1^c$. Combining this inequality with $\omega(\bar{s}^A, x_1^c, x_2) > \omega(0, x_1^c, x_2)$ concludes the proof of the claim.■

Summarizing the above three claims yields $s^A = \bar{s}^A$ for all $x_1 > x_2$. By following qualitatively similar steps as the above, it is straightforward to verify that $s^A = 0$ for all $x_1 < x_2$.

I next establish useful properties of s^{DSO}. Since

$$\begin{aligned}\frac{\partial\Omega(\tilde{s}^{DSO}, \bar{s}^A)}{\partial\tilde{s}^{DSO}}|_{\tilde{s}^{DSO}=\frac{1}{2}\bar{s}^{DSO}} = \; & -[P(x_1 + \tfrac{1}{2}\bar{s}^{DSO} + \bar{s}^A) - P(x_2 + \tfrac{1}{2}\bar{s}^{DSO})] \\ & -[P'(x_1 + \tfrac{1}{2}\bar{s}^{DSO} + \bar{s}^A) - P'(x_2 + \tfrac{1}{2}\bar{s}^{DSO})]\tfrac{1}{2}\bar{s}^{DSO} < 0\end{aligned}$$

by $x_1 + \frac{1}{2}\bar{s}^{DSO} + \bar{s}^A - (x_2 + \frac{1}{2}\bar{s}^{DSO}) = x_1 - x_2 + \bar{s}^A > 0$, it follows that $s^{DSO} < \frac{1}{2}\bar{s}^{DSO}$. If also $\frac{\partial\Omega(\tilde{s}^{DSO},\bar{s}^A,\mathbf{x})}{\partial\tilde{s}^{DSO}}|_{\tilde{s}^{DSO}=0} \leq 0$, then $s^{DSO} = 0$. This condition is equivalent to

$$P(x_1 + \bar{s}^A) \geq P(x_2 + \bar{s}^{DSO}) + P'(x_2 + \bar{s}^{DSO})\bar{s}^{DSO}.$$

The inequality is satisfied, for instance, if $\bar{s}^{DSO}$ is sufficiently small. Otherwise, $s^{DSO} \in (0, \frac{1}{2}\bar{s}^{DSO})$ and characterized by the point at which marginal expenditures are equated across the two periods:

$$P(x_1+s^{DSO}+\bar{s}^A)+P'(x_1+s^{DSO}+\bar{s}^A)s^{DSO} = P(x_2+\bar{s}^{DSO}-s^{DSO})+P'(x_2+\bar{s}^{DSO}-s^{DSO})(\bar{s}^{DSO}-s^{DSO}).$$

In interior optimum,

$$\frac{\partial s^{DSO}}{\partial \bar{s}^{DSO}} = \frac{2P_2' + P_2''(\bar{s}^{DSO} - s^{DSO})}{P_1' + P_1'' s^{DSO} + P_2' + P_2''(\bar{s}^{DSO} - s^{DSO})} > 0,$$

where $P_1' = P'(x_1 + s^{DSO} + \bar{s}^A)$, $P_1'' = P''(x_1 + s^{DSO} + \bar{s}^A)$, and P_2' and P_2'' are similarly defined. Observe also that

$$\frac{\partial s^{DSO}}{\partial \bar{s}^{DSO}} - 1 = \frac{-(P_1' + P_1'' s^{DSO})}{P_1' + P_1'' s^{DSO} + P_2' + P_2''(\bar{s}^{DSO} - s^{DSO})} < 0.$$

Part (ii) of the proposition: If $x_1 > x_2$, then $s - \frac{1}{2}\bar{s} = s^{DSO} + \frac{1}{2}(\bar{s}^A - \bar{s}^{DSO})$, which is positive if $\bar{s}^{DSO} < \bar{s}^A$. If $x_1 < x_2$, then $s - \frac{1}{2}\bar{s} = s^{DSO} - \bar{s}^{DSO} - \frac{1}{2}(\bar{s}^A - \bar{s}^{DSO})$, which is negative if $\bar{s}^{DSO} < \bar{s}^A$. Hence, $\bar{s}^{DSO} < \bar{s}^A$ implies $(s - \frac{1}{2}\bar{s})(x_1 - x_2) > 0$ for all $x_1 \neq x_2$. $\bar{s}^{DSO} < \bar{s}^A$ is equivalent to $L^{DSO} < \frac{1}{2}$. If $L^{DSO} = 1$, then $\bar{s}^{DSO} = \bar{s}$ and $\bar{s}^A = 0$, in which case the DSO maximizes $\Omega(s, 0) = \Pi^{DSO}(s) - F - \bar{t}$ over s. In this case, $(\frac{1}{2}\bar{s} - s)(x_1 - x_2) > 0$ for all $x_1 \neq x_2$ by Proposition 2. Since

$$\frac{\partial s}{\partial \bar{s}^{DSO}} = \frac{\partial}{\partial \bar{s}^{DSO}}(s^{DSO} + \bar{s} - \bar{s}^{DSO}) = \frac{\partial s^{DSO}}{\partial \bar{s}^{DSO}} - 1 < 0,$$

it follows that there exists some $\hat{L}^{DSO} \in [\frac{1}{2}, 1)$ such that $(s - \frac{1}{2}\bar{s})(x_1 - x_2) > 0$ for all $x_1 \neq x_2$ if and only if $L^{DSO} < \hat{L}^{DSO}$.to $L^{DSO} \leq \frac{1}{2}$.■

A.6 Proof of Proposition 5

The Lagrangian $\Pi^I(s) + \underline{\lambda} s + \bar{\lambda}(\bar{s} - s)$ of the integrated generator/aggregator is strictly concave by strict concavity of $\Pi^I(s)$:

$$\Pi^{I\prime\prime}(s) = -\sum_{i=1,2}[2C''(q_i) - C'''(q_i)(y - s_i)] = -\sum_{i=1,2} C''(q_i)[2 - \frac{C'''(q_i) q_i}{C''(q_i)} + \frac{C'''(q_i)}{C''(q_i)}(x_i - 2y + 2s_i)] < 0,$$

where I have applied assumption (9). Hence, the profit-maximizing solution $(s^I, \underline{\lambda}^I, \bar{\lambda}^I)$ is uniquely determined by the first-order condition

$$p_2^I - p_1^I + P'(q_1^I)(y - s^I) - P'(q_2^I)(y - \bar{s} + s^I) + \underline{\lambda}^I - \bar{\lambda}^I = 0$$

and complementary slackness conditions

$$s^I \in [0, \bar{s}],\ \underline{\lambda}^I \geq 0,\ \bar{\lambda}^I \geq 0,\ \underline{\lambda}^I s^I = \bar{\lambda}^I(\bar{s} - s^I) = 0,$$

where $q_1^I = x_1 + s^I - y$, $q_2^I = x_2 + \bar{s} - s^I - y$, and $p_i^I = P(q_i^I)$, $i = 1, 2$.

The first property of the optimum is $(\frac{1}{2}\bar{s} - s^I)(x_1 - x_2) \geq 0$, with strict inequality if $x_1 \neq x_2$. To establish this result, observe that marginal profit equals

$$\Pi^{I\prime}(\frac{1}{2}\bar{s}) = P'(x_1 + \frac{1}{2}\bar{s} - y)(y - \frac{1}{2}\bar{s}) - P(x_1 + \frac{1}{2}\bar{s} - y) - [P'(x_2 + \frac{1}{2}\bar{s} - y)(y - \frac{1}{2}\bar{s}) - P(x_2 + \frac{1}{2}\bar{s} - y)]$$

evaluated at $s = \frac{1}{2}\bar{s}$. By way of assumption (9),

$$\frac{\partial}{\partial x_1}\Pi^{I\prime}(\frac{1}{2}\bar{s}) = -C''(x_1 + \frac{1}{2}\bar{s} - y)[1 - \frac{C'''(x_1 + \frac{1}{2}\bar{s} - y)(y - \frac{1}{2}\bar{s})}{C''(x_1 + \frac{1}{2}\bar{s} - y)}] < 0.$$

Since $\Pi^{I\prime}(\frac{1}{2}\bar{s})|_{x_1=x_2} = 0$, it follows that $\Pi^{I\prime}(\frac{1}{2}\bar{s})(x_1 - x_2) < 0$ for $x_1 \neq x_2$. This property of marginal profit implies $(\frac{1}{2}\bar{s} - s^I)(x_1 - x_2) > 0$ for all $x_1 \neq x_2$.

The second property of the optimum is $(s^I - s^{DSO})(x_1 - x_2) \geq 0$. If $s^{DSO} = 0$, then $x_1 \geq x_2$ by Proposition 2, and $(s^I - s^{DSO})(x_1 - x_2) = s^I(x_1 - x_2) \geq 0$. If $s^{DSO} = \bar{s}$, then $x_2 \geq x_1$ again by Proposition 2, and $(s^I - s^{DSO})(x_1 - x_2) = (\bar{s} - s^I)(x_2 - x_1) \geq 0$. The marginal profit has the following property

$$(x_1 - x_2)\Pi^{I\prime}(s^{DSO})|_{s=s^{DSO}} = [C''(q_1^{DSO}) - C''(q_2^{DSO})](x_1 - x_2)y. \tag{24}$$

if $s^{DSO} \in (0, 1)$. If $x_1 > \bar{s} + x_2$, then $s^* < 0$, and the right-hand side of

$$(q_1^{DSO} - q_2^{DSO})(x_1 - x_2) = 2(s^{DSO} - s^*)(x_1 - x_2) \tag{25}$$

is strictly positive. If $x_1 \in (x_2, \bar{s} + x_2]$, then $s^{fb} = s^* \in [0, \frac{1}{2}\bar{s})$ and $|x_1 - x_2| \in (0, \bar{s}]$. Proposition 2 has established $(s^{DSO} - s^{fb})(x_1 - x_2) \geq 0$, with strict inequality if $|x_1 - x_2| \in (0, \bar{s}]$. By implication, the right-hand side of (25) is strictly positive also in this range. Analogous arguments imply that the right-hand side of (25) is strictly positive also for all $x_2 > x_1$. Hence, the right-hand side of (24) is non-negative, and strictly positive if $C''' > 0$ and $x_1 \neq x_2$. I conclude that $(s^I - s^{DSO})(x_1 - x_2) \geq 0$ also for $s^{DSO} \in (0, 1)$. In particular, the inequality is strict if $C''' > 0$, $x_1 \neq x_2$ and $s^{DSO} \in (0, 1)$.■