

Zerrahn, Alexander; Huppmann, Daniel

Conference Paper

Network Expansion to Mitigate Market Power: How Increased Integration Fosters Welfare

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2014: Evidenzbasierte Wirtschaftspolitik
- Session: Energy Markets, No. G11-V4

Provided in Cooperation with:

Verein für Socialpolitik / German Economic Association

Suggested Citation: Zerrahn, Alexander; Huppmann, Daniel (2014) : Network Expansion to Mitigate Market Power: How Increased Integration Fosters Welfare, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2014: Evidenzbasierte Wirtschaftspolitik - Session: Energy Markets, No. G11-V4, ZBW - Deutsche Zentralbibliothek für Wirtschaftswissenschaften, Leibniz-Informationszentrum Wirtschaft, Kiel und Hamburg

This Version is available at:

<https://hdl.handle.net/10419/100459>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Network Expansion to Mitigate Market Power: How Increased Integration Fosters Welfare

Alexander Zerrahn^{*}, Daniel Huppmann[†]

February 28, 2014

Preliminary version, please do not publish or quote

Abstract

Lack of transmission capacity hampers the efficient integration of the European electricity market, and thereby precludes reaping the full benefits of competition. We investigate to what extent the expansion of the transmission grid promotes competition, efficiency, and welfare. This work proposes a three-stage model for grid investment: a benevolent planner decides on network upgrades; she considers the welfare benefits of investment through a reduction of market power exertion by strategic generators. These firms anticipate their impact on the Independent System Operator and are able to exert market power, in particular when lines are congested.

We illustrate the model on a simple three-node network. Results indicate that network expansion indeed provides a suitable way of enhancing welfare due to a reduction of market power potential.

Keywords: market power, transmission grid expansion, electricity, integration, EPEC, multistage games, Cournot competition

JEL classification: L13, L51, C61

^{*}German Institute for Economic Research (DIW Berlin), Mohrenstr. 58, D-10117 Berlin, Germany, azerrahn@diw.de

[†]German Institute for Economic Research (DIW Berlin), Mohrenstr. 58, D-10117 Berlin, Germany, dhuppmann@diw.de

1 Market Power and Network Expansion in Europe

Starting in the mid-nineties, the European Union has been developing the institutional basis for an internal energy market (IEM). Initiated by numerous directives, formerly integrated companies were unbundled, and the transmission grid opened up for competitive generation. As a consequence, regional energy exchanges were established, providing market places with transparent wholesale prices. Interconnectors between those single market areas had originally been constructed for contingencies, not with the aim of facilitating cross-border trade – national markets basically remained in autarky. Today, in ten Member States the largest generating company has a market share above 70% (European Commission, 2012b, data for 2010), and in eight Member States more than 80% of generating assets are held by the former incumbent (European Commission, 2012a, data for 2012). By means of further European integration, this high degree of concentration could be reduced, thereby mitigating the potential to exert market power, enhancing efficiency, and increasing welfare. Along this basic economic rationale, policymakers acknowledge the gains from further integration:

The European Union needs an internal energy market that is competitive, integrated and fluid providing a solid backbone for electricity and gas flowing where it is needed. [...] Despite major advances in recent years in the way the energy market works, more must be done to integrate markets, improve competition and respond to new challenges. (European Commission, 2012a, p.2)

The formation of the Council of European Energy Regulators (CEER) in 2000, and the Agency for the Cooperation of Energy Regulators (ACER) in 2010 laid institutional cornerstones. The organizational framework to accomplish enhanced integration and competition, is provided by different specifications of market coupling, eventually creating a common electricity market. In the Central Western Europe (CWE) region comprising France, the Benelux states, and Germany/Austria, for example, the current market coupling algorithm was introduced in 2010, extending a similar mechanism comprising France, the Netherlands, and Belgium.¹ Unlike the prior explicit auctioning of interconnector capacities, the CWE coupling mechanism employs an implicit auction algorithm in order to utilize existing capacities efficiently, and directing flows from low to high price areas. Throughout Europe, further national markets are also on their way to integration through several regional coupling initiatives. The ultimate goal is laid down by the so-called Electricity Target Model (ETM) of market integration, envisaging a single pan-European price coupling mechanism (ACER/CEER, 2013).

What is missing, however, are sufficient physical interconnector capacities to accomplish a truly integrated market. Strong indicators for factually incomplete integration within Europe come from analyses of wholesale price spreads. Zachmann (2008) econometrically examines the period from 2002 to 2006: he identifies significant convergence between some bilateral wholesale price pairs. Especially during peak hours of high flows within the network, however, tendencies of divergence are observable.² A similar detection for more recent data from 2008 to 2012 is made by ACER/CEER (2013): in 2012, full price convergence in CWE, for example, is achieved in about 50% of hours, whereas in roughly a quarter of the year wholesale prices differed by more than 10 Euro per MWh. It is important to note that these trends coincide with an increasing use of commercial transfer capacities, and a declining number of flows against the price differential. Böckers et al. (2013) provide analogous figures. These findings, thus, lend stark evidence that, first, physical capacities are still insufficient

¹Before the introduction of market coupling, interconnector capacities were auctioned explicitly, often weeks or months ahead, which inevitably led to inefficiencies like flows against the prevailing price differential (cf. Booz & Company et al., 2013).

²He also identifies the explicit auctioning of cross-border capacities as significant obstacle to integration, as discussed above.

to achieve full integration, and, second, inefficiencies in the allocation mechanism are no predominant driver of this incomplete convergence.

From a pure gains-from-trade perspective, ACER launched attempts to quantify welfare benefits from integration. To this end, several European Power Exchanges were prompted to carry out counterfactual analyses, holding historically received bids constant while varying interconnector capacities. According to these calculations, gross welfare gains up to 25 million Euro for a single interconnector would have accrued for a virtual increase of 100 MW in 2012 (ACER/CEER, 2013, pp.73-75). In an analogous fashion, Booz & Company et al. (2013) quantifies the efficiency losses due to congestion on the interconnector between France and the UK with more than 20 million Euro in 2011, and more than 70 million Euro in 2012. A similar analysis is provided on a monthly basis by the European power exchange EPEX Spot: in its *Social Welfare Report*, historical data are compared to a hypothetical world with infinite cross-border capacities. For the CWE region, gross welfare increases of more than 250 million Euro would have originated from the absence of any congestion (EPEX Spot, 2013). Beyond these (rough approximations of) efficiency gains from trade, enhanced security of supply or more efficient feed-in of variable renewables³, integration is essential to ensure vigorous inter-regional competition where concentration on single markets remains high – which is the focus of our contribution.

Modeling the interaction between strategic firms and the expansion of constraining networks is, however, still challenging – especially to properly account for gaming opportunities between generators and network operators. In our paper, we investigate to which extent transmission grid expansion between national markets offers a way to realize welfare improvements due to reduced market power. The major contribution consists in directly incorporating the tradeoff between costs and benefits of network expansion into one integrated model. In a three-stage game, we account for the fact that each agent anticipates the impact of her actions on subsequent decisions: the transmission system planner for her effect on strategic firms' behavior, and those, in turn, for their effect on the network operator. Thereby, we endogenize transmission grid expansion decisions and do not have to rely on the exogenous variation of parameters or the analysis of scenarios or cases. To the best of our knowledge, this is the first contribution exhibiting this tradeoff, and provides a consistent framework for future applications.

The remainder of this paper is structured as follows: Section 2 gives an overview of the existing literature and relates it to our research, Section 3 presents the methodological formulation, Section 4 delivers and discusses numerical results, and Section 5 provides a wrap-up.

2 Modeling Market Power in Electricity Networks

We develop a three-stage model capturing the interaction of strategic generators and electricity grid expansion. On the third stage an Independent System Operator (ISO) dispatches competitive fringe plants, and ensures feasible network flows, assigning locational prices to each node in the system, taking all upper level decisions as given. Reformulating her Karush-Kuhn-Tucker (KKT) optimality conditions with help of duality theory establishes the solution space for the second stage: strategic generators maximize profit in Cournot competition, anticipating their impact on the ISO's actions, but treating first level decision as exogenous. These two lower stages of the model describe the spot market phase – mathematically they constitute an Equilibrium Problem under Equilibrium Constraints (EPEC). Deriving its KKT conditions renders a set of inequalities capturing all stationary points, among which a multitude of extrema and saddle points may be found. Among this set of vectors fulfilling the necessary conditions for optimality, the first model stage serves as

³A basic exposition of the economics of European electricity market integration can be found in Böcker et al. (2013).

a selection device: a benevolent planner maximizes social welfare over expanding links in the network, thereby picking that stationary point rendering the welfare-optimal solution. Beyond necessity, we have to ensure that proposed solutions are indeed Nash equilibria. To this end, a post-solve check confirms incentive compatibility for the strategic players in order to ensure sufficiency. To account for multiple optima, we implement an iterative algorithm systematically exploring the solution space.

The analysis of strategic generator behaviour in constraining networks has been drawing the attention of the research community for several years. Since Neuhoff et al. (2005)⁴ we know that *the devil is in the details*: comparing different approaches concerning the treatment of transmission constraints in two-stage models, the authors identify difficulties in rendering a realistic representation of interactions between strategic generation and clearing of multiple markets. Specifically, there are two methods to incorporate the TSOs' optimization programs (cf. Hobbs et al., 2005), depending on whether strategic players anticipate their impact on the network operation (Stackelberg) or not (Bertrand or naïve). The latter perspective considerably reduces model complexity at the cost of excluding strategic effects. Examples comprise the exogenous assumption of rationing mechanisms in case transmission capacity is sparse (Willems, 2002), or, more recently, strategic players treating transmission charges arising from the TSO optimization as exogenous in their constraint sets (Tanaka, 2009). The Stackelberg assumption is pursued, for example, in Hobbs et al. (2000), proposing an algorithmic solution of the arising EPEC, or Cunningham et al. (2002), who explicitly derive reaction functions for special cases. For all of these approaches, however, network expansion remains exogenous to the model and is restricted to a limited number of cases in varying line constraint parameters.

For a reduced special case, the assumption of strategic firms anticipating their impact on the network situation found reflection in the theoretical industrial organization literature: in their seminal contribution, sometimes referred to as the *thin-line paper*, Borenstein et al. (2000) demonstrate that in a simple two-node network even a line with relatively low capacity may be enough to foster competition and evoke substantial welfare gains. In this context, it is irrelevant whether the line capacity does in fact suffice to realize the threat of harsh competition, or is even utilized. For certain capacity levels, however, there may exist no Nash equilibrium in pure strategies at all.

Combining the so-called peak load pricing literature with the presence of a potentially limiting transmission grid, another, more recent strand of the electricity economic literature on producer market power focused on plant investments from a theoretical perspective: Ruderer and Zöttl (2012) compare the outcomes of different transmission management systems⁵ on investment incentives into generation and transmission. They derive that the social optimum is achieved under nodal pricing. This approach is extended to strategic Cournot firms by Léautier (2013). He finds that under certain assumptions the marginal social value of augmenting the capacity of a congested transmission line is positive, and isolates the fraction that can be attributed to increased competition intensity.

The recent literature also features some empirical approaches, each with a distinct regional scope: for cross-border trade flows between particular European countries, Gebhardt and Höfler (2013) compare a theoretical benchmark with actual data in order to investigate whether wholesale price spreads between countries originate from limited interconnector capacities or rather from missing competition incentives. They presume that prima facie irrational non-participation in cross-border trade – although positive profits could be expected – could be incentivized by dynamically upholding market power instead of engaging in potential retaliation in supergame structures.⁶ For the Indian electricity market, Ryan

⁴The article also provides a comprehensive literature review on strategic two-stage games in constraining networks up to the date of publication

⁵Nodal pricing directly incorporating scarce transmission capacity into locational prices, as in our model, and a redispatch system in which a subsequent market corrects for potentially infeasible flows resulting from a spot dispatch not taking network capacities into account.

⁶Note that their analysis is based on the European market organization in the years 2002 to 2006, where

(2013) analyzes the impact of transmission expansion between regions which are frequently separated by congestion. To this end, he imposes hypothetical line upgrades and presents a counterfactual analysis disentangling welfare gains from different sources. He finds that 72% of the detected welfare increase can be traced back to decreased market power exertion. Finally, Wolak (2012) studies supplier behaviour in the wholesale electricity market of Alberta (Canada). Within the framework of a theoretical model of bidding strategies, he establishes a causality running from firms' perception of the likelihood of transmission constraints to bind to substantially reduced market power exertion – realized as wedge between actual marginal costs and submitted bidding curves. This perception of increased competition intensity can, naturally, be promoted by actual transmission expansion. A model calibration finds consumer surplus effects between roughly 30 and 700 million Canadian Dollars per year.

There is a modeling stream similar to the work presented here: under the label *proactive planning*, Sauma and Oren (2006) present a methodology to evaluate network expansion projects such that their effect on strategic players is taken into account. Variations and extensions of this paper are brought forward in Pozo et al. (2013a) and Pozo et al. (2013b), who propose an improved solution technique allowing for more flexibility concerning line expansions. The results of all three analyses are in line with intuition: network expansion has the potential to enhance social welfare, and this increase is greater in case the network planning entity proactively takes strategic behaviour into account. Although the studies are based on three-stage models, and incorporate strategic interactions among generating firms, we depart from their approach in multiple significant ways, and aim at filling several research gaps:

First, by employing results from duality theory following Ruiz et al. (2012), we are able to reformulate the market model EPEC such that first order conditions can be explicitly derived. The virtue of this methodological contribution lies in circumventing the inconvenient usage of iterative algorithms or discretizations of decision variables. Thereby, we gain flexibility in employing a first-stage optimization function as selection mechanism choosing among equilibrium points of the EPEC – compared to the analysis of a small number of predefined cases. Second, we decidedly focus on market power exertion and strategic withholding. To this end, we contrast our model results with those of several benchmark runs excluding network expansions or strategic firm behaviour⁷.

Third, we suggest a formulation allowing to employ network as well as demand and generation data as a stylized representation of actual power systems.⁸ In this respect, we incorporate different generation technologies and fringe suppliers, aiming at providing more realistic results. As a further important difference, note that we do not consider investments in generation capacity while inelastic demand is satisfied by a perfectly competitive dispatch as in Pozo et al. (2013a) and Pozo et al. (2013b), but rather strategic interaction on a spot market with existing capacities. This choice deserves some explanation why we are convinced it is a good approach: in Sauma and Oren (2006), and Pozo et al. (2013a) there is only one generation technology whose marginal production costs decrease in capacity invested, an assumption needed to justify investment activities by firms.⁹ This might be a viable approximation if one only regarded one stylized technology comprising the entire generation

cross-border transmission capacity was auctioned off explicitly. Moreover, they do not model generators as strategic, but rather hypothesize strategic behaviour as described.

⁷Pozo et al. (2013a), Pozo et al. (2013b), and to a greater extent Sauma and Oren (2006) do indeed regard market power, the latter as a leitmotif, as their analyses center on strategic players. Their discussions, however, do not detail the impact of network expansion on market power exertion. Sauma and Oren (2006) rather focus on comparing different network planning paradigms. Pozo et al. (2013a) and Pozo et al. (2013b) mention positive profits of competitive firms originating in strategic generation investment. However, neither do they provide a deeper discussion nor present a tradeoff between market power rents and potentially costly mitigation due to transmission expansion.

⁸Pozo et al. (2013b) do also provide a case study of the Chilean power system.

⁹Pozo et al. (2013b) incorporate different conventional technologies. There is, however, only one distinct technology per node and investment is possible only into that technology. The other assumptions remain.

mix. As we go for a representation of the existing technology mix, however, considerably lower marginal costs due to a “larger” generating unit do not appear convincing. Moreover, we believe that decisions to invest in certain technologies are also driven by other substantial factors, which are not captured by a model of this kind, than the potential exertion of market power due to limited network capacities. Among these are the dynamic patterns of demand and renewables, the existing fleet of a firm, aspects of financing and fuel prices, or innovation processes. For example, the publication by Léautier (2013), discussed above, and in a more reduced form Zöttl (2011), exhibit that the analysis of strategic investment behaviour even with only two distinct technologies (base and peak) is already quite involved, and results are substantially related to assumptions on the occurrence of different spot markets¹⁰ or the policy framework.

As a final remark, Kreps and Scheinkman (1983) pointed out in their seminal contribution that Cournot competition can be regarded as equivalent to a two-stage game of strategic capacity investments with subsequent price competition. Pozo et al. (2013a) note this analogy – due to their approach of discretizing decision variables, however, they admit that an interpretation in that fashion does not hold true any more. As we do not apply such model reformulations, our approach lends itself to evoke that parallel and reinterpret the Cournot market as investment game, in case one wants to pursue that view.

Subsuming, if the potential to exert market power with an existing power plant park does exist, strategic firms will have the incentive to exploit it. We therefore restrict ourselves to the analysis of strategic spot market behaviour under endogenous network configurations.

3 The Three-Stage Model

We propose a three-stage model in which a benevolent social planner anticipates the reactions of the lower stage strategic players to transmission grid expansion: on the third model stage, an ISO optimizes the dispatch of non-strategic fringe power plants and ensures that resulting flows are feasible with given network capacities – taking strategic firms’ generation and network expansion as given.

On the second model stage, the strategic generators, in turn, anticipate the effect of their generation decisions on the ISO. In particular, they can potentially generate excess returns by appropriating rents through congesting the network. Among each other, the strategic generators play a Nash-Cournot game; that is, they anticipate the ISO’s reaction on the third stage, while taking the generation levels of other strategic firms and the first stage network expansion decisions as given. By withholding capacity, they are able to increase prices above the competitive level. The spot market therefore constitutes a two-stage game, with strategic firms on the upper-level and the ISO representing the lower level. We use the notation of an ISO for convenience, but it is equivalent to the equilibrium of an otherwise competitive market. To this end, consider the ISO as that entity which assigns nodal prices after having received bids by strategic players, and thus clears the market by dispatching strategic and competitive fringe plants. Respecting network constraints, she directs flows in a welfare-optimal manner, mimicking the coupling of nodal markets in the most efficient way. Actual market coupling procedures are, in some European regions, based on net transfer capacity (NTC)-based systems, in which commercial flows may underuse existing capacities. Our approach thus renders an upper bound for efficiency and welfare.¹¹

¹⁰More precisely, the expectations about the occurrence of different demand scenarios. Related to that point, there is an ongoing debate about the appropriate market design to provide firms with sufficient incentives to invest into capacity such that security of supply is maintained. Compare Cramton et al. (2013) for an overview.

¹¹In an NTC system, a fixed calculatory amount of interconnector capacity between two countries is quoted in advance, and the cheapest bids from one market to another are accepted as long as this capacity limit is reached. Actually resulting physical flows, however, may diverge from merchant flows due to Kirchhoff’s laws. In case physical interconnector capacities are violated, counter-trading has to be carried out. Compare Oggioni et al. (2012) for an analysis of this issue.

Level	Time	Player(s) and decisions
I	Network expansion	<i>Welfare-maximizing planner</i> transmission capacity investment
II	Spot Market	<i>Strategic Generators</i> generation level(s) at each node
III		<i>Independent System Operator (ISO)</i> dispatch of competitive generation, load, prices, network flows subject to network feasibility

Table 1: Model Structure

At the top stage of our model, a benevolent social planner maximizes total welfare by deciding upon the level of transmission grid expansion. While anticipating how changes in the network topology will influence the Nash equilibrium outcome on the lower stages, she faces a trade-off between costs of grid expansion and the welfare-enhancing effect of integration between the different regions: by increasing competition among the strategic generators, their potentially dominant positions are weakened and thus the potential to exert market power is reduced. Table 1 illustrates the model structure.

To solve for the Nash equilibrium of this three stage game, we follow the methodology proposed by Ruiz et al. (2012). We reformulate the ISO’s optimization problem on the lowest stage (III) using strong duality, which enables us to get rid of bilinearities arising from complementarity slackness constraints, and replace it by a system of equalities and inequalities. As we employ an elastic demand curve — an important assumption when it comes to analyzing market power — we extend the approach from the literature by incorporating nonconvexities. Our ISO problem is hence quadratic and not linear, slightly complicating the duality reformulation. These third stage constraints then provide the solution space for the second stage strategic generators’ Generalized Nash game (GNE). For each of those, we derive the KKT conditions, and combine them to form one Equilibrium Problem under Equilibrium Constraints (EPEC). We now face two problems: first, there may exist multiple Nash equilibria – as demonstrated by Borenstein et al. (2000). Second, due to the nonconvexity of the EPEC, stationary points do not necessarily capture a Nash equilibrium, but may also describe minima for some strategic players or saddle points. To address these issues, we utilize the overall game’s first stage as selection mechanism: the benevolent social planner optimizes overall welfare deciding on network expansion, and trading off its costs against the positive welfare effects on the underlying spot market, while the lower level is rendered by the Karush-Kuhn-Tucker (KKT) first-order optimality conditions of all strategic generators. Our approach thus picks that level of transmission grid expansion which yields a welfare second best solution, considering that generators behave strategically. Using a disjunctive constraints reformulation, the resulting model is a non-convex Mixed-Integer Quadratic Problem, which can be solved using standard approaches (Tawarmalani and Sahinidis, 2005).

In order to address the potential multiplicity of solutions, we implement an iterative algorithm, exploring the solution space by rendering candidate vectors solving the overall problem, but differing in quantities produced and the magnitude of network expansion. In a second step, we solve for each solution candidate an optimization program for each strategic firm, holding decisions of other firms constant, and check whether the particular quantity is incentive compatible.

Throughout the analysis, keep in mind that we assume a system with nodal prices, and a welfare-optimal exchange between single nodes accomplished by the ISO. In the following, we will present our model from a mathematical point of view. For load flows, we use a DC approximation based on H and B matrices, which is equivalent to a *ptdf* formulation (for a detailed exposition, cf. Leuthold et al., 2012).

Stage III: A Competitive Market – or the ISO

On the third model stage, the ISO maximizes welfare — the sum of consumer surplus, generator profit and congestion rent, minus variable production costs of utilized plants — over the dispatch of nonstrategic fringe plants at all nodes n , consumed quantities d_n and voltage angles δ_n , while rendering feasible network flows. We employ a lossless DC load flow approximation incorporating loop flows by voltage angles at each node (Schweppe et al., 1988). Its constraint set consists of the nodal balances, feasibility flows on each line as well as maximum generation and positivity restrictions, each with its respective dual variable. Inverse nodal demand is assumed linear with reservation price a_n and slope parameter b_n . Note that index s indicates an individual plant, where each plant is exogenously mapped to a node n . An additional superscript F denotes that the respective plant belongs to the fringe supply being optimized by the ISO. All other plants, indicated by superscript S , are owned by strategic generators – therefore their generation is treated as exogenous parameter by the ISO. Generation quantities for each plant s are rendered by g_s^S or g_s^F , respectively. From the ISO's point of view, the line expansion e_l is an exogenous parameter. For notational convenience, we introduce set \mathfrak{S}_n containing all plants at node n , and assemble all dual variables in set \mathfrak{D} , all fringe plants in \mathfrak{F} , and all nodes at which nonzero demand is located in \mathfrak{N} . The optimization problem reads as follows:

$$\min_{d, \delta, g^F} - \sum_{n \in \mathfrak{N}} [(a_n - \frac{1}{2} b_n d_n) d_n] + \sum_{s \in \mathfrak{F}} c_s^G g_s^F \quad (1a)$$

$$\text{s.t.} \quad - \sum_{s \in \mathfrak{S}_n} (g_s^S + g_s^F) + \sum_k B_{nk} \delta_k + d_n = 0 \quad (p_n) \quad \forall n \quad (1b)$$

$$-\bar{f}_l - e_l + \sum_k H_{lk} \delta_k \leq 0 \quad (\bar{\mu}_l) \quad \forall l \quad (1c)$$

$$-\bar{f}_l - e_l - \sum_k H_{lk} \delta_k \leq 0 \quad (\underline{\mu}_l) \quad \forall l \quad (1d)$$

$$-\bar{g}_s^F + g_s^F \leq 0 \quad (\beta_s) \quad \forall s \in \mathfrak{F} \quad (1e)$$

$$-g_s^F \leq 0 \quad (\psi_s) \quad \forall s \in \mathfrak{F} \quad (1f)$$

$$-d_n \leq 0 \quad (\phi_n) \quad \forall n \in \mathfrak{N} \quad (1g)$$

$$\delta_{\hat{n}} = 0 \quad (\gamma) \quad (1h)$$

The hub node of the network is given by \hat{n} . In case there is no demand at node n , we fix d_n and ϕ_n at zero. Differentiating yields the respective first order KKT conditions:

$$c_s^G - p_{n, s \in \mathfrak{S}_n} + \beta_s - \psi_s = 0 \quad \perp \quad g_s^F \text{ (free)} \quad \forall s \in \mathfrak{F} \quad (2a)$$

$$-a_n + b_n d_n + p_n - \phi_n = 0 \quad \perp \quad d_n \text{ (free)} \quad \forall n \in \mathfrak{N} \quad (2b)$$

$$\sum_k B_{kn} p_k + \sum_l H_{ln} (\bar{\mu}_l - \underline{\mu}_l) + \begin{cases} \gamma & \text{if } n = \hat{n} \\ 0 & \text{else} \end{cases} = 0 \quad \perp \quad \delta_n \text{ (free)} \quad \forall n \quad (2c)$$

$$\sum_{s \in \mathfrak{S}_n} (g_s^S + g_s^F) - \sum_k B_{nk} \delta_k - d_n = 0 \quad \perp \quad p_n \text{ (free)} \quad \forall n \quad (2d)$$

$$\bar{f}_l + e_l - \sum_k H_{lk} \delta_k \geq 0 \quad \perp \quad \bar{\mu}_l \geq 0 \quad \forall l \quad (2e)$$

$$\bar{f}_l + e_l + \sum_k H_{lk} \delta_k \leq 0 \quad \perp \quad \underline{\mu}_l \geq 0 \quad \forall l \quad (2f)$$

$$\bar{g}_s^F - g_s^F \geq 0 \quad \perp \quad \beta_s \geq 0 \quad \forall s \in \mathfrak{F} \quad (2g)$$

$$g_s^F \geq 0 \quad \perp \quad \psi_s \geq 0 \quad \forall s \in \mathfrak{F} \quad (2h)$$

$$d_n \geq 0 \perp \phi_n \geq 0 \quad \forall n \in \mathfrak{N} \quad (2i)$$

$$\delta_n = 0 \perp \gamma \text{ (free)} \quad (2j)$$

Recall that the ISO's problem constitutes the third stage of our multistage optimization model. On the second stage, the strategic firms come into play and seek to optimize their profits over quantities, at the same time anticipating the subsequent ISO's actions. At this point, note the formal problem we have to overcome: conditions (2a) – (2j) come in complementarity form and thus cannot serve as constraint set for the second stage. In order to transform the third-stage equilibrium conditions into an operational form that can directly be utilized as feasible set for the strategic players, we have to get rid of the complementary slackness conditions, represented by the perpendicular operator \perp . There are three options: first, a disjunctive constraints reformulation, expressing the complementarity requirement with help of dummy variables, as for example employed by Gabriel and Leuthold (2010). Second, the so-called SOS1 method (Siddiqui and Gabriel, 2013), and third the application of duality theory. The first two alternatives come with the drawback of introducing a multitude of additional constraints and binary variables, complicating the analysis computationally, and requiring difficult calibration of large scalars serving as upper bounds for new auxiliary inequalities. More importantly, we then would be left with a discretely-constrained Nash-Cournot game for which it is, again, not possible to derive KKT conditions without further modifications. Although theoretical research has made progress in that field – Gabriel et al. (2013) propose a method to tackle that issue by relaxing both integrality and complementarity – it is, however, still ongoing and does not provide a ready-made toolbox yet. We therefore pursue the third option and make use of properties stemming from duality theory (cf. Boyd and Vandenberghe, 2004, ch.5). To this end, we first set up the ISO's equivalent dual problem.

Proposition 1

The equivalent dual problem to (1a) - (1h) is given by

$$\begin{aligned} \max_{p, \bar{\mu}, \underline{\mu}, \beta, \phi, \psi, \gamma} \quad & -\frac{1}{2} \sum_{n \in \mathfrak{N}} \frac{1}{b_n} [a_n - p_n + \phi_n]^2 - \sum_l (\bar{f}_l + e_l)(\bar{\mu}_l + \underline{\mu}_l) \\ & - \sum_n p_n \left(\sum_{s \in \mathfrak{S}_n} g_s^S \right) - \sum_{s \in \mathfrak{S}} \beta_s \bar{g}_s^F \end{aligned} \quad (3a)$$

$$\text{s.t.} \quad (2a), (2b), (2c) \text{ without complementarity,} \quad (3b)$$

$$\bar{\mu}, \underline{\mu}, \phi, \beta, \psi \geq 0, p, \gamma \in \mathbb{R} \quad (3c)$$

Proof:

See Appendix A.1 ■

For a primal problem, here the ISO minimization of negative welfare, the dual problem consists in finding that dual variables which maximize the Lagrangian value function – that is in detecting the greatest minimum of the primal problem over all elements of \mathfrak{D} . In any case, the optimal value of the dual problem is no larger than the optimal value of the primal problem, a characteristic that holds by definition. The difference between these two values is called duality gap. Under certain conditions, the duality gap collapses to zero – a property referred to as strong duality. One condition triggering strong duality is fulfilled when the primal's objective is convex, and Slater's constraint qualification holds. Both is the case in our setup. The first requirement is easily verified, the second follows from all inequality

constraints being affine. The optimal values of the primal function and the dual function, thus, are identical. As, moreover, the primal objective is strictly convex, the KKT conditions are necessary and sufficient to describe the unique global optimum of the ISO problem.

Therefore, both the KKT conditions and the identity of the primal and dual functions are equivalent descriptions of the unique global solution of the ISO welfare maximization problem. Replacing the former by the latter, hence, leaves us with a representation of optimality on stage three without explicitly incorporating mathematically inconvenient complementarity conditions. Concretely, instead of (2a) - (2j), we employ:

$$(1b) - (1h), \quad (2a) - (2c) \text{ without complementarity} \quad (4a)$$

$$- \sum_{n \in \mathfrak{N}} \left[\left(a_n - \frac{1}{2} b_n d_n \right) d_n \right] + \sum_{s \in \mathfrak{S}} c_s^G g_s^F + \frac{1}{2} \sum_{n \in \mathfrak{N}} \frac{1}{b_n} [a_n - p_n + \phi_n]^2 +$$

$$+ \sum_n p_n \left(\sum_{s \in \mathfrak{S}_n} g_s^S \right) + \sum_l (\bar{f}_l + e_l) (\bar{\mu}_l + \underline{\mu}_l) + \sum_{s \in \mathfrak{S}} \beta_s \bar{g}_s^F \leq 0 \quad (4b)$$

where (4a) simply ensures feasibility for both the primal and dual problem, and (4b) imposes a zero optimality gap¹². We thus express the problem's solution by means of a set of equalities and inequalities without having to make use of complementarity slackness conditions. Thereby, we are able to explicitly derive KKT conditions of the second-level players on the set of vectors capturing ISO optimal behaviour.

Stage II: Strategic Firms

Now, we introduce the second model stage: strategic firms, i , own plants s at nodes n . Letting \mathfrak{S}_i denote the set of all plants owned by strategic firm i , g_i^S the respective production vector, and \mathfrak{D}_i^S the set of Lagrange multipliers attached to i 's optimization program. Firm i 's minimization task thus reads:

$$\min_{g_i^S, g^F, \delta, d, \mathfrak{D}} - \sum_n \sum_{s \in (\mathfrak{S}_n \cap \mathfrak{S}_i)} (p_n - c_s^G) g_s^S \quad (5a)$$

where each firm i faces:

$$-\bar{g}_s^S + g_s^S \leq 0 \quad (\beta_s^S) \quad \forall s \in \mathfrak{S}_i \quad (5b)$$

$$-g_s^S \leq 0 \quad (\psi_s^S) \quad \forall s \in \mathfrak{S}_i \quad (5c)$$

and the following constraints arising from the ISO problem:

$$\text{s.t.} \quad c_s^G - p_{n, s \in \mathfrak{S}_n} + \beta_s - \psi_s = 0 \quad (\zeta_{si}^S) \quad \forall s \in \mathfrak{S} \quad (5d)$$

$$-a_n + b_n d_n + p_n - \phi_n = 0 \quad (\eta_{ni}^S) \quad \forall n \quad (5e)$$

$$\sum_k B_{kn} p_k + \sum_l H_{ln} (\bar{\mu}_l - \underline{\mu}_l) + \begin{cases} \gamma & \text{if } n = \hat{n} \\ 0 & \text{else} \end{cases} = 0 \quad (\theta_{ni}^S) \quad \forall n \quad (5f)$$

$$- \sum_{s \in \mathfrak{S}_n} (g_s^S + g_s^F) + \sum_k B_{nk} \delta_k + d_n = 0 \quad (\iota_{ni}^S) \quad \forall n \quad (5g)$$

$$-\bar{f}_l - e_l + \sum_k H_{lk} \delta_k \leq 0 \quad (\bar{\mu}_{li}^S) \quad \forall l \quad (5h)$$

$$-\bar{f}_l - e_l - \sum_k H_{lk} \delta_k \leq 0 \quad (\underline{\mu}_{li}^S) \quad \forall l \quad (5i)$$

¹²Note that for convenience we do not state an equality here. By definition, the primal's objective is weakly greater than the dual's. Requiring the reverse inequality to hold thus imposes identity

$$-\bar{g}_s^F + g_s^F \leq 0 \quad (\beta_{si}^{SF}) \quad \forall s \in \mathfrak{F} \quad (5j)$$

$$-g_s^F \leq 0 \quad (\psi_{si}^{SF}) \quad \forall s \in \mathfrak{F} \quad (5k)$$

$$-d_n \leq 0 \quad (\phi_{ni}^S) \quad \forall n \quad (5l)$$

$$\delta_{\hat{n}} = 0 \quad (\gamma_i^S) \quad (5m)$$

$$\begin{aligned} & - \sum_{n \in \mathfrak{N}} \left[(a_n - \frac{1}{2} b_n d_n) d_n \right] + \sum_{s \in \mathfrak{F}} c_s^G g_s^F + \frac{1}{2} \sum_{n \in \mathfrak{N}} \frac{1}{b_n} [a_n - p_n + \phi_n]^2 + \\ & + \sum_n p_n \left(\sum_{s \in \mathfrak{S}_n} g_s^S \right) + \sum_l (\bar{f}_l + e_l) (\bar{\mu}_l + \underline{\mu}_l) + \sum_{s \in \mathfrak{F}} \beta_s \bar{g}_s^F \leq 0 \quad (\xi_i^S) \end{aligned} \quad (5n)$$

where superscript S indicates a strategic firm's problem. As each firm i faces the analogous optimization, the spot market, captured by stages II and III, mathematically constitutes an Equilibrium Problem under Equilibrium Constraints (EPEC). Due to our duality reformulation of the third-stage optimality constraints, as derived in the previous section, however, we circumvent difficulties in solving this type of model. Observe that constraints (5d) - (5n) are identical for each player whereas the attached Lagrange multipliers in brackets are specific to each firm i . Formally, this constitutes a Generalized Nash Equilibrium (GNE) setup of shared constraints, which comes along with the feature that there are more endogenous variables, that is Lagrange multipliers, than distinct equations. Mathematically, an underdetermined system of equations emerges. The implied degrees of freedom, therefore, potentially admit a multitude of solutions. Note, however, that the strategic firms are not bound in their decisions by the output choice of their competitors although the quantities of all firms enter the constraint set of each of them. Formally, they do not only interact over their objective functions, but also over the constraint sets – practically, however, the subsequent ISO stage accommodates all strategic decisions such that the market clears in a feasible way. As a second point, by (5n) the constraint set is nonconvex such that KKT points are necessary, but not sufficient for an equilibrium. We will return to both issues later, and first derive the KKT conditions for all firms. Consider:

$$-p_{n,s \in \mathfrak{S}_n} + c_s^G + \beta_s^S - \psi_s^S - \iota_{ni,s \in \mathfrak{S}_n,s \in \mathfrak{S}_i}^S + \xi_{i,s \in \mathfrak{S}_i}^S p_{n,s \in \mathfrak{S}_n} = 0 \perp g_s^S \quad (\text{free}) \quad \forall s \notin \mathfrak{F} \quad (6a)$$

$$-\iota_{ni,s \in \mathfrak{S}_n}^S + \beta_{si}^{SF} - \psi_{si}^{SF} + \xi_i^S (c_s^G + \beta_s) = 0 \perp g_s^F \quad (\text{free}) \quad \forall s \in \mathfrak{F}, i \quad (6b)$$

$$\begin{aligned} & - \sum_{s \in (\mathfrak{S}_n \cap \mathfrak{S}_i)} g_s^S - \sum_{s \in (\mathfrak{S}_n \cap \mathfrak{F})} \zeta_{si}^S + \eta_{ni}^S + \sum_k \theta_{ki}^S B_{nk} + \\ & + \xi_i^S \frac{1}{b_n} (p_n - a_n - \phi_n) + \xi_i^S \sum_{s \in \mathfrak{S}_n} g_s^S = 0 \perp p_n \quad (\text{free}) \quad \forall n, i \end{aligned} \quad (6c)$$

$$b_n \eta_{ni}^S + \iota_{ni}^S - \phi_{ni}^S - \xi_i^S (a_n - b_n d_n) = 0 \perp d_n \quad (\text{free}) \quad \forall n, i \quad (6d)$$

$$\sum_k B_{kn} \iota_{ki}^S + \sum_l H_{ln} (\bar{\mu}_{li}^S - \underline{\mu}_{li}^S) + \begin{cases} \gamma_i^S & \text{if } n = \hat{n} \\ 0 & \text{else} \end{cases} = 0 \perp \delta_n \quad (\text{free}) \quad \forall n, i \quad (6e)$$

$$\zeta_{si}^S + \xi_i^S \bar{g}_s^F \geq 0 \perp \beta_s \geq 0 \quad \forall s \in \mathfrak{F}, i \quad (6f)$$

$$-\zeta_{si}^S \geq 0 \perp \psi_s \geq 0 \quad \forall s \in \mathfrak{F}, i \quad (6g)$$

$$-\eta_{ni}^S + \xi_i^S \left[\frac{1}{b_n} (\phi_n + a_n - p_n) \right] \geq 0 \perp \phi_n \geq 0 \quad \forall n, i \quad (6h)$$

$$\sum_n \theta_{ni}^S H_{ln} + \xi_i^S (\bar{f}_l + e_l) \geq 0 \perp \bar{\mu}_l \geq 0 \quad \forall l, i \quad (6i)$$

$$- \sum_n \theta_{ni}^S H_{ln} + \xi_i^S (\bar{f}_l + e_l) \geq 0 \perp \underline{\mu}_l \geq 0 \quad \forall l, i \quad (6j)$$

$$\theta_{\hat{n}i}^S = 0 \perp \gamma \quad (\text{free}) \quad \forall i \quad (6k)$$

$$\bar{g}_s^S - g_s^S \geq 0 \perp \beta_s^S \geq 0 \quad \forall s \notin \mathfrak{F} \quad (6l)$$

$$g_s^S \geq 0 \perp \psi_s^S \geq 0 \quad \forall s \notin \mathfrak{F} \quad (6m)$$

$$c_s^G - p_{n,s \in \mathfrak{S}_n} + \beta_s - \psi_s = 0 \perp \zeta_{si}^S \quad (\text{free}) \quad \forall s \in \mathfrak{F}, i \quad (6n)$$

$$-a_n + b_n d_n + p_n - \phi_n = 0 \perp \eta_{ni}^S \quad (\text{free}) \quad \forall n, i \quad (6o)$$

$$\sum_k B_{kn} p_k + \sum_l H_{ln} (\bar{\mu}_l - \underline{\mu}_l) + \begin{cases} \gamma & \text{if } n = \hat{n} \\ 0 & \text{else} \end{cases} = 0 \perp \theta_{ni}^S \quad (\text{free}) \quad \forall n, i \quad (6p)$$

$$- \sum_{s \in \mathfrak{S}_n} (g_s^S + g_s^F) + \sum_k B_{nk} \delta_k + d_n = 0 \perp \iota_{ni}^S \quad (\text{free}) \quad \forall n, i \quad (6q)$$

$$\bar{f}_l + e_l - \sum_k H_{lk} \delta_k \geq 0 \perp \bar{\mu}_{li}^S \geq 0 \quad \forall l, i \quad (6r)$$

$$\bar{f}_l + e_l + \sum_k H_{lk} \delta_k \geq 0 \perp \underline{\mu}_{li}^S \geq 0 \quad \forall l, i \quad (6s)$$

$$\bar{g}_s^F - g_s^F \geq 0 \perp \beta_{si}^{SF} \geq 0 \quad \forall s \in \mathfrak{F}, i \quad (6t)$$

$$g_s^F \geq 0 \perp \psi_{si}^{SF} \geq 0 \quad \forall s \in \mathfrak{F}, i \quad (6u)$$

$$d_n \geq 0 \perp \phi_{ni}^S \geq 0 \quad \forall n, i \quad (6v)$$

$$\delta_{\hat{n}} = 0 \perp \gamma_i^S \quad (\text{free}) \quad \forall i \quad (6w)$$

$$\begin{aligned} & \sum_{n \in \mathfrak{N}} \left[\left(a_n - \frac{1}{2} b_n d_n \right) d_n \right] - \sum_{s \in \mathfrak{F}} c_s^G g_s^F - \frac{1}{2} \sum_{n \in \mathfrak{S}_n} \frac{1}{b_n} [a_n - p_n + \phi_n]^2 + \\ & - \sum_n p_n \left(\sum_{s \in \mathfrak{S}_n} g_s^S \right) - \sum_l (\bar{f}_l + e_l) (\bar{\mu}_l + \underline{\mu}_l) - \sum_{s \in \mathfrak{F}} \beta_s \bar{g}_s^F \geq 0 \perp \xi_i^S \quad \forall i \end{aligned} \quad (6x)$$

The bilinearities in (6x) are nonconvex and inconvenient to deal with. Recalling the equivalence between strong duality and the ISO's optimality conditions, we replace this constraint by the according first order conditions (2a) - (2j).¹³

$$c_s^G - p_{n,s \in \mathfrak{S}_n} + \beta_s - \psi_s = 0 \perp g_s^F \quad (\text{free}) \quad \forall s \in \mathfrak{F} \quad (7a)$$

$$-a_n + b_n d_n + p_n - \phi_n = 0 \perp d_n \quad (\text{free}) \quad \forall n \quad (7b)$$

$$\sum_k B_{kn} p_k + \sum_l H_{ln} (\bar{\mu}_l - \underline{\mu}_l) + \begin{cases} \gamma & \text{if } n = \hat{n} \\ 0 & \text{else} \end{cases} = 0 \perp \delta_n \quad (\text{free}) \quad \forall n \quad (7c)$$

$$\sum_{s \in \mathfrak{S}_n} (g_s^S + g_s^F) - \sum_k B_{nk} \delta_k - d_n = 0 \perp p_n \quad (\text{free}) \quad \forall n \quad (7d)$$

$$\bar{f}_l + e_l - \sum_k H_{lk} \delta_k \geq 0 \perp \bar{\mu}_l \geq 0 \quad \forall l \quad (7e)$$

$$\bar{f}_l + e_l + \sum_k H_{lk} \delta_k \leq 0 \perp \underline{\mu}_l \geq 0 \quad \forall l \quad (7f)$$

$$\bar{g}_s^F - g_s^F \geq 0 \perp \beta_s \geq 0 \quad \forall s \in \mathfrak{F} \quad (7g)$$

$$g_s^F \geq 0 \perp \psi_s \geq 0 \quad \forall s \in \mathfrak{F} \quad (7h)$$

$$d_n \geq 0 \perp \phi_n \geq 0 \quad \forall n \quad (7i)$$

$$\delta_{\hat{n}} = 0 \perp \gamma \quad (\text{free}) \quad (7j)$$

The only elements in (7a) - (7j) it is not accounted for so far are the complementarity conditions. Bringing them in disjunctive constraints formulation (cf. Fortuny-Amat and McCarl, 1981) allows for integrating them into (6a) - (6w). Applying the same procedure to all other instances of complementarity, we have the following necessary optimality conditions for the spot market of stages two and three:

$$(6a) - (6w) \quad (8a)$$

¹³Note that we can ignore the complementarity requirement in (6x) here because the left-hand side cannot be different from zero in any case.

$$\sum_i \left(\zeta_{si}^S + \xi_i^S \bar{g}_s^F \right) \leq r_s^\beta K_s^\beta, \quad \beta_s \leq \left(1 - r_s^\beta \right) K_s^\beta \quad \forall s \in \mathfrak{F} \quad (8b)$$

$$\sum_i \left(-\zeta_{si}^S \right) \leq r_s^\psi K_s^\psi, \quad \psi_s \leq \left(1 - r_s^\psi \right) K_s^\psi \quad \forall s \in \mathfrak{F} \quad (8c)$$

$$\sum_i \left(-\eta_{ni}^S + \xi_i^S \left[\frac{1}{b_n} (\phi_n + a_n - p_n) \right] \right) \leq r_n^\phi K_n^\phi, \quad \phi_n \leq \left(1 - r_n^\phi \right) K_n^\phi \quad \forall n \quad (8d)$$

$$\sum_i \left(\sum_n \theta_{ni}^S H_{ln} + \xi_i^S (\bar{f}_l + e_l) \right) \leq r_l^{\bar{\mu}} K_l^{\bar{\mu}}, \quad \bar{\mu}_l \leq \left(1 - r_l^{\bar{\mu}} \right) K_l^{\bar{\mu}} \quad \forall l \quad (8e)$$

$$\sum_i \left(\sum_n \theta_{ni}^S H_{ln} + \xi_i^S (\bar{f}_l + e_l) \right) \leq r_l^{\underline{\mu}} K_l^{\underline{\mu}}, \quad \underline{\mu}_l \leq \left(1 - r_l^{\underline{\mu}} \right) K_l^{\underline{\mu}} \quad \forall l \quad (8f)$$

$$\bar{g}_s^S - g_s^S \leq r_s^{\beta^S} K_s^{\beta^S}, \quad \beta_s^S \leq \left(1 - r_s^{\beta^S} \right) K_s^{\beta^S} \quad \forall s \notin \mathfrak{F} \quad (8g)$$

$$g_s^S \leq r_s^{\psi^S} K_s^{\psi^S}, \quad \psi_s^S \leq \left(1 - r_s^{\psi^S} \right) K_s^{\psi^S} \quad \forall s \notin \mathfrak{F} \quad (8h)$$

$$\bar{f}_l + e_l - \sum_k H_{lk} \delta_k \leq r_l^{\bar{\mu}^S} K_l^{\bar{\mu}^S}, \quad \sum_i \bar{\mu}_{li}^S + \bar{\mu}_l \leq \left(1 - r_l^{\bar{\mu}^S} \right) K_l^{\bar{\mu}^S} \quad \forall l \quad (8i)$$

$$\bar{f}_l + e_l - \sum_k H_{lk} \delta_k \leq r_l^{\underline{\mu}^S} K_l^{\underline{\mu}^S}, \quad \sum_i \underline{\mu}_{li}^S + \underline{\mu}_l \leq \left(1 - r_l^{\underline{\mu}^S} \right) K_l^{\underline{\mu}^S} \quad \forall l \quad (8j)$$

$$\bar{g}_s^F - g_s^F \leq r_s^{\beta^{SF}} K_s^{\beta^{SF}}, \quad \sum_i \beta_{si}^{SF} + \beta_s \leq \left(1 - r_s^{\beta^{SF}} \right) K_s^{\beta^{SF}} \quad \forall s \in \mathfrak{F} \quad (8k)$$

$$g_s^F \leq r_s^{\psi^{SF}} K_s^{\psi^{SF}}, \quad \sum_i \psi_{si}^{SF} + \psi_s \leq \left(1 - r_s^{\psi^{SF}} \right) K_s^{\psi^{SF}} \quad \forall s \in \mathfrak{F} \quad (8l)$$

$$d_n \leq r_n^{\phi^S} K_n^{\phi^S}, \quad \sum_i \phi_{ni}^S + \phi_n \leq \left(1 - r_n^{\phi^S} \right) K_n^{\phi^S} \quad \forall n \quad (8m)$$

where K together with the according superscripts and subscripts denotes a sufficiently large scalar.

Stage I: Welfare-Optimal Network Expansion

Equations and inequalities (8a) - (8m) capture all vectors fulfilling the necessary conditions for an optimum of the EPEC spot market model. Among those, global and local maxima, minima, and saddle points may be found. The first stage now serves as a selection device, in which a benevolent planner optimizes global welfare over expansion of network capacity, demand and generation costs. Letting c_l^E denote the costs, and e_l the amount of network expansion for line l , consider her welfare-maximization:

$$\min_{e, g, d, \delta, \mathfrak{D}, \mathfrak{D}^S}, - \sum_{n \in \mathfrak{N}} \left[(a_n - \frac{1}{2} b_n d_n) d_n \right] + \sum_{s \in \mathfrak{F}} c_s^G g_s^F + \sum_l c_l^E e_l \quad (9a)$$

$$\text{s.t.} \quad 0 \leq e_l \leq \bar{e}_l \quad \forall l, \quad (9b)$$

$$(8a) - (8m) \quad (9c)$$

where line expansion is bounded between zero and an exogenous maximum \bar{e}_l . Program (9a) - (9c) can be solved as Mixed Integer Quadratic Problem using standard approaches and commercial solvers (Tawarmalani and Sahinidis, 2005).

Post-Solve Checks

Recall that the first stage picks the solution vector out of an admissible space only rendering necessary conditions. Therefore the vector returned by the first-stage program may neither fulfil the sufficiency conditions – that is in fact constitute a saddle point, maximum or local

minimum of the overall problem – nor be incentive compatible – that is be in fact no Nash equilibrium solution. To tackle these issues, we implement a two-step procedure, where step one consists of an algorithm systematically exploring the solution space, and step two checks for incentive compatibility. For step one, we first solve (9a) - (9c) for one candidate solution, and fix the outcomes on generation and line expansion. In order to introduce variation, we then impose the requirement that for the next solution to be found exactly one of these variables has to be sufficiently different from the current solution's value, and solve the program again. To impose this restriction, we extend (9a) - (9c) by:

$$g_s \geq D_s^{g+} (\hat{g}_s + \epsilon) \quad \forall s, \hat{g}_s \quad (10a)$$

$$g_s \leq \bar{g}_s - D_s^{g-} (\bar{g}_s - \hat{g}_s + \epsilon) \quad \forall s, \hat{g}_s \quad (10b)$$

$$e_l \geq D_l^{e+} (\hat{e}_l + \epsilon) \quad \forall l, \hat{e}_l \quad (10c)$$

$$e_l \leq \bar{e}_l - D_l^{e-} (\bar{e}_l - \hat{e}_l + \epsilon) \quad \forall l, \hat{e}_l \quad (10d)$$

$$\sum_s [D_s^{g+} + D_s^{g-}] + \sum_l [D_l^{e+} + D_l^{e-}] \geq 1 \quad (10e)$$

where D are binary variables, and ϵ represents the tolerance distance level for the next solution to be different from all preceding ones. Once a solution is found, we keep it and denote it – with slide abuse of notation – by a tilde for future reference. Inequalities (10a) - (10e) therefore implement that no two solution vectors are alike, but differ at least by tolerance ϵ for at least one variable¹⁴. Iterating over this procedure successively cuts 'holes' into the admissible space and delivers a number of candidate solutions that are ranked according to their welfare objective value.

In the second step, we test for deviation stability: for each candidate solution, we fix all variables except generation for one strategic firm and let her optimize profits to check whether she wants to depart from what was proposed by the actual model. Repeating this procedure for all firms reveals whether a candidate solution is incentive compatible, or must be discarded otherwise.¹⁵

4 Results and Discussion

Teh results demonstrate the prevailing effects of the tradeoff concerning welfare-enhancing but costly network expansion when strategic firms are present. We choose a straightforward three-node network, representing the simplest possible case capturing loop flows.¹⁶ The three nodes are mutually linked by transmission lines with limited capacities of $f_1 = f_2 = 0.5, f_3 = 2.5$ that are subject to potential expansion. At node one, there is linear-elastic demand represented by $P_1 = 10 - q_1$, where q_1 captures the quantity consumed at that node. The other two nodes, at each of which one strategic firm with zero marginal costs is located at, feature no consumption. By disregarding potentially asymmetric production costs, we preclude results that are driven by pure efficiency gains: a welfare increase due to substitution effects from expensive to previously unaccessible cheap production could flaw the stylized assessment of reduced market power rents. For convenience, moreover, we abstract from competitive fringe generation. Consider figure 1 depicting the network.

To frame the problem, we first calculate results for two benchmark cases: 1) *Copperplate*, in which all line capacities are exogenously set sufficiently high to accommodate all potentially emerging flows, and 2) *No expansion*, in which network expansion is exogenously set to zero. Table 2 presents the results:

¹⁴Pozo et al. (2013a) implement a comparable approach

¹⁵A uniqueness discussion for the solutions of this MPEC can be found in the results section ??.

¹⁶For interpretational purposes, each node can selectively be regarded as a national market, where we abstract from internal congestion, and the links as international interconnectors.

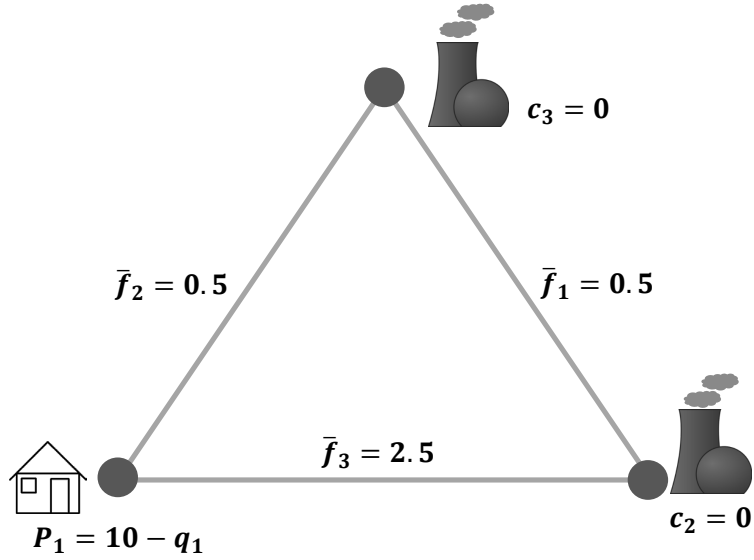


Figure 1: Sample network. Pictograms are under public domain free licence

		1) <i>Copperplate</i>	2) <i>No expansion</i>
Generation	node 2	3.33	1.5
	node 3	3.33	0
Price	node 1	3.33	8.5
	node 2	3.33	8.5
	node 3	3.33	8.5
Demand	node 1	6.67	1.5
Network flows	line 1	0	0.5
	line 2	3.33	0.5
	line 3	-3.33	-1
Consumer rent	total	22.22	1.13
Profit	firm 2	11.11	12.75
	firm 3	11.11	0
Congestion rent	total	0	0
Welfare	total	44.44	13.88

Table 2: Results for the benchmark cases

The *copperplate* results set the upper welfare benchmark: if line restrictions are never binding, the Cournot solution emerges, representing the maximum degree of competition among the two firms, unconstrained by any network limitations. The lower bound is rendered by the *no expansion* results: in that case, firm two behaves aggressively by actively congesting the transmission lines that connect the passive firm three to the demand node. By that strategy, a stable passive-aggressive equilibrium emerges as also detected by Borenstein et al. (2000). From an overall welfare perspective, thus, there is scope for network expansion to push welfare closer to the upper bound. Assuming expansion costs of $c_l^E = 1 \forall l$, stage I of our actual model selects candidates for optimal points, three distinct types of which are rendered in table 3

		<i>Asymmetric</i>	<i>Cournot 2</i>	<i>Cournot 1</i>
Generation	firm 2	1.5	3.33	3.33
	firm 3	4.25	3.33	3.33
Price		4.25	3.33	3.33
Demand	node 1	5.75	6.67	6.67
Network capacity (initial + expansion)	line 1	1.23 (0.5 + 0.73)	0.5 (0.5 + 0)	0.8 (0.5 + 0.3)
	line 2	3.33 (0.5 + 2.83)	3.33 (0.5 + 2.83)	3.33 (0.5 + 2.83)
	line 3	2.5 (2.5 + 0)	3.33 (2.5 + 0.83)	3.33 (2.5 + 0.83)
Total expansion		3.57	3.67	3.97
Network flows	line 1	-0.92	0	0
	line 2	3.33	3.33	3.33
	line 3	-2.42	-3.33	-3.33
Consumer rent	total	16.53	22.22	22.22
Profit	firm 2	6.38	11.11	11.11
	firm 3	18.06	11.11	11.11
Congestion rent	total	0	0	0
Expansion cost	total	3.57	3.67	3.97
Welfare	total	37.40	40.78	40.48

Table 3: Results for the network expansion model

For a low level of network expansion, $\sum_l e_l = 3.57$, a first type of solutions is attained, which we denote by *asymmetric (A1)*. This fosters an equilibrium in which the previously passive firms enters the aggressive position due to now being connected to demand by a line with higher capacity. The character of the equilibrium, however, is of the same passive-aggressive nature as in the *no expansion* benchmark, although overall welfare raises considerably from 13.88 to 37.4. The post-model check for deviation stability does not indicate any deviation incentives for the strategic players.

Further network expansion, then, promotes the emergence of a second, symmetric type of KKT points, which we call – due to the quantities produced by the strategic firms – *Cournot 2 (C2)*, and *Cournot 1 (C1)*. The level of network expansion, $\sum_l e_l = 3.67$ for solution candidate *C2*, renders the lowest level of line capacities exactly necessary to accommodate the prevailing flows in the Cournot solution, see table 3, thereby delivering the theoretically highest amount of social welfare that can be attained in this numerical setup. This KKT point, however, turns out to be not stable against deviations of the strategic firms, and therefore must be discarded. The rationale that prevails here found also reflection in (Pozo et al., 2013a) who discuss it under optimistic versus pessimistic solutions – a label we want to adopt here. Holding firm 3’s generation fixed at $g_3^S = 3.33$, the post-solve MPEC check

delivers the result summarized in table 4¹⁷:

	Firm 2	Firm 3	n_1	n_2	n_3	l_1	l_2	l_3	Total
Generation	1.83	3.33							
Demand			5.37	0	0				
Price			4.83	9.67	0				
Network capacity						0.5	3.33	3.33	
Network flows						-0.5	2.83	-2.83	
Consumer rent			13.35						13.35
Profit	17.72	0							17.72
Congestion rent									10.91
Expansion costs						0	2.83	0.83	3.67
Welfare									38.32

Table 4: Post-solve MPEC check: Results for a deviation by firm 2

Firm 2 deviates by lowering its generation to $g_2^S = 1.83$, and increases its profits from 11.11 to 17.22 whereas the other firm 3 yields zero profits. This outcome is induced by the emerging price pattern with $p_2 = 9.67$ and $p_3 = 0$. The re-optimizing player, thus, picks a generation level such that an equilibrium reached that is most profitable to him. This mechanism is a very feature of the MPEC: recall that the ISO's equilibrium conditions on model stage III merely describe a set of stationary points, and admit multiple equilibria. If all decision variables except generation of one firm are fixed, then this firm will pick the equilibrium most profitable to her. In this case, lowering production, evoking congestion and profiting from the resulting high price at her node. Insofar the label optimistic applies – hypothetically, the ISO could also behave in a way such that the worst outcome would emerge, for example when attaching nonequal weights to the components of welfare¹⁸. A solution that is stable under the optimistic assumption, however, is not necessarily stable under the pessimistic one. This most favourable setup for firms, thus, triggers a conservative perspective on our research question. If no market power can be exerted under the optimistic assumption, then this holds true for all other setups. For a stable solution, line 1 as well has to be expanded, yielding a total network expansion of $\sum_l e_l = 3.97$, as proposed by solution *C1*, which is stable against deviations. Two effects are at work here: firstly, line l_1 has sufficient capacity such that a unilateral reduction of generation does not evoke a congested equilibrium with substantially differentiated prices, as in *C2*, any more. Secondly, the *thin line*-effect (Borenstein et al., 2000): although connection line l_2 between the two firms is not actually used at all, compare the respective entries for flows in table 3, its expansion is required so that each firm can credibly maintain the threat of harsh competition to keep the Cournot equilibrium stable.

We now turn to the distributional implications. The bars in 2 indicate the outcomes for the three model cases. Comparing the *no expansion* benchmark them with the asymmetric equilibrium *A1* illustrates that the considerable welfare increase goes along with higher producer and consumer rents, where the latter profit to a higher extent. For the stable Cournot equilibrium *C1* total welfare increases further as already discussed. The distribution of rents, however, reveals an interesting effect: the consumer rent increases whereas the producer rent decreases. We thus detect a shift of rents from suppliers towards the demand side, underlining the decreased potential of firms to extract rents due to their dominant position.

The effect of network expansion on welfare when strategic firms are present, therefore, is twofold: on the one hand, overall rents increase in the availability of transmission resources,

¹⁷Due to symmetry, holding firm 2's generation fixed and letting firm 3 re-optimize delivers a parallel result.

¹⁸In a similar formal setup, Huppmann and Egerer (2014) discuss the issue of weighting the components of the welfare objective unequally.

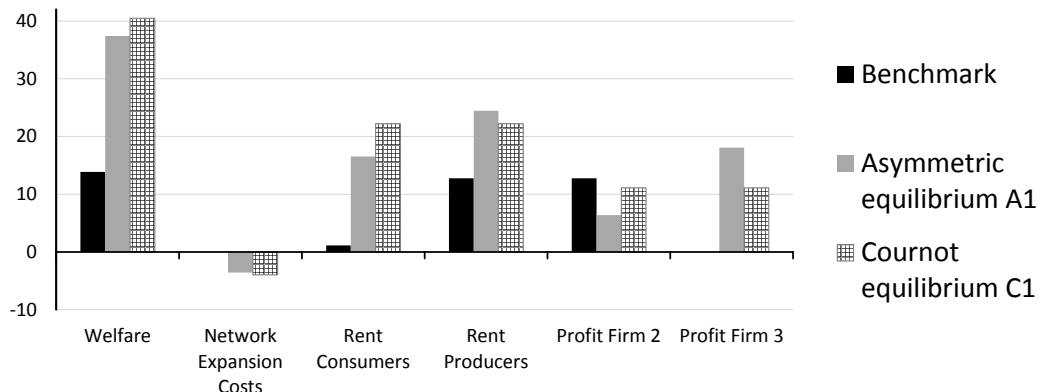


Figure 2: Distribution of rents

where consumers and producers profit. On the other hand, at some point, however, the network features enough capacity to accommodate the Cournot solution, the highest degree of competition in this sample case. Along with the generation of additional rents, there is a redistribution of from generators towards consumers – network expansion, thus, has the potential to limit the exertion of market power, and prevents firms from extracting rents that arise due to limited transmission capacity

5 Conclusion & Outlook

Insufficient transmission grid capacities may impede reaping the full benefits of a competitive electricity market. To assess whether and to which extent costly network expansion promotes social welfare due to enhanced competition, we set up a three-stage model mimicking the interplay of electricity grid expansion, and strategic generation: on its third stage, an ISO dispatches competitive fringe plants and ensures feasible flows. Reformulating her KKT conditions with help of duality theory yields the solution space for the profit-maximization programs of strategic firms engaging in Cournot competition. The first-order KKT conditions of this EPEC capture the set of all vectors fulfilling the necessary conditions for an optimum of the spot market game. Finally, to pick the welfare-optimal outcome out of these, on the first stage a benevolent planner maximizes total welfare over network expansion decisions. To ensure sufficiency and incentive compatibility, we implement a two-step algorithm exploring the solution space in an iterative procedure, and checking deviation incentives for the strategic players.

Numerical results for a basic three-node network illustrate the strategic effects: firstly, network expansion has the potential to mitigate market power by banning firms from selecting into equilibria that congest the transmission lines, and thereby foster competition. Secondly, higher levels of transmission grid expansion prevent asymmetric passive-aggressive equilibria as described by Borenstein et al. (2000), and induce a relative shift of rents towards the demand side. Stability of the suggested solutions is tested assuming an *optimistic* perspective, putting a natural lower bound on the behaviour of strategic firms within their feasible sets: if they assume that the ISO will always act to the best of their interests and still do not find an incentive to deviate from a suggested solution, they cannot find such an incentive in case the ISO “works against them”.

In this contribution, we lay the theoretical foundations of a model that can be suited to larger applied work, opening up several avenues for future research in this field: for instance, an application of the model to a representation of the European electricity system based on the aggregation of realistic data will be able to assess benefits of further electricity

market integration in terms of the welfare-enhancing effects of more vigorous competition. A further elaboration on the distributional implications can, moreover, allow for an analysis of gains and losses from integration, which is, in turn, connected to the question of incentive compatibility of integration among sovereign states.¹⁹

References

- ACER/CEER. Annual Report on the Results of Monitoring the Internal Electricity and Natural Gas Markets in 2012, 2013. URL http://www.acer.europa.eu/Official_documents/Acts_of_the_Agency/Publication/ACER%20Market%20Monitoring%20Report%202013.pdf.
- V. Böckers, J. Haucap, and U. Heimeshoff. Benefits of an Integrated European Electricity Market. *DICE Discussion Paper*, 109, 2013.
- Booz & Company, D. Newbery, G. Strbac, D. Pudjianto, P. Noel, and LeighFisher. Benefits of an Integrated European Energy Market. Prepared for: Directorate-General Energy, European Commission, 2013. URL http://ec.europa.eu/energy/infrastructure/studies/doc/20130902_energy_integration_benefits.pdf.
- S. Borenstein, J. Bushnell, and S. Stoft. The Competitive Effects of Transmission Capacity in a Deregulated Electricity Industry. *The RAND Journal of Economics*, 31(2):294–325, Summer 2000.
- S. P. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.
- P. Cramton, A. Ockenfels, and S. Stoft. Capacity Market Fundamentals. *Economics of Energy & Environmental Policy*, 2(2):27–46, 2013.
- L.B. Cunningham, R. Baldick, and M.L. Baughman. An Empirical Study of Applied Game Theory: Transmission Constrained Cournot Behavior. *IEEE Transactions on Power Systems*, 17(1):166–172, 2002.
- EPEX Spot. Social Welfare Report 01-12/2013, 12 2013. URL https://www.epexspot.com/document/25315/Social%20Welfare%20Report_01-12%202013.pdf.
- European Commission. Making the internal energy market work. Communication (2012) 663 final, November 2012a.
- European Commission. *Energy, transport and environment Indicators 2012 edition*. Publications Office of the European Union, 2012b.
- J. Fortuny-Amat and B. McCarl. A Representation and Economic Interpretation of a Two-Level Programming Problem. *The Journal of the Operational Research Society*, 32(9):783–792, 1981.
- S. A. Gabriel and F. U. Leuthold. Solving discretely-constrained MPEC problems with applications in electric power markets. *Energy Economics*, 32:3–14, 2010.
- S. A. Gabriel, S. A. Siddiqui, A. J. Conejo, and C. Ruiz. Solving Discretely-Constrained Nash-Cournot Games with an Application to Power Markets. *Networks and Spatial Economics*, 13(3):307–326, 2013.
- G. Gebhardt and F. Höffler. How Competitive is Cross-border Trade of Electricity? Theory and Evidence from European Electricity Markets. *The Energy Journal*, 34(1):125–154, 2013.

¹⁹Neglecting strategic generation, this issue is pursued in Huppmann and Egerer (2014) using a similar modeling framework.

- B.F. Hobbs, C.B. Metzler, and J.-S. Pang. Strategic Gaming Analysis for Electric Power Systems: An MPEC Approach. *IEEE Transactions on Power Systems*, 15(2):638–645, 2000.
- B.F. Hobbs, F.A.M. Rijkers, and M.G. Boots. The More Cooperation, The More Competition? A Cournot Analysis of the Benefits of Electric Market Coupling. *The Energy Journal*, 26(4):69–98, 2005.
- D. Huppmann and J Egerer. National-strategic investment in European cross-border transmission capacity. *DIW Discussion Paper*, forthcoming, 2014.
- D. M. Kreps and J. A. Scheinkman. Quantity Precommitment and Bertrand Competition Yield Cournot Outcomes. *The Bell Journal of Economics*, 14(2):326–337, 1983.
- Th.-O. Léautier. Fred Scheppe meets Marcel Boiteux and Antoine-Augustin Cournot: transmission constraints and strategic underinvestment in electric power generation. *Toulouse School of Economics Working Paper*, 432, 2013.
- F.U. Leuthold, H. Weigt, and C. v. Hirschhausen. A Large-Scale Spatial Optimization Model of the European Electricity Market. *Networks and Spatial Economics*, 12(1):75–107, 2012.
- K. Neuhoff, J. Barquin, M. G. Boots, A. Ehrenmann, B. F. Hobbs, F. A. M Rijkers, and M. Vazquez. Network-constrained Cournot models of liberalized electricity markets: the devil is in the details. *Energy Economics*, 27:495–525, 2005.
- G. Oggioni, Y. Smers, E. Allevi, and S. Schaible. A Generalized Nash Equilibrium Model of Market Coupling in the European Power System. *Networks and Spatial Economics*, 12(4):503–560, 2012.
- D. Pozo, J. Contreras, and E. Sauma. If you build it, he will come: Anticipative power transmission planning. *Energy Economics*, 36:135–146, 2013a.
- D. Pozo, E. E. Sauma, and J. Contreras. A Three-Level Static MILP Model for Generation and Transmission Expansion Planning. *IEEE Transactions on Power Systems*, 28(1):202–210, 2013b.
- D. Ruderer and G. Zöttl. The Impact of Transmission Pricing in Network Industries. *EPRG Working Paper*, (1214), 2012.
- C. Ruiz, Antonio J. Conejo, and Yves Smeers. Equilibria in an Oligopolistic Electricity Pool With Stepwise Offer Curves. *IEEE Transactions on Power Systems*, 27(2):752–761, 2012.
- N. Ryan. The Competitive Effects of Transmission Infrastructure in the Indian Electricity Market. *Working Paper*, 2013. URL <http://economics.mit.edu/files/9302>.
- E. E. Sauma and S. S. Oren. Proactive planning and valuation of transmission investments in restructured electricity markets. *Journal of Regulatory Economics*, 30:358–387, 2006.
- F.C. Scheppe, M.C. Caramanis, R.D. Tabors, and R.E. Bohn. *Spot Pricing of Electricity*. Kluwer, Boston, 1988.
- S. Siddiqui and S. A. Gabriel. An SOS1-Based Approach for Solving MPECs with a Natural Gas Market Application. *Networks and Spatial Economics*, 13(2):205–227, 2013.
- M. Tanaka. Transmission-constrained oligopoly in the Japanese electricity market. *Energy Economics*, 31:690–701, 2009.
- M. Tawarmalani and N. V. Sahinidis. A polyhedral branch-and-cut approach to global optimization. *Mathematical Programming*, 103(2):225–249, 2005.

- B. Willems. Model Cournot Competition in an Electricity Market with Transmission Constraints. *The Energy Journal*, 23(3):95–125, 2002.
- F. A. Wolak. Measuring the Competitive Benefits of a Transmission Investment Policy: The Case of The Alberta Electricity Market. *Working Paper*, 2012. URL http://www.stanford.edu/group/fwolak/cgi-bin/sites/default/files/files/alberta_transmission_benefits_aug_2012.pdf.
- G. Zachmann. Electricity wholesale market prices in Europe: Convergence? *Energy Economics*, 30:1659–1671, 2008.
- G. Zöttl. On optimal scarcity prices. *International Journal of Industrial Organization*, 29:589–605, 2011.

A Appendix

A.1 Proof of Proposition 1

First, we derive the Lagrangian dual problem for a quadratic problem with a positive definite matrix in a general form:

$$\min_{x,y} \quad \frac{1}{2}x^T D_x x + c^T \begin{bmatrix} x \\ y \end{bmatrix} \quad \text{s.t.} \quad A \begin{bmatrix} x \\ y \end{bmatrix} \leq b \quad (\lambda), \quad B \begin{bmatrix} x \\ y \end{bmatrix} = e \quad (\nu) \quad (11)$$

where x is the vector of variables that enter the quadratic part of the objective function, and y is the vector of variables that only appear in linear terms. By assumption, D_x is symmetric and strictly positive definite.

The Lagrangian (primal) function is then:

$$\mathcal{L}(x, y, \lambda, \nu) = \frac{1}{2}x^T D_x x + c^T \begin{bmatrix} x \\ y \end{bmatrix} + \left(A \begin{bmatrix} x \\ y \end{bmatrix} - b \right)^T \lambda + \left(B \begin{bmatrix} x \\ y \end{bmatrix} - e \right)^T \nu \quad (12)$$

Lemma A1

The Lagrangian dual function for (12) is given by

$$\hat{\mathcal{L}}(\lambda, \nu, x, y) = -\frac{1}{2} \left(c_x + A_x^T \lambda + B_x^T \nu \right)^T (D_x)^{-1} \left(c_x + A_x^T \lambda + B_x^T \nu \right) - b^T \lambda - e^T \nu \quad (13a)$$

$$\text{where } D_x x + c_x + (A_x)^T \lambda + (B_x)^T \nu = 0, \quad (13b)$$

$$c_y + (A_y)^T \lambda + (B_y)^T \nu = 0, \quad (13c)$$

where subscripts indicate that parts of matrices that are multiplied with the respective variable vector.

Proof

The Lagrange dual function is defined as the infimum of the Lagrange primal function over the primal decision variables²⁰. Differentiating (12) with respect to x and y yields

$$\frac{\partial \mathcal{L}}{\partial x} = D_x x + c_x + (A_x)^T \lambda + (B_x)^T \nu = 0 \quad (14a)$$

$$\Rightarrow x^* = -(D_x)^{-1} \left(c_x + (A_x)^T \lambda + B_x^T \nu \right) \quad (14b)$$

$$\frac{\partial \mathcal{L}}{\partial y} = c_y + (A_y)^T \lambda + B_y^T \nu = 0 \quad (14c)$$

Note that \mathcal{L} is strictly convex in x and y such that the first order equality conditions are necessary and sufficient to ensure a global minimum. Plugging optimality condition (14b) into the (primal) Lagrangian (12)

$$\begin{aligned} \mathcal{L} &= \frac{1}{2} \left[\left[(-D_x^{-1})(c_x + A_x^T \lambda + B_x^T \nu) \right]^T D_x \left[(-D_x^{-1})(c_x + A_x^T \lambda + B_x^T \nu) \right] \right] \\ &\quad + c_x^T \left[(-D_x^{-1})(c_x + A_x^T \lambda + B_x^T \nu) \right] + c_y^T y \\ &\quad + \left[A_x \left[(-D_x^{-1})(c_x + A_x^T \lambda + B_x^T \nu) \right] + A_y y - b \right]^T \lambda \\ &\quad + \left[B_x \left[(-D_x^{-1})(c_x + A_x^T \lambda + B_x^T \nu) \right] + B_y y - e \right]^T \nu \\ &= \frac{1}{2} \left[(c_x + A_x^T \lambda + B_x^T \nu)^T D_x^{-1} (c_x + A_x^T \lambda + B_x^T \nu) \right] \\ &\quad - \left[c_x^T D_x^{-1} (c_x + A_x^T \lambda + B_x^T \nu) + (c_x + A_x^T \lambda + B_x^T \nu)^T D_x^{-1} A_x^T \lambda + (c_x + A_x^T \lambda + B_x^T \nu)^T D_x^{-1} B_x^T \nu \right] \end{aligned} \quad (14d)$$

²⁰Consult Boyd and Vandenberghe (2004, chapter 5) for a theoretical exhibition.

$$+ c_y^T y + y^T A_y^T \lambda + y^T B_y^T \nu - b^T \lambda - e^T \nu \quad (14e)$$

where we have made use of $(D_x^{-1})^T = D_x^{-1}$ due to symmetry of D_x . Now observe that from (14c) $A_y^T \lambda + B_y^T \nu = -c_y$ and obtain

$$\begin{aligned} \mathcal{L} = & \frac{1}{2} \left[(c_x + A_x^T \lambda + B_x^T \nu)^T D_x^{-1} (c_x + A_x^T \lambda + B_x^T \nu) \right] \\ & - \left[c_x^T D_x^{-1} (c_x + A_x^T \lambda + B_x^T \nu) + (c_x + A_x^T \lambda + B_x^T \nu)^T D_x^{-1} A_x^T \lambda + (c_x + A_x^T \lambda + B_x^T \nu)^T D_x^{-1} B_x^T \nu \right] \\ & + c_y^T y - y^T c_y - b^T \lambda - e^T \nu \end{aligned} \quad (14f)$$

The inner products $c_y^T y$ and $y^T c_y$ in the last row of (14f) are of dimension (1×1) , thus necessarily symmetric and cancel out. Moreover, straightforward algebra yields that the two terms in brackets in the first and second line of (14f) are identical. Therefore,

$$\hat{\mathcal{L}}(\lambda, \nu) = -\frac{1}{2} \left[(c_x + A_x^T \lambda + B_x^T \nu)^T D_x^{-1} (c_x + A_x^T \lambda + B_x^T \nu) \right] - b^T \lambda - e^T \nu \quad (14g)$$

■

Now, we can set up the dual Lagrangian problem, consisting in maximizing L^* over λ and ν , while optimality conditions concerning x and y are added as constraints. By construction, further constraints comprise weak positivity of decision variables attached to inequality constraints

$$\max_{\lambda, \nu} \quad -\frac{1}{2} \left[(c_x + A_x^T \lambda + B_x^T \nu)^T D_x^{-1} (c_x + A_x^T \lambda + B_x^T \nu) \right] - b^T \lambda - e^T \nu \quad (15a)$$

$$\text{s.t.} \quad D_x x + c_x + (A_x)^T \lambda + (B_x)^T \nu = 0, \quad (15b)$$

$$c_y + (A_y)^T \lambda + (B_y)^T \nu = 0 \quad (15c)$$

$$\lambda \geq 0, \nu \in \mathbb{R} \quad (15d)$$

Next, we apply this result to our problem. Rewriting the primal problem (1a) - (1h) in matrix notation in the fashion of (11) yields for the objective function:

$$\min_{g^F, d, \delta, \mathfrak{D}} - \sum_{n \in \mathfrak{N}} \left[(a_n - \frac{1}{2} b_n d_n) d_n \right] - \sum_{s \in \mathfrak{F}} c_s^G g_s^F \hat{=} \min_{g^F, d, \delta, \mathfrak{D}} \frac{1}{2} d^T \Delta d + \gamma^T \begin{bmatrix} d \\ g^F \\ \delta \end{bmatrix} \quad (16a)$$

where the variable vectors are composed as follows:

$$d_{(\tilde{N} \times 1)} = \begin{bmatrix} d_1 \\ \vdots \\ d_{\tilde{N}} \end{bmatrix}, \quad g^F_{(\tilde{F} \times 1)} = \begin{bmatrix} g_1^F \\ \vdots \\ g_{\tilde{F}}^F \end{bmatrix}, \quad \delta_{(N \times 1)} = \begin{bmatrix} \delta_1 \\ \vdots \\ \delta_N \end{bmatrix} \quad (16b)$$

and the parameter matrices as²¹:

$$\Delta_{(\tilde{N} \times \tilde{N})} = \begin{bmatrix} b_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & b_{\tilde{N}} \end{bmatrix}, \quad \gamma_{((\tilde{N} + \tilde{F} + N) \times 1)} = [-a_1 \dots -a_{\tilde{N}} \quad c_1^G \dots c_{\tilde{F}}^G \quad 0 \dots 0]^T \quad (16c)$$

For convenience, let $|\mathfrak{N}| = \tilde{N}$, and $|\mathfrak{F}| = \tilde{F}$ render the number of nodes at which demand is located, and number of fringe plants respectively. Moreover, observe that Δ fulfils our

²¹We assume for convenience here that vector g^F is nonempty, i.e. that there are \tilde{F} fringe suppliers, beginning with plant $s = 1$. The results can easily be reduced to the case without fringe, i.e. $\tilde{F} = 0$. Moreover, formally excluding nodes with no demand makes matrix Δ strictly positive definite and allows inverting it, which is necessary for Lemma A1 to hold and to follow our proposed procedure.

assumptions of symmetry and strict positive definiteness. Expressing constraints (1b) - (1h) in matrix notation:

$$A \begin{bmatrix} d \\ g^F \\ \delta \end{bmatrix} \leq b \quad (\lambda) \quad \hat{=} \quad [A_d : A_{g^F \delta}] \begin{bmatrix} d \\ g^F \\ \delta \end{bmatrix} \leq b \quad (\lambda), \quad (16d)$$

$$B \begin{bmatrix} d \\ g^F \\ \delta \end{bmatrix} = e \quad (\nu) \quad \hat{=} \quad [B_d : B_{g^F \delta}] \begin{bmatrix} d \\ g^F \\ \delta \end{bmatrix} = e \quad (\nu) \quad (16e)$$

with submatrices $A_d, A_{g^F \delta}, B_d, B_{g^F \delta}$ capturing the relevant entries for the respective constraints, and accordingly composed vectors b, e, λ , and ν . We now just have to plug the matrices and vectors into the formulation of the dual Lagrangian problem (15a) - (15d). Straightforward algebra yields for the expression $(c_x + A_x^T \lambda + B_x^T \nu)$:

$$[c_x + A_x^T \lambda + B_x^T \nu] \hat{=} \left[\begin{bmatrix} -a_1 \\ \vdots \\ -a_{\tilde{N}} \end{bmatrix} + A_d^T \lambda + B_d^T \nu \right] = \begin{bmatrix} -a_1 + p_1 - \phi_1 \\ \vdots \\ -a_{\tilde{N}} + p_{\tilde{N}} - \phi_{\tilde{N}} \end{bmatrix} \quad (17a)$$

obviously rendering the first term of the dual objective:

$$\begin{aligned} & -\frac{1}{2} [(c_x + A_x^T \lambda + B_x^T \nu)^T D_x^{-1} (c_x + A_x^T \lambda + B_x^T \nu)] \\ & \hat{=} -\frac{1}{2} \left[\begin{bmatrix} -a_1 + p_1 - \phi_1 \\ \vdots \\ -a_{\tilde{N}} + p_{\tilde{N}} - \phi_{\tilde{N}} \end{bmatrix} \right]^T \Delta^{-1} \left[\begin{bmatrix} -a_1 + p_1 - \phi_1 \\ \vdots \\ -a_{\tilde{N}} + p_{\tilde{N}} - \phi_{\tilde{N}} \end{bmatrix} \right] \\ & \hat{=} -\frac{1}{2} \sum_{n \in \mathfrak{N}} \frac{1}{b_n} (a_n - p_n + \phi_n)^2 \end{aligned} \quad (17b)$$

For the second term of (15a) it immediately follows:

$$b^T \lambda + e^T \nu = \sum_l (\bar{f}_l + e_l) (\bar{\mu}_l + \underline{\mu}_l) + \sum_n p_n \left(\sum_{s \in \mathfrak{S}_n} g_s^S \right) + \sum_{s \in \mathfrak{F}} \beta_s \bar{g}_s^F \quad (17c)$$

Bringing (17b) and (17c) together thus yields the dual objective. The constraint set of the dual is abstractly given by:

$$D_x x + c_x + (A_x)^T \lambda + (B_x)^T \nu \hat{=} \Delta d - [a_1 \dots a_{\tilde{N}}] + (A_d)^T \lambda + (B_d)^T \nu = 0, \quad (17d)$$

$$c_y + A_y^T \lambda + B_y^T \nu \hat{=} [c_1^G \dots c_{\tilde{F}}^G \quad 0 \dots 0]^T + A_{g^F \delta}^T \lambda + B_{g^F \delta}^T \nu = 0, \quad (17e)$$

cf. (15b) - (15c). Plugging in the relevant matrices reproduces KKT conditions (2a) - (2c) of the original problem:

$$c_s^G - p_{n, s \in \mathfrak{S}_n} + \beta_s - \psi_s = 0 \quad \forall s \in \mathfrak{F} \quad (17f)$$

$$-a_n + b_n d_n + p_n - \phi_n = 0 \quad \forall n \in \mathfrak{N} \quad (17g)$$

$$\sum_k B_{kn} p_n + \sum_l H_{ln} (\bar{\mu}_l - \underline{\mu}_l) + \begin{cases} \gamma & \text{if } n = \hat{n} \\ 0 & \text{else} \end{cases} = 0 \quad \forall n \quad (17h)$$

Therefore, the equivalent Lagrangian dual problem to (1a) - (1h) is given by

$$\begin{aligned} \max_{p, \bar{\mu}, \underline{\mu}, \beta, \psi, \phi, \gamma} & -\frac{1}{2} \sum_{n \in \mathfrak{N}} \frac{1}{b_n} [a_n - p_n + \phi_n]^2 - \sum_l (\bar{f}_l + e_l) (\bar{\mu}_l + \underline{\mu}_l) \\ & - \sum_n p_n \left(\sum_{s \in \mathfrak{S}_n} g_s^S \right) - \sum_{s \in \mathfrak{F}} \beta_s \bar{g}_s^F \end{aligned} \quad (18)$$

$$\begin{aligned} \text{s.t.} \quad & (17\text{f}), (17\text{g}), (17\text{h}) & (19) \\ & \bar{\mu}, \underline{\mu}, \beta, \psi, \phi \geq 0, p, \gamma \in \mathbb{R} & (20) \end{aligned}$$

■