THE SHORT-RUN APPROACH TO LONG-RUN EQUILIBRIUM: A GENERAL THEORY WITH APPLICATIONS

CDAM RESEARCH REPORT LSE-CDAM-2005-02

ANTHONY HORSLEY AND ANDREW J. WROBEL

ABSTRACT. This is a new formal framework for the theory of competitive equilibrium and its applications. Our "short-run approach" means the calculation of long-run producer optima and general equilibria from the short-run solutions to the producer's profit maximization programme and its dual. The marginal interpretation of the dual solution means that it can be used to value the capital and other fixed inputs, whose levels are then adjusted accordingly (where possible). But short-run profit can be a nondifferentiable function of the fixed quantities, and the short-run cost is nondifferentiable whenever there is a rigid capacity constraint. Nondifferentiability of the optimal value requires the introduction of nonsmooth calculus into equilibrium analysis, and subdifferential generalizations of smooth-calculus results of microeconomics are given, including the key Wong-Viner Envelope Theorem. This resolves long-standing discrepancies between "textbook theory" and industrial experience. The other tool employed to characterise long-run producer optima is a primal-dual pair of programmes. Both marginalist and programming characterizations of producer optima are given in a taxonomy of seventeen equivalent systems of conditions. When the technology is described by production sets, the most useful system for the short-run approach is that using the short-run profit programme and its dual. This programme pair is employed to set up a formal framework for long-run general-equilibrium pricing of a range of commodities with joint costs of production. This gives a practical method that finds the short-run general equilibrium en route to the long-run equilibrium, exploiting the operating policies and plant valuations that must be determined anyway. These critical short-run solutions have relatively simple forms that can greatly ease the fixed-point problem of solving for equilibrium, as is shown on an electricity pricing example. Applicable criteria are given for the existence of the short-run solutions and for the absence of a duality gap. The general analysis is spelt out for technologies with conditionally fixed coefficients, a concept extending that of the fixed-coefficients production function to the case of multiple outputs. The short-run approach is applied to the peak-load pricing of electricity generated by thermal, hydro and pumped-storage plants. This gives, for the first time, a sound method of valuing the fixed assets—in this case, river flows and the sites suitable for reservoirs.

ADDRESS: Department of Economics, London School of Economics, Houghton Street, London WC2A 2AE, United Kingdom.
E-MAIL: LSEecon123@mac.com

Date: 14 February 2005 (revised 4 May 2005).

2000 *Mathematics Subject Classification.* Primary 91B50; Secondary 90B05, 90C25.

1991 *Journal of Economic Literature Classification*: C61, D24, D46, D58, L94.

Key words and phrases. general equilibrium, fixed-input valuation, nondifferentiable joint costs,Wong-Viner Envelope Theorem, public utility pricing.

This work was started on ESRC grant R000232822; their financial support is gratefully acknowledged.

1. Introduction

This is a new formal framework for the theory of competitive equilibrium and its applications. Our "short-run approach" is a scheme for calculating long-run producer optima and general equilibria by building on short-run solutions to the producer's profit maximization problem, in which the capital inputs and natural resources are treated as fixed. These fixed inputs are valued at their marginal contributions to the operating profits and, where possible, their levels are then adjusted accordingly.[1] Since short-run profit is a concave but generally nondifferentiable function of the fixed inputs, their marginal values are defined as nonunique supergradient vectors. Also, they usually have to be obtained as solutions to the dual programme of fixed-input valuation because there is rarely an explicit formula for the operating profit. Thus the key property of the dual solution is its marginal interpretation, but this requires the use of a generalized, multi-valued derivative—viz., the subdifferential—because an optimal-value function, such as profit or cost, is commonly nondifferentiable.

Differential calculus is essential for applications, but it has been purged from geometric treatments of the Arrow-Debreu model, which are limited to equilibrium existence and Pareto optimality results. Our use of subgradients rehabilitates it as a rigorous method for equilibrium theory. The mathematical tools we employ—convex programming and subdifferential calculus—enable us to reformulate some basic microeconomic results. In addition to stating the known subdifferential versions of the Shephard-Hotelling Lemmas, we have devised a subdifferential version of the Wong-Viner Envelope Theorem for the short-run approach especially (Section 11). This facilitates economic analysis and resolves long-standing discrepancies between "textbook theory" and industrial experience.[2]

We use these methods to set up a framework for the general-equilibrium pricing of multiple outputs with joint production costs. This is applied to the pricing, operation and investment problems of an electricity supply industry with a technology that can include hydroelectric generation and pumped storage of energy, in addition to thermal generation (Sections 15 to 17). This application draws on the much simpler case of purely thermal generation (Section 2) and on our studies of the operation and valuation of hydroelectric and pumped-storage plants in [24] and [27]. Those results are summarized and "fed into" the short-run approach.

The short-run approach starts with fixing the producer's capacities k and optimizing the variable quantities, viz., the outputs y and the variable inputs v. For a competitive, price-taking producer, the optimum quantities, $\hat{y}$ and $\hat{v}$, depend on their given prices, p and w, as well as on k.[3] The primal solution ($\hat{y}$ and $\hat{v}$) is associated with the dual solution $\hat{r}$, which gives the unit values of the fixed inputs (with $\hat{r} \cdot k$ as their total value); the optima are, for the moment, taken to be unique for simplicity. When the goal is limited

[1]When carried out by iterations, the calculations might also be seen as modelling the real processes of price and quantity adjustments.

[2]The usual theory of differentiable convex functions is, of course, included in subdifferential calculus as a special case. Furthermore, the subgradient concept can also be used to prove—rather than assume—that a convex function is differentiable by showing that it has a unique subgradient. We use this method in [24] and [27, Section 9].

[3]From Section 4 on, short-run cost minimization is split off as a subprogramme, whose solution is $\check{v}(y, k, w)$. In these terms, $\hat{v}(p, k, w) = \check{v}(\hat{y}(p, k, w), k, w)$.

to finding the producer's long-run profit maxima, it can be achieved by part-inverting the short-run solution map of (p, k, w) to $(y, -v; r)$ so that the prices (p, r, w) are mapped to the quantities $(y, -k, -v)$. This is done by solving the equation $\hat{r}(p, k, w) = r$ for k and substituting any solution into $\hat{y}(p, k, w)$ and $\hat{v}(p, k, w)$ to complete a long-run profit-maximizing input-output bundle. Such a bundle may be unique but only up to scale if the returns to scale are constant (making $\hat{r}$ homogeneous of degree zero in k).

Even within the confines of the producer problem, this approach saves effort by building on the short-run solutions that have to be found anyway: the problems of plant operation and plant valuation are of central practical interest and always have to be tackled by producers. But the short-run approach is even more important as a practical method for calculating market equilibria. For this, with the input prices r and w taken as fixed for simplicity, the short-run profit-maximizing supply $\hat{y}(p, k, w)$ is equated to the demand for the products $\hat{x}(p)$ to determine the short-run equilibrium output prices $p^{\star}_{\mathrm{SR}}(k, w)$. The capacity values $\hat{r}$, calculated at the equilibrium prices $p^{\star}_{\mathrm{SR}}(k, w)$ with the given k and w, are only then equated to the given capacity prices r to determine the long-run equilibrium capacities $k^{\star}(r, w)$, and hence also the long-run equilibrium output prices and quantities (by substituting $k^{\star}$ in the short-run equilibrium solution).[4] In other words, the determination of investment is postponed until *after* the equilibrium in the product markets has been found: the producer's long-run problem is split into two—that of operation and that of investment—and the equilibrium problem is "inserted" in-between. Since the operating solutions usually have relatively simple forms, doing things in this order can greatly ease the fixed-point problem of solving for equilibrium: indeed, the problem can even be elementary if approached in this way (Section 2). Furthermore, unlike the optimal investment of the pure producer problem, the equilibrium investment $k^{\star}$ has a definite scale (determined by demand for the products). Put another way: $\hat{r}(p^{\star}_{\mathrm{SR}}(k, w), k, w)$, the value to be equated to r, is *not* homogeneous of degree zero in k like $\hat{r}(p, k, w)$. Thus one can keep mostly to single-valued maps, and avoid dealing with multi-valued correspondences. And finally, like the short-run producer optimum, the short-run general equilibrium is of interest in itself.

This exposition comes in three parts. The first and main part (Sections 2 to 19) contains various characterizations of long-run producer optima, but its core is a framework for the short-run approach to the long-run general-equilibrium pricing of a range of commodities with joint costs of production (Sections 12 and 13). This is applied to the peak-load pricing of electricity generated by a variety of techniques (Sections 15 to 17); a greatly simplified version of this problem serves also as an introductory example (Section 2). The characterizations of producer optima (needed for the short-run approach) are complemented by criteria for the existence of optimum quantities and shadow prices for the short-run profit maximization and cost minimization problems, and for the equality of total values of the variable quantities and the fixed quantities, i.e., for the absence of a gap between the primal and dual solutions. These results form the second part (Sections 20 to 23). The third and last part (Sections 24 to 28) introduces the concept

[4]The short-run approach to equilibrium might also be based on short-run cost minimization, in which not only the capital inputs (k) but also the outputs (y) are kept fixed and are shadow-priced in the dual problem, but such cost-based calculations are usually much more complicated than those using profit maximization: see Section 12.

of technologies with conditionally fixed coefficients, and the preceding general analysis is specialized to this class. Two short appendices (A and B) provide contextual examples of mathematical complications, one possible but exceptional (a duality gap), the other typical (a nonfactorable joint subdifferential of a nondifferentiable bivariate convex function). Appendix C gives the required results of convex calculus (with one innovation, viz., Lemma C.5 on subdifferential sections).

As a simple but instructive introduction to the short-run approach to long-run equilibrium, we rehearse Boiteux's treatment of the simplest peak-load pricing problem, viz., the problem of pricing the services of a homogeneous capacity that produces a nonstorable good with cyclic demands (such as electricity). A direct calculation of the long-run equilibrium poses a fixed-point problem, but, with cross-price independent demands, the short-run equilibrium is obtained by the elementary method of intersecting the supply and demand curves for each time instant separately. At each time t, the short-run equilibrium output price $p^\star_{\mathrm{SR}}(t)$ is the sum of the unit operating cost w and a capacity charge $\kappa^\star_{\mathrm{SR}}(t) \geq 0$ that is nonzero only at times of full capacity utilization, i.e., when the output rate $y^\star_{\mathrm{SR}}(t)$ equals the given capacity k. Finally, the long-run equilibrium is found by adjusting the capacity k so that its unit cost r equals its unit value, defined as the unit operating profit, which equals the total capacity charge over the cycle, $\int_0^T \kappa^\star_{\mathrm{SR}}\,\mathrm{d}t = \int_0^T (p^\star_{\mathrm{SR}} - w)\,\mathrm{d}t$. This solution is given by Boiteux with discretized time [9, 3.2–3.3].[5] Its continuous-time version is given in Section 2.

We develop Boiteux's idea into a frame for the analysis of investment and pricing by an industry that supplies a range of commodities—such as a good differentiated over time, locations or events (Sections 12 and 13). In Sections 15 to 17, this is applied to augment the rudimentary one-station model to a continuous-time equilibrium model of electricity pricing with a diverse technology, including energy storage and hydro as well as thermal generation. Such a plant mix makes supply cross-price dependent, even in the short run (i.e., with the capacities fixed). Demand, too, is allowed to be cross-price dependent.

The setting up of the short-run approach to pricing and investment (Sections 12 and 13) is the most novel part of this study. Unlike the characterization and existence results about producer optima, this is not fully formalized into theorems: we assume, rather than prove, that the short-run equilibrium is unique, and we merely note that its existence cannot be guaranteed unless the fixed capacities are all positive (i.e., unless $k \gg 0$).[6] The question of a general method of computing short-run market equilibria is only touched upon, in Figure 3, where the use of Walrasian tatonnement is suggested.[7] And we do not establish any qualitative properties of the long-run condition $\hat{r}\left(p^\star_{\mathrm{SR}}(k, w), k, w\right) = r$,

[5]Boiteux's work is also presented by Drèze [15, pp. 10–16], but the short-run character of the approach is more evident from the original [9, 3.2–3.3] because Boiteux discusses the short-run equilibrium first, before using it as part of the long-run equilibrium system. Drèze mentions the short-run equilibrium on its own only as an afterthought [15, p. 16].

[6]This is not an unacceptable condition, but some capacities could of course be zero in long-run equilibrium. The long-run model meets the usual adequacy assumption, as does the short-run model with positive capacities, and so the existence of an equilibrium follows from results such as Bewley's [7, Theorem 1], which is amplified in [31, Section 3] and [29] by a proof using the continuity of demand in prices.

[7]As is well known, this process does not always converge, but there are other iterative methods.

as an equation for the investment k (or, more generally, as a subdifferential inclusion, viz., (13.18)). But it is shown that our *SRP programme-based system*, consisting of Conditions (13.11)–(13.15) together with (13.18)–(13.19), is a full characterization of long-run market equilibrium. Furthermore, it is clear already from the introductory example of Section 2 that the short-run approach can greatly simplify the problem of solving for long-run equilibrium (as well as finding the short-run equilibrium on the way). It is apparent that the approach is worth applying not only to the case of electricity but also to the supply of other time-differentiated commodities (such as water, natural gas, etc.). The questions of uniqueness, stability and iterative computation of equilibria, though important, are not specific to the short-run approach; also, they have been much studied and are well understood (at least in infinite-dimensional commodity spaces). The central and distinctive quantitative elements of the approach are valuation and operation of plants, and these are problems that we have fully solved for the various types of plant in the electricity supply industry (see Section 16 and its references). The priorities in developing the approach are: (i) to analyze the valuation and operation problems for other technologies and industries, and (ii) to compute numerical solutions from real data by using, at least to start with, the standard methods (viz., linear programming for producer optima and tatonnement for market equilibria). It would seem sensible to address the theoretical questions of uniqueness and stability in the light of future computational experience (in which more elaborate iterative methods could be employed if necessary). These questions are potentially important for practice as well as for completing the theory, but they are not priorities for this study, and are left for further research.

Sections 3 to 11, between the introductory example and the setup for the short-run approach, give characterizations of long-run producer optima. Each is either an *optimization system* or a *differential system*, i.e., it is a set of conditions formulated in terms of either the marginal optimal values or the optimal solutions to a primal-dual pair of programmes (although one can also mix the two kinds of condition in one system).

Though equivalent, the various systems are not equally usable, and the best choice of system depends on one's purpose as well as on the available mathematical description of the technology. In our application to electricity pricing with non-thermal as well as thermal generation, the technology is given by production sets (rather than profit or cost functions), and so the best tool for the short-run approach is the system using the programme of maximizing the short-run profit (SRP), together with the dual programme for shadow-pricing the fixed inputs. For each *individual* plant type,[8] the problem of minimizing the short-run cost (SRC) is typically easy (if it arises at all); therefore, it can be split off as a subprogramme (of profit maximization). The resulting *split SRP optimization system* serves as the basis of our framework for the short-run approach to pricing and industrial investment (Section 13). Because of its importance to our applications, this system is introduced as soon as possible, in Section 4—not only before the differential systems (Sections 7, 8 and 11), but also before the other optimization

[8]By contrast, SRC minimization for a *system* of plants can be difficult because it involves allocating the system's given output among the plants. Its complexity shows in, e.g., the case of a hydro-thermal electricity-generating system [35]. Our decentralized approach avoids having to deal directly with the formidable problem of minimizing the entire system's cost: see the Comments containing Formulae (12.3) and (12.4).

systems (Sections 6 and 11), and even before a discussion of the dual programmes (in Section 5).

Of the differential systems, the first one to be presented formally, in Section 7, is that which generalizes Boiteux's original set of conditions, limited though it is to technologies that are simple enough to allow explicit formulae not only for the SRC function but also for the SRP function. Another differential system, introduced informally in Section 2 and formally in Section 11, has the same mathematical form but uses the LRC instead of the SRP function (with the variables suitably switched). The two systems' equivalence extends the Wong-Viner Envelope Theorem (on the equality of SRMC and LRMC) to convex technologies with nondifferentiable cost functions by Formula (11.1)—and this is the result outlined earlier in Section 2 (where it is exemplified by our account of Boiteux's short-run approach to the simple peak-load pricing problem). The extension is made possible by using the subdifferential (a.k.a. the subgradient set) as a generalized, multi-valued derivative. This is necessary because the joint-cost functions may lose differentiability at crucial points. For example, in the simplest peak-load pricing problem, the long-run cost is nondifferentiable at every output bundle with multiple global peaks because, although the total capacity charge is determinate (being equal to r, the given rental price of capacity), its distribution over the peaks cannot be determined purely by cost calculations. And, far from being exceptional, multiple peaks forming an output plateau do arise in equilibrium as a solution to the shifting-peak problem, as we show in [26] under appropriate assumptions about demand.[9] The short-run marginal cost is even less determinate: whenever the output rate reaches full capacity, an SRMC exceeds the unit operating cost w by an arbitrary amount κ—which makes the capacity charge indeterminate in total as well as in its distribution. This is an example of the inclusion between the subdifferentials of the two costs, as functions of the output bundle: the set of SRMCs is larger than the set of LRMCs when the capital inputs are at an optimum (i.e., minimize the total cost). It then takes a stronger condition to ensure that a particular SRMC is actually an LRMC. What is needed is the equality of rental prices to the *profit*-imputed values of the fixed inputs (which are the fixed inputs' marginal contributions to the operating profit). This equality is the required generalization of Boiteux's long-run optimum condition, which, for his one-station technology, equates the capacity price r to the unit operating profit $\int \kappa \, \mathrm{d}t = \int (p(t) - w) \, \mathrm{d}t$ [9, 3.3, and Appendix: 12]. The valuations *must* be based on increments to the operating profit: it is generally ineffective to try to value capacity increments by any reductions in the operating cost. The one-station example shows just how futile such an attempt can be: excess capacity does not reduce the operating cost at all, but any capacity shortage makes the required output infeasible. This leaves the capacity value completely indeterminate by

[9]This shows how mistaken is the widespread but unexamined view that nondifferentiabilities of convex functions are of little consequence: the very points which are *a priori* exceptional turn out to be the rule rather than the exception in equilibrium. Also, it is only on finite-dimensional spaces that convex functions are "generically smooth" or, more precisely, twice differentiable almost everywhere with respect to the Lebesgue measure (Alexandroff's Theorem). On an infinite-dimensional space, a convex function can be nondifferentiable everywhere.

SRC calculations—in contrast to the definite value $\int (p(t) - w)\,\mathrm{d}t$ obtained by calculating the SRP. Only with differentiable costs is the SRC as good as the SRP for the purpose of capital-input valuation.

Our extension of the Wong-Viner Envelope Theorem uses the SRP function and thus achieves for any convex technology what Boiteux [9, 1.1–1.2 and 3.2–3.3] in effect does with the very simple but nondifferentiable cost functions of his problem, which are spelt out here in (2.5) and (2.6). He realizes that there is something wrong with the supposed equality of SRMC and LRMC [9, 1.1.4 and 1.2.2]. As he puts it,

> "It seems practically out of the question that these costs should be equal; it is difficult to imagine, for instance, how the marginal cost of operating a thermal power station could become high enough to equal the development cost (including plant) of the thermal energy [its long-term marginal cost]. The paradox is due to the fact that most industrial plants are in reality very 'rigid'. . . .
>
> There is no. . . question of equating the development cost to the cost of overloading the plant, since any such overloading is precluded by the assumption of rigidity. . . . The more usual types of plant have some slight flexibility in the region of their limit capacities. . . but. . . any 'overloading' which might be contemplated in practice would never be sufficient to equate its cost with the development cost; hence the paradox referred to above."

Its resolution starts with his

> "new notion which will play an essential part in 'peak-load pricing': for output equal to maximum, the differential cost [the SRMC] is indeterminate: it may be equal to, or less or greater than the development cost [the LRMC]."

In the language of subdifferentials, Boiteux's "new notion"—that the LRMC is just one of many SRMCs—is a case of the afore-mentioned general inclusion between the LRMCs and SRMCs, which is usually strict: $\partial_y C_{\mathrm{LR}}(y, r) \subsetneq \partial_y C_{\mathrm{SR}}(y, k)$ when $r \in -\partial_k C_{\mathrm{SR}}(y, k)$, i.e., when the bundle of capital inputs k minimizes the total cost of an output bundle y, given their prices r (and given also the variable-input prices w, which, being kept fixed, are suppressed from the notation). For differentiable costs, this reduces to the Wong-Viner equality of gradient vectors: $\nabla_y C_{\mathrm{LR}} = \nabla_y C_{\mathrm{SR}}$ (when the capital inputs are at an optimum). But for nondifferentiable costs, all it shows is that each LRMC is an SRMC—which is the *reverse* of what is required for the short-run approach. The way out of this difficulty is to bring in the SRP function, Π_{SR}, and require that the given prices for the capital inputs are equal to their profit-imputed values, i.e., that $r = \nabla_k \Pi_{\mathrm{SR}}(p, k)$ or, should the gradient not exist, that $r \in \widehat{\partial}_k \Pi_{\mathrm{SR}}$ (which is the superdifferential a.k.a. the supergradient set). This condition is stronger than cost-optimality of the fixed inputs when the output price system p is an SRMC, i.e., if $p \in \partial_y C_{\mathrm{SR}}(y, k)$ then $\widehat{\partial}_k \Pi_{\mathrm{SR}}(p, k) \subseteq -\partial_k C_{\mathrm{SR}}(y, k)$, generally with a strict inclusion (indeed, $\nabla_k \Pi_{\mathrm{SR}}$ can exist even when $\nabla_k C_{\mathrm{SR}}$ does not, in which case $\nabla_k \Pi_{\mathrm{SR}} \in -\partial_k C_{\mathrm{SR}}$). And the new condition—that $r \in \widehat{\partial}_k \Pi_{\mathrm{SR}}(p, k)$—is no stronger than it need be: it is just strong enough to do the job and guarantee that if $p \in \partial_y C_{\mathrm{SR}}(y, k)$ then $p \in \partial_y C_{\mathrm{LR}}(y, r)$.

Thus our analysis of the relationship between SRMC and LRMC bears out, amplifies and develops Boiteux's ideas, which, at the time, he allowed, with a hint of exasperation, were "false in the theoretical general case, but more or less true of ordinary industrial plant". We accommodate both cases: the industrial reality of fixed coefficients and rigid capacities as well as the unrealistic textbook supposition of smooth costs. By bridging the gap between the inadequate existing theory and its intended applications, we put an end to its disturbing and unnecessary divorce from reality. This allows peak-load pricing to be put, for the first time, on a sound and rigorous theoretical basis (Sections 15 to 17).

From our perspective, Boiteux's long-run optimum condition, that $r = \int (p(t) - w)\, \mathrm{d}t$, should be viewed as a special case, for the one-station technology, of the equation $r = \nabla_k \Pi_{\mathrm{SR}}$. But staying within the confines of this particular example, Boiteux interprets his condition merely as recovery of the total cost of production, including the capital cost [9, 3.4.2: (2) and Conclusions: 4]. This is correct, but only in the case of a single capital input, and it cannot provide a basis for dealing with a production technique that uses a number of interdependent capital inputs.[10] In such a case, our generalization of Boiteux's long-run optimum condition is stronger than capital-cost recovery: i.e., under constant returns to scale, if $r \in \widehat{\partial}_k \Pi_{\mathrm{SR}}$ (or $r = \nabla_k \Pi_{\mathrm{SR}}$), then $r \cdot k = \Pi_{\mathrm{SR}}$, but not *vice versa* (though the converse *is* of course true when k is a positive scalar). To think purely in terms of marginal costs and cost recovery is a dead end: with multiple capital inputs, cost recovery is *not* sufficient to guarantee that a short-run equilibrium is also a long-run equilibrium or, equivalently, that an SRMC tariff is also an LRMC tariff. The SRP function with its marginals (derivatives w.r.t. k), or the SRP programme with the dual solution, have to be brought into the short-run approach. This is done here for the first time.

In mathematical terms, the Extended Wong-Viner Theorem (11.1) comes from what we call the Subdifferential Sections Lemma (SSL), which gives the joint subdifferential of a bivariate convex function ($\partial_{y,k} C$) in terms of one of *its* partial subdifferentials ($\partial_y C$) and a partial superdifferential, $\widehat{\partial}_k \Pi (p, k)$, of the relevant partial conjugate (which is a saddle function): see (9.3), and Lemma C.5 in Appendix C. This is applied, twice, to either the SRP or the LRC as a saddle function obtained by partial conjugacy from the SRC, which is a jointly convex function (C) of the output bundle y and the fixed-input bundle k, with the variable-input prices w kept fixed (Section 11). The SSL can be regarded as a direct precursor of a well-known result of convex calculus, viz., the Partial Inversion Rule (PIR), which relates the partial sub/super-differentials of a saddle function ($\partial_p \Pi$ and $\widehat{\partial}_k \Pi$) to the joint subdifferential of its bivariate convex "parent" function ($\partial_{y,k} C$): see Lemmas C.6 and C.8 (whose proofs derive the PIR from the SSL). One well-known application of this fundamental principle is the equivalence of two optimality conditions,

[10]Capital inputs are called *independent* if the SRP function (Π_{SR}) is linear in the capital-input bundle $k = (k_1, k_2, \ldots)$; an example is the multi-station technology of thermal electricity generation. Such a technology effectively separates into a number of production techniques with a single capital input each, and Boiteux's analysis applies readily: to ensure that the short-run equilibrium is also a long-run one, it suffices to require cost recovery for each production technique θ with $k_\theta > 0$, although one must also remember to check that any unused production technique (one with $k_\theta = 0$) cannot be profitable (e.g., that $r_\theta \geq \int (p(t) - w_\theta)\, \mathrm{d}t$ for any unused type of thermal station).

viz., the parametric version of Fermat's Rule and the Kuhn-Tucker characterization of primal and dual optima as a saddle-point of the Lagrange function: see, e.g., [45, 11.39 (d) and 11.50]. Another well-known use of the PIR establishes the equivalence of Hamiltonian and Lagrangian systems in convex variational calculus; when the Lagrange integrand is nondifferentiable, this usefully splits the Euler-Lagrange inclusion (a generalized equation system) into the pair of Hamiltonian differential inclusions, and it may even transform the inclusion into ordinary equations because the Hamiltonian can be differentiable even when the Lagrangian is not: see, e.g., [44, (10.38) and (10.40)], [43, Theorem 6] or [4, 4.8.2].[11] Our own use of the PIR or the SSL relates the marginal optimal values for a programme to those of a subprogramme, to put it in general terms. In the specific context of extending the Wong-Viner Theorem, SRC minimization is a subprogramme both of SRP maximization and of LRC minimization; their optimal values are $C_{\mathrm{SR}}(y,k)$, $\Pi_{\mathrm{SR}}(p,k)$ and $C_{\mathrm{LR}}(y,r)$, respectively. This is a new use of what is, in Rockafellar's words, "a striking relationship...at the heart of programming theory" [41, p. 604].

One half of this argument (the application of the SSL to the saddle function Π_{SR} as a partial conjugate of the bivariate convex function C_{SR} to prove the first equivalence in (11.1)) is given already in Section 9. It comes along with other applications of the PIR and the SSL that establish the equivalence of saddle differential systems to the systems with joint subdifferentials of Section 8.

Like all optimization, economic theory has to deal with the nondifferentiability of optimal values that commonly arises even when the programmes' objective and constraint functions are all smooth. This has led to the eschewing of marginal concepts in rigorous equilibrium analysis, but any need for this disappeared with the advent of nonsmooth calculus. Of course, in using generalized derivatives such as the subdifferential, one cannot expect to transcribe familiar theorems from the smooth to the subdifferentiable case simply by replacing the ordinary single gradients with multi-valued subdifferentials—proper subdifferential calculus must be applied. This not only extends the scope for marginal analysis, but also leads to a rethinking and reinterpretation that can give a new economic content to well-known results. The Wong-Viner Theorem is a case in point: a useful extension depends on recasting its fixed-input optimality assumption in terms of profit-based valuations (i.e., on restating the optimality of fixed inputs as equality of their rental prices to their marginal contributions to the operating profit). After this reformulation of optimality in terms of SRP marginals—but not before—the "smooth" version can be transcribed to the case of subdifferentiable costs (by replacing each ∇ with a ∂). Without this preparatory step, all extension attempts are doomed: a direct transcription of the original Wong-Viner equality of SRMC and LRMC to the subdifferentiable case is plainly false, and although it can be changed to a true inclusion without bringing in the SRP function, that kind of result fails to attain the goal of identifying an SRMC as an LRMC.[12]

[11] To distinguish the two quite different meanings of the word "Lagrangian", we occasionally expand it into either "Lagrange function" (in the multiplier method of optimization) or "Lagrange integrand" (in the calculus of variations only).

[12] Without involving Π_{SR}, the inclusion ($\partial_y C_{\mathrm{LR}} \subseteq \partial_y C_{\mathrm{SR}}$) can be improved only by making it more precise but no more useful: $\partial_y C_{\mathrm{SR}}(y,k)$ can be shown to equal the union of $\partial_y C_{\mathrm{LR}}(y,r)$ over $r \in -\partial_k C_{\mathrm{SR}}(y,k)$, i.e., over all those fixed-input price systems r for which k is an optimal fixed-input bundle

Herein, one well-known optimality condition is, in the main part, conspicuous by its absence. The Lagrangian saddle-point condition of Kuhn and Tucker is central to the duality theory of convex programmes (CPs)—and we do use it in our studies of hydro and energy storage [21], [23], [24] and [27], which serve the short-run framework's application to electricity supply in Sections 15 to 17—but we do not use the Kuhn-Tucker system in the main part of this analysis (Sections 3 to 17). Instead, for a general analysis with an abstract production cone, we prefer the Complementarity Conditions on the price system and the input-output bundle (3.5). This system is a case of what we call the FFE Conditions, which consist of primal feasibility, dual feasibility and equality of the primal and dual objectives (at the feasible points in question). The FFE Conditions form an effective system whenever the dual programme can be worked out from the primal explicitly. This is so with the profit and cost problems because they become linear programmes (LPs) once the production cone is represented by linear inequalities. For an LP, the FFE system is linear in the primal and dual variables jointly—unlike the Kuhn-Tucker system (which is nonlinear because of the quadratic term in the complementary slackness condition): compare (5.3) with (5.2). And a linear system (i.e., a system of linear equalities and inequalities) is much simpler to deal with: in particular, it can be solved numerically by the simplex method (or another LP algorithm). The problem's size is smaller, though, when the method is applied directly to the relevant LP (or to its dual), rather than to its FFE system.[13] Either way, there is no need for the Kuhn-Tucker system in solving the SRP programmes with their fixed-input valuation duals—although it is instrumental in proving uniqueness of their solutions, in [21], [27] and [24].

In the LP formulation of a profit or cost programme, the fixed quantities are primal parameters but need not be the same as the standard "right-hand side" parameters—and so their shadow prices, which are the dual variables, need not be identical to the standard dual variables. Yet the usual theory of linear programming works with the standard parameterization, and it is the *standard* dual solution that the simplex method provides along with the primal solution. But, as we show in Section 14, this is not much of a complication because any other dual variables can be expressed in terms of the standard dual variables, i.e., in terms of the usual Lagrange multipliers for constraints. We use this in valuing the fixed inputs for electricity generation, in Section 16. The principle has also a counterpart beyond the linear or convex duality framework: it is the Generalized Envelope Theorem for smooth optimization, whereby the marginal values of all parameters, including any nonstandard ones, are equal to the corresponding partial derivatives of the ordinary Lagrangian—and are thus expressed in terms of the constraints' multipliers. See [1, (10.8)] or [47, 1.F.b].

Our exposition of the producer optimum pauses for "stock-taking" in Section 10. In particular, Tables 1 and 2 summarize the various characterizations of the long-run optimum, though not their "mirror images" which result from a formal substitution of the LRC for the SRP. These tables record also the methods employed to transform these systems into one another. This shows a unity: the same methods are applied to systems of the same type, even though the exposition gives special places to the two systems of

---

for the output bundle y (given also the omitted variable-input price system w). See the Comments at the end of Appendix B.

[13] See the Comment at the end of Section 6 for a count of variables and constraints.

importance for our application of the short-run approach, viz., the split SRP optimization system of Section 4 and the SRC-P saddle differential system of Section 7. The latter system's "mirror image", the L-SRC saddle differential system of Section 11, is also directly involved in our application when its conditions of LRMC pricing and LRC minimization serve as the definition of long-run optimum—as is often the case in public utility pricing, including Boiteux's work and our account of it in Section 2. The other fourteen systems are not used here, but any can be the best tool, for the short-run approach as for other purposes, if the technology is described most simply in the system's own terms; see also the Comments at the end of Section 12. In particular, one should not be trapped by the language into thinking that a system using the LRC programme or function is somehow fundamentally unsuitable for the short-run approach.

Section 10 ends by noting that some of the systems—including the two "special" ones—can be partitioned into a *short-run subsystem* (which characterizes SRP maxima) and a supplementary condition that generalizes Boiteux's long-run optimum condition and requires that investment be at a profit maximum.

A complete formalization of all the duality-based systems is deferred to Sections 18 and 19, where the programmes' duality and the systems' equivalence are cast as rigorous results with proofs. To this end, we restate formally the subdifferential versions of the Shephard-Hotelling Lemmas (announced in Section 6). As has long been known [14, pp. 555 and 583], these are cases of the derivative property of the optimal value, which transcribes to the subdifferentiable case directly (by replacing ∇ with ∂).

Our characterizations of the producer optimum are complemented by results on the equality of the primal and dual values and the solubility of both programmes. Such an analysis is given in Sections 20 to 23; it yields sufficient criteria for the existence of a pair of solutions with equal values. First, we recall from the general theory of CPs that the absence of a duality gap is equivalent to the semicontinuity of the optimal value, and we apply this to the profit and cost programmes (Section 20). To make this criterion applicable, we give some sufficient conditions for the required semicontinuity of SRP as well as LRC and SRC as functions of the programmes' quantity data (Section 21). When the commodity space for the variable quantities (the programme's decision variables) is infinite-dimensional, we utilize its weak* topology as well as its vector order. It is therefore taken to be a dual Banach lattice (i.e., the dual of a completely normed vector lattice). One example is $L^{\infty}[0,T]$, which serves as the output space in our application to peak-load pricing. With this or any other nonreflexive commodity space (for the programme's variable quantities), our results on the semicontinuity of profit or cost (as a function of the fixed quantities) apply only when the given price system (for the variable quantities) lies not just in the dual but actually in the smaller predual space. Such a criterion is adequate only when the general-equilibrium price system is known to lie in the predual—as is the case for the commodity space L^{∞} under Bewley's assumptions [7], which we weaken in [26] to make that result applicable to continuous-time problems. But even the weakened assumption is restrictive: it requires that brief interruptions of a consumption or input flow cause only small losses of utility or output (interruptibility of demand). When this is not so and the programme's price system cannot be taken to lie in the predual, a duality gap can be precluded by imposing a generalized form of Slater's Condition (Section 23). This guarantees not only the semicontinuity, but even the continuity of profit or cost as a function of quantities, and thus also its subdifferentiability (i.e., the existence of a subgradient) or, equivalently, solubility of the dual programme for shadow-pricing the fixed quantities. The primal programme of optimal operation is shown to be soluble in Section 22, when the given price system (for the variable quantities) lies in the predual of the commodity space. When it does not, the programme can still be soluble in some, though not all, cases (it must be soluble in general equilibrium even when the supporting price system does not lie in the predual space).[14]

Thermal generation and pumped storage of electricity are examples of production techniques with conditionally fixed coefficients (c.f.c.). Introduced in Section 24, this concept extends the notion of a fixed-coefficients technology to the case of a multi-dimensional output bundle. The convex programme of SRP maximization for a c.f.c. technique, its dual and the Kuhn-Tucker Conditions are also spelt out in Section 24, although fully formalized statements and proofs are deferred to Section 25. In Section 26, the assumptions of Sections 21 to 23 are verified for c.f.c. techniques. Therefore, the solubility and no-gap results of Sections 21, 22 and 23 can be applied to the profit and cost programmes with such a technology, and this is done for the SRP programme (with its dual) in Section 27. Finally, in Section 28, we note that c.f.c. techniques can also be handled by linear programming (as is done for the electricity generation and storage techniques in Section 16).

Notation is explained when first used, but it is also listed below in several categories. Later, Table 3 shows the correspondence of notation between our general duality scheme of Sections 5 and 14 and its application to electricity supply.[15]

List of Notation

Profit and cost optimization and shadow-pricing programmes: parameters and decision variables, solutions, optimal values and marginal values

$y \in Y$ an output bundle, in a space Y

$k \in K$ a fixed-input bundle, in a space K

$v \in V$ a variable-input bundle, in a space V

$p \in P$ an output price system, in a space P

$r \in R$ a fixed-input price system, in a space R

$w \in W$ a variable-input price system, in a space W

$\mathit{\Delta} y$, $\mathit{\Delta} k$, etc. increments to y, k, etc. ($\mathit{\Delta}$ differs from the upright Δ)

$\mathbb{Y}$ a production set (in the commodity space $Y \times K \times V$)

A, B and C matrices or linear operations, esp. such that $(y, -k, -v) \in \mathbb{Y}$ if and only if $Ay - Bk - Cv \leq 0$

A^{T} the transpose of a matrix A

$\delta(\cdot \mid \mathbb{Y})$ the 0-∞ indicator function of the set $\mathbb{Y}$ (equal to 0 on $\mathbb{Y}$)

---

[14] In [21] and [24], we give examples of an SRP programme in which the output space is $L^{\infty}[0,T]$ and a "singular" price term places the price system outside the predual $L^{1}[0,T]$, but it is the timing of the singularity, and not just its presence, that decides whether the programme is soluble or not.

[15] Note the two different uses of the symbols s and σ: in Sections 5 and 14, these mean the standard parameters and dual variables, but in Section 15 they mean the energy stock and water spillage. Also, the n_{θ}, n_{St} and n_{Tu} of Section 16 are lower constraint parameters (whose original, unperturbed values are zeros). In Sections 13 and 17, n means an input of the numeraire.

$\mathbb{Y}^{\circ}$ the polar cone of $\mathbb{Y}$ (a cone in $P \times R \times W$ when $\mathbb{Y}$ is a cone in $Y \times K \times V$)

$\mathbb{Y}^{\circ}_{p,w}$ the polar cone's section through (p, w)

$\mathcal{G}'$ and $\mathcal{G}''$ respectively, the sets of generators and of spanning vectors of $\mathbb{Y}^{\circ}$, when $\mathbb{Y}$ is a polyhedral cone in a finite-dimensional space

$\mathrm{proj}_Y(\mathbb{Y})$ projection on Y of a subset, $\mathbb{Y}$, of $Y \times K \times V$

$\mathbb{Y}_{\mathrm{SR}}(k)$ short-run production set (the section of $\mathbb{Y}$ through $-k$)

$\mathbb{I}_{\mathrm{LR}}(y)$ long-run input requirement set (the negative of the section of $\mathbb{Y}$ through y)

$\mathbb{I}_{\mathrm{SR}}(y, k)$ short-run input requirement set (the negative of the section of $\mathbb{Y}$ through $(y, -k)$)

$\mathrm{vmax}\, Z$ and $\mathrm{vmin}\, Z$ sets of all the maximal and of all the minimal points of a subset, Z, of an ordered vector space (used with $\mathbb{Y}_{\mathrm{SR}}(k)$, $\mathbb{I}_{\mathrm{LR}}(y)$ or $\mathbb{I}_{\mathrm{SR}}(y, k)$ as Z)

Π_{LR} the maximum long-run profit, a function of (p, r, w)

Π_{SR} the maximum short-run a.k.a. operating profit, a function of (p, k, w)

C_{LR} the minimum long-run cost, a function of (y, r, w)

C_{SR} the minimum short-run cost, a function of (y, k, w)

∂C the subdifferential of a convex function C

$\widehat{\partial}\Pi$ the superdifferential of a concave function Π

$\nabla\Pi$ the (Gateaux) gradient vector of a function Π

$\partial/\partial k$ partial differentiation with respect to a scalar variable k

$\check{V}(y, k, v)$ the set of all variable-input bundles that minimize the short-run cost

$\check{v}(y, k, v)$ *the* variable-input bundle such as above (i.e., minimizing the short-run cost), if it is unique

$\hat{Y}(p, k, w)$ the set of all output bundles that maximize the short-run profit (i.e., maximize the function $\langle p \,|\, \cdot \rangle - C_{\mathrm{SR}}(\cdot, k, w)$)

$\hat{y}(p, k, w)$ *the* output bundle such as above (i.e., maximizing the function $\langle p \,|\, \cdot \rangle - C_{\mathrm{SR}}(\cdot, k, w)$), if it is unique

$\hat{K}(p, r, w)$ the set of all fixed-input bundles that maximize the long-run profit

$\hat{k}(p, r, w)$ *the* fixed-input bundle such as above (i.e., maximizing the long-run profit), if it is unique (under decreasing returns to scale)

$\underline{C}_{\mathrm{SR}}(y, k, w)$ the maximum, over shadow prices, of total output value less fixed-input value (and less Π_{LR} when $\mathbb{Y}$ is not a cone)

$\underline{C}_{\mathrm{LR}}(y, r, w)$ the maximum, over shadow prices, of total output value (less Π_{LR} when $\mathbb{Y}$ is not a cone)

$\overline{\Pi}_{\mathrm{SR}}(p, k, w)$ the minimum, over shadow prices, of total fixed-input value (plus Π_{LR} when $\mathbb{Y}$ is not a cone)

$\hat{R}(p, k, w)$ the set of all fixed-input price systems that minimize the total fixed-input value (plus Π_{LR} when $\mathbb{Y}$ is not a cone)

$\hat{r}(p, k, w)$ *the* fixed-input price system such as above (i.e., minimizing the total fixed-input value), if it is unique

$\check{P}(y, k, w)$ the set of all output price systems that maximize the total output value less fixed-input value, $\langle \cdot \,|\, y \rangle - \overline{\Pi}_{\mathrm{SR}}(\cdot, k, w)$, less Π_{LR} when $\mathbb{Y}$ is not a cone

$\check{p}(y, k, w)$ *the* output price system such as above (i.e., maximizing $\langle \cdot \,|\, y \rangle - \overline{\Pi}_{\mathrm{SR}}(\cdot, k, w)$), if it is unique

s vector of the standard primal parameters for a convex or linear programme (paired to its equality and inequality constraints)

σ vector of the standard dual variables (Lagrange multipliers of the constraints) for a convex or linear programme

$\hat{\Sigma}(p, s)$ the set of all the standard dual solutions (Lagrange multiplier systems), when the primal is a linear programme with s as its primal parameters and $\langle p \mid \cdot \rangle$ as its linear objective function

$\hat{\sigma}(p, s)$ *the* standard dual solution such as above, if it is unique

$\mathcal{L}$ the Lagrangian (the Lagrange function of the primal and dual variables and parameters)

Characteristics of the Supply Industry

θ a production technique of the Supply Industry

Φ_θ the set of fixed inputs of production technique θ

Ξ_θ the set of variable inputs of production technique θ

$\mathbb{Y}_\theta$ the production set of technique θ, a cone in $Y \times \mathbb{R}^{\Phi(\theta)} \times \mathbb{R}^{\Xi(\theta)}$

ξ a variable input, with a price w_ξ

ϕ a fixed input, with a price r_ϕ

Φ^{F} the set of fixed inputs with *given* prices r^{F}

Φ^{E} the set of fixed inputs with prices r^{E} to be determined in long-run equilibrium

G_ϕ the supply cost of an equilibrium-priced input $\phi \in \Phi^{\mathrm{E}}$, a function of the supplied quantity k_ϕ

Characteristics of consumer and factor demands (from Industrial User)

F production function of the Industrial User—a function of inputs: n of the numeraire and z of the differentiated good (e.g., electricity)

U_h consumer h's utility, a function of consumptions: φ of the Industrial User's product, m of the numeraire and x of the differentiated good (e.g., electricity)

$u(t, \mathsf{x})$ the consumer's instantaneous utility from the consumption rate x at time t (when U is additively separable)

m_h^{En} consumer h's initial endowment of the numeraire

$\varsigma_{h\phi}$ consumer h's share of profit Π_ϕ from the supply of input $\phi \in \Phi^{\mathrm{E}}$

$\varsigma_{h\,\mathrm{IU}}$ consumer h's share in the Industrial User's profit, Π_{IU}

$\varpi_{h\theta}$ consumer h's share in the operating profit from production technique θ of the Supply Industry

$B(p, \varrho, M)$ consumer's budget set when his income is M, the differentiated good (electricity) price is p and the Industrial User's product price is ϱ

$\hat{M}_{\mathrm{SR}\,h}\left(p; r^{\mathrm{E}}, r^{\mathrm{F}}; w, \varrho \mid k\right)$ consumer's income in the short run

$\hat{M}_{\mathrm{LR}\,h}\left(p, r^{\mathrm{E}}, \varrho\right)$ consumer's income in the long run (Supply Industry's pure profit is zero)

$\hat{x}_h(p, \varrho; M)$ consumer h's demand for the differentiated good (electricity) when its price is p, the Industrial User's product price is ϱ, and the income is M

$\hat{\varphi}_h(p, \varrho; M)$ consumer h's demand for the Industrial User's product when its price is ϱ, the differentiated good's (electricity) price is p, and the consumer's income is M

$\hat{z}(p, \varrho)$ the Industrial User's factor demand for the differentiated good (electricity)

$\hat{n}(p, \varrho)$ the Industrial User's factor demand for the numeraire

Short-run general-equilibrium prices and quantities

$p^{\star}_{\mathrm{SR}}$, $\varrho^{\star}_{\mathrm{SR}}$ prices for the differentiated good (electricity) and for the IU's product
$y^{\star}_{\mathrm{SR}\,\theta}$ output of the differentiated good (electricity) by production technique θ
$v^{\star}_{\mathrm{SR}\,\theta}$ variable input into production technique θ
$x^{\star}_{\mathrm{SR}\,h}$, $z^{\star}_{\mathrm{SR}}$ consumer demand and factor demand for the differentiated good (electricity)
$m^{\star}_{\mathrm{SR}\,h}$, $n^{\star}_{\mathrm{SR}}$ consumer demand and factor demand for the numeraire
$\varphi^{\star}_{\mathrm{SR}}$ the Industrial User's output

Long-run general-equilibrium prices and quantities

w the given prices of the Supply Industry's variable inputs
r^{F} the *given* rental prices of the Supply Industry's fixed-priced capital inputs
r^{E} rental prices of the Supply Industry's equilibrium-priced capital inputs—to be determined in long-run equilibrium
$r^{\star}$ the equilibrium prices of the equilibrium-priced inputs (i.e., the equilibrium value of r^{E})
$k^{\star}_{\theta}$ equilibrium capacities of producer θ in the Supply Industry
$p^{\star}_{\mathrm{LR}}$, $y^{\star}_{\mathrm{LR}\,\theta}$, etc. equilibrium prices and quantities—as above, but for the long-run equilibrium

Electricity generation (all techniques)

$p(t)$ electricity price at time t (in $\$$/kWh), i.e., p is a time-of-use tariff
$D_t(\mathsf{p})$ cross-price independent demand for electricity (in kW) at time t, if the current price is p

Thermal generation

$S(\mathsf{p})$ in the short run, the cross-price independent rate of supply (in kW) of thermally generated electricity, if the current price is p
$c_{\mathrm{SR}}(\mathsf{y})$ the instantaneous short-run thermal cost per unit time (in $\$$/kWh), if the current output rate is y (in kW); the common graph of the correspondences S and ∂c_{SR} is the thermal SRMC curve
θ a type of thermal plant
$\widetilde{\xi}(\theta)$ fuel type used by plant type θ
w_{ξ} price of fuel of type ξ (in $ per kWh of heat)
v_{θ} fuel input of plant type θ (in kWh of heat)
η_{θ} technical efficiency of plant type θ, i.e., $1/\eta_{\theta}$ is the heat rate
$\widetilde{w}_{\theta} = w_{\widetilde{\xi}_{\theta}}/\eta_{\theta}$ unit running cost of plant type θ (in $/kWh of electricity output)
w_{θ} abbreviation for $\widetilde{w}_{\theta}$ when plant types correspond to fuels one-to-one
k_{θ} thermal generating capacity of type θ (in kW)
$\kappa_{\theta}(t)$ unit value of the generating capacity at time t, per unit time (in $/kWh)
$r_{\theta} = \int_0^T \kappa_{\theta}(t)\,\mathrm{d}t$ unit value of the generating capacity in total for the cycle (in $/kW)
$\gamma(t) = \kappa(t) / \int_0^T \kappa(t)\,\mathrm{d}t$ density, at time t, of the distribution of capacity charges over the cycle, i.e., a subgradient of the function EssSup (more generally, a subgradient of any capacity requirement function)
r^{F}_{θ} the *given* rental price of the generating capacity of type θ (in $/kW)

$\nu_\theta(t)$ unit value of nonnegativity constraint on output at time t, per unit time (in \$/kWh)

$y_\theta(t)$ rate of electricity output from plant type θ at time t (in kW)

Pumped-storage

k_{St} the plant's storage a.k.a. reservoir capacity (in kWh)

$\kappa_{\mathrm{St}}(\mathrm{d}t)$ unit value of storage capacity on a time interval of length $\mathrm{d}t$ (in \$/kWh)

$r_{\mathrm{St}} = \int_0^T \kappa_{\mathrm{St}}(\mathrm{d}t)$ unit value of storage capacity in total for the cycle (in \$/kWh)

$r^\star_{\mathrm{St}}$ the (long-run) equilibrium rental price of storage capacity (in \$/kWh)

$\nu_{\mathrm{St}}(\mathrm{d}t)$ unit value of nonnegativity constraint on energy stock on an interval of length $\mathrm{d}t$ (in \$/kWh)

k_{Co} the plant's conversion capacity (in kW)

$\kappa_{\mathrm{Pu}}(t)$ unit value of converter's pump capacity at time t, per unit time (in \$/kWh)

$\kappa_{\mathrm{Tu}}(t)$ unit value of converter's turbine capacity at time t, per unit time (in \$/kWh)

$\kappa_{\mathrm{Co}}(t) = \kappa_{\mathrm{Pu}}(t) + \kappa_{\mathrm{Tu}}(t)$ unit value of converter's capacity at time t, per unit time (in \$/kWh)

$r_{\mathrm{Co}} = \int_0^T \kappa_{\mathrm{Co}}(t)\,\mathrm{d}t$ unit value of conversion capacity in total for the cycle (in \$/kW)

$r^{\mathrm{F}}_{\mathrm{Co}}$ the *given* rental price of conversion capacity (in \$/kW)

$y_{\mathrm{PS}}(t)$ rate of electricity output from the pumped-storage plant at time t (in kW)

$\hat{Y}_{\mathrm{PS}}(p; k_{\mathrm{St}}, k_{\mathrm{Co}})$ the set of all the electricity output bundles that maximize the operating profit of a pumped-storage plant with capacities $(k_{\mathrm{St}}, k_{\mathrm{Co}})$, when the electricity tariff is p

$\hat{y}_{\mathrm{PS}}(p; k_{\mathrm{St}}, k_{\mathrm{Co}})$ *the* electricity output bundle such as above (i.e., the one maximizing the storage plant's operating profit), if it is unique

s_0 energy stock at time 0 and T (in kWh)

λ unit value of energy stock at time 0 and T (in \$/kWh)

$s(t)$ energy stock at time t (in kWh)

$\varsigma_{h\,\mathrm{St}}$ household's share of profit from supplying the storage capacity

$\psi(t)$ unit value of energy stock at time t (in \$/kWh)

$\hat{\Psi}_{\mathrm{PS}}(p; k_{\mathrm{St}}, k_{\mathrm{Co}})$ the set of all the imputed time-of-use values of energy stock (shadow-price functions for energy stock) in a pumped-storage plant with capacities $(k_{\mathrm{St}}, k_{\mathrm{Co}})$, when the electricity tariff is p

$\hat{\psi}_{\mathrm{PS}}(p; k_{\mathrm{St}}, k_{\mathrm{Co}})$ *the* imputed time-of-use value (shadow price) of energy stock, if it is unique (as a function of time)

Hydro

k_{St} the plant's storage a.k.a. reservoir capacity (in kWh)

$\kappa_{\mathrm{St}}(\mathrm{d}t)$ unit value of storage capacity on a time interval of length $\mathrm{d}t$ (in \$/kWh)

$r_{\mathrm{St}} = \int_0^T \kappa_{\mathrm{St}}(\mathrm{d}t)$ unit value of storage capacity in total for the cycle (in \$/kWh)

$r^\star_{\mathrm{St}}$ the (long-run) equilibrium rental price of storage capacity (in \$/kWh)

$G(k_{\mathrm{St}})$ the supply cost of reservoir of capacity k_{St}

$\nu_{\mathrm{St}}(\mathrm{d}t)$ unit value of nonnegativity constraint on water stock on an interval of length $\mathrm{d}t$ (in \$/kWh)

k_{Tu} the plant's turbine-generator capacity (in kW)

$\kappa_{\mathrm{Tu}}(t)$ unit value of turbine capacity at time t, per unit time (in \$/kWh)

$r_{\mathrm{Tu}} = \int_0^T \kappa_{\mathrm{Tu}}(t)\,\mathrm{d}t$ as the unit value of turbine capacity in total for the cycle (in \$/kW)

$r_{\mathrm{Tu}}^{\mathrm{F}}$ the *given* rental price of turbine capacity (in \$/kW)

$\nu_{\mathrm{Tu}}(t)$ unit value of nonnegativity constraint on turbine's output at time t, per unit time (in \$/kWh)

$e(t)$ rate of river flow at time t (in kW)

$y_{\mathrm{H}}(t)$ rate of electricity output from the hydro plant at time t (in kW)

$\hat{Y}_{\mathrm{H}}(p; k_{\mathrm{St}}, k_{\mathrm{Tu}}; e)$ the set of all the electricity output bundles that maximize the operating profit of a hydro plant with capacities $(k_{\mathrm{St}}, k_{\mathrm{Tu}})$ and river inflow function e when the electricity tariff is p

$\hat{y}_{\mathrm{H}}(p; k_{\mathrm{St}}, k_{\mathrm{Tu}}; e)$ *the* electricity output bundle such as above (i.e., the one maximizing the hydro plant's operating profit), if it is unique

$\sigma(t)$ rate of spillage from the reservoir at time t (in kW)

s_0 water stock at time 0 and T (in kWh)

λ unit value of water stock at time 0 and T (in \$/kWh)

$s(t)$ water stock at time t (in kWh)

$\psi(t)$ unit value of water stock at time t (in \$/kWh)

$\hat{\Psi}_{\mathrm{H}}(p; k_{\mathrm{St}}, k_{\mathrm{Tu}}; e)$ the set of all the imputed time-of-use water values (shadow water-price functions) in a hydro plant with capacities $(k_{\mathrm{St}}, k_{\mathrm{Tu}})$ and river inflow function e, when the electricity tariff is p

$\hat{\psi}_{\mathrm{H}}(p; k_{\mathrm{St}}, k_{\mathrm{Tu}}; e)$ *the* imputed time-of-use water value (shadow price), if it is unique (as a function of time)

$\varsigma_{h\,\mathrm{St}}$ household's share of profit from supplying the reservoir capacity

Specific vector spaces, norms and functionals

meas the Lebesgue measure, on an interval $[0, T]$ of the real line $\mathbb{R}$

$L^1[0, T]$ the space of meas-integrable real-valued functions on $[0, T]$

$L^\infty[0, T]$ the space of essentially bounded real-valued functions on $[0, T]$

$\mathrm{EssSup}(y) = \operatorname{ess\,sup}_{t\in[0,T]} y(t)$ the essential supremum of a $y \in L^\infty[0, T]$

$\|y\|_\infty := \mathrm{EssSup}\,|y|$ the supremum norm on L^∞

$\mathcal{C}[0, T]$ the space of continuous real-valued functions on $[0, T]$

$\mathcal{M}[0, T]$ the space of Borel measures on $[0, T]$

$\int_{[0,T]} s(t)\,\mu(\mathrm{d}t)$ the integral of a continuous function s with respect to a measure μ

ε_t the Dirac measure at t (i.e., a unit mass concentrated at the single point t)

$\mathrm{BV}(0, T)$ the space of functions of bounded variation on $(0, T)$

$\mathrm{Var}^+(\psi)$ the total positive variation (upper variation) of a $\psi \in \mathrm{BV}(0, T)$

$\mathrm{Var}_{\mathrm{c}}^+(\psi) := \mathrm{Var}^+(\psi) + (\psi(0) - \psi(T))^+$ the cyclic positive variation of ψ

Norms and topologies on vector spaces, dual spaces, order and nonnegativity, scalar product

Y^* the norm-dual of a Banach space $(Y, \|\cdot\|)$

$\|\cdot\|^*$ the dual norm on Y^*

Y' the Banach predual of $(Y, \|\cdot\|)$, when Y is a dual Banach space

$\|\cdot\|'$ the predual norm on Y'

Y_+, Y_+^* and Y_+' the nonnegative cones in Y, Y^* and Y' (when these are Banach lattices), e.g., L_+^∞ and L_+^1 are the nonnegative cones in L^∞ and L^1

y^+ and y^- the nonnegative and nonpositive parts of a $y \in Y$ (when Y is a vector lattice)

$k \gg 0$ means that k is a strictly positive vector (in a lattice paired with another one); here, used only with a finite-dimensional k

$\langle \cdot \mid \cdot \rangle$ a bilinear form (scalar product) on the Cartesian product, $P \times Y$, of two vector spaces (when $P = \mathbb{R}^n = Y$, $p \cdot y$ is an alternative notation for the scalar product $\langle p \mid y \rangle := p^{\mathrm{T}} y$, where y is a column vector and p^{T} is a row of the same, finite dimension n)

$\mathrm{w}(Y, P)$ the weak topology on a vector space Y for its pairing with another vector space P (e.g., with Y^* or Y' when Y is a dual Banach space)

$\mathrm{m}(Y, P)$ the Mackey topology on Y for its pairing with P (e.g., with $P = Y^*$ or with $P = Y'$ when Y is a dual Banach space)

w^* and m^* abbreviations for $\mathrm{w}(P^*, P)$ and $\mathrm{m}(P^*, P)$, the weak* and the Mackey topologies on the norm-dual of a Banach space P

bw^* the bounded weak* topology (on a dual Banach space)

$\mathrm{cl}_{Y,\mathcal{T}} Z$ the closure of a set Z relative to a (larger) set Y with a topology $\mathcal{T}$

$\mathrm{int}_{Y,\mathcal{T}} Z$ the interior of a set Z relative to a (larger) set Y with a topology $\mathcal{T}$

Y^{a} the algebraic dual of a vector space Y

$\mathcal{T}_{\mathrm{SLC}} = \mathrm{m}(Y, Y^{\mathrm{a}})$ the strongest locally convex topology on a vector space Y

Sets derived from a set in a vector space

$\mathrm{cone}\, Z$ the cone generated by a subset, Z, of a vector space (i.e., the smallest cone containing Z)

$\mathrm{conv}\, Z$ the convex hull of a subset, Z, of a vector space (i.e., the smallest convex set containing Z)

$\mathrm{cor}\, Z$ the core of a subset, Z, of a vector space

$\mathrm{ext}\, Z$ the set of all the extreme points of a subset, Z, of a vector space

$\mathrm{span}\, Z$ the linear span of a subset, Z, of a vector space

$\mathrm{N}(y \mid Z) = \partial \delta (y \mid Z)$ the outward normal cone to a convex set Z at a point $y \in Z$ (a cone in the dual space)

$\mathrm{N}^{\mathrm{a}}(y \mid Z) = \partial^{\mathrm{a}} \delta (y \mid Z)$ the algebraic normal cone to Z at y (a cone in algebraic dual space); ∂^{a} is the algebraic subdifferential

Sets and functions derived from functions or operations on a vector space

$\mathrm{argmax}_Z f$ means the set of all maximum points of an extended-real-valued function f on a set Z

$\mathrm{dom}\, C$ the effective domain of a convex extended-real-valued function C

$\widehat{\mathrm{dom}} \Pi$ the effective domain of a concave extended-real-valued function Π

$\mathrm{epi}\, C$ the epigraph of a convex extended-real-valued function C (on a vector space)

$\ker A$ the kernel of a linear operation, A

$\mathrm{lsc}\, C$ the lower semicontinuous envelope of C (the greatest l.s.c. minorant of C)

$\mathrm{usc}\, \Pi$ the upper semicontinuous envelope of Π (the least u.s.c. majorant of Π)

$C^{\#}$ the Fenchel-Legendre convex conjugate (of a convex function C)

$\Pi_{\#}$ the concave conjugate (of a concave function Π)

$C^{\#_{1,2}}$, etc. the partial conjugate, of a multi-variate function, w.r.t. all the variables shown (here, w.r.t. the first and the second variables together)

$C' \triangle C''$ the infimal convolution of convex functions, C' and C''

Other notation

card Φ the number of elements in a (finite) set Φ
$\emptyset$ the empty set
1_A the 0-1 indicator function of a set A (equal to 1 on A)
$\liminf$, $\limsup$ respectively, the lower and upper limits (of a real-valued) function)
$\mathbb{R}$ the real line

2. Peak-load pricing with cross-price independent demands

We illustrate the short-run approach to solving for long-run general equilibrium with the example of pricing, over the demand cycle, the services of a homogeneous productive capacity with a unit capital cost r and a unit running cost w. The technology can be interpreted as, e.g., electricity generation from a single type of thermal station with a fuel cost w (in \$/kWh) and a capacity cost r (in \$/kW) per period. The cycle is represented by a continuous time interval $[0, T]$. Demand for the time-differentiated, nonstorable product, $D_t(\mathsf{p})$, is assumed to depend only on the time t and the current price p. As a result, the short-run equilibrium can be found separately at each instant t, by intersecting the demand and supply curves in the price-quantity plane. This is because, with this technology, short-run supply is cross-price independent: given a capacity k, the supply is

$$S(\mathsf{p}, k, w) = \begin{cases} 0 & \text{for } \mathsf{p} < w \\ [0, k] & \text{for } \mathsf{p} = w \\ k & \text{for } \mathsf{p} > w \end{cases} \tag{2.1}$$

where p is the current price. That is, given a time-of-use (TOU) tariff p (i.e., given a price $p(t)$ at each time t), the set of profit-maximizing output trajectories, $\hat{Y}(p, k, w)$, consists of selections from the correspondence $t \mapsto S(p(t), k, w)$. When $D_t(w) > k$, the short-run equilibrium TOU price, $p^{\star}_{\mathrm{SR}}(t, k, w)$, exceeds w by whatever is required to bring the demand down to k (Figure 1a). The total premium over the cycle is the unit operating profit, which in the long run should equal the unit capacity cost r—i.e., the long-run equilibrium capacity, $k^{\star}(r, w)$, can be determined by solving for k the equation

$$r = \int_0^T (p^{\star}_{\mathrm{SR}}(t, k, w) - w)^+ \mathrm{d}t \tag{2.2}$$

where $\pi^+ = \max\{\pi, 0\}$ is the nonnegative part of π (i.e., by equating to r the shaded area in Figure 1b). Put into the short-run equilibrium price function, the equilibrium capacity gives the long-run equilibrium price

$$p^{\star}_{\mathrm{LR}}(t; r, w) = p^{\star}_{\mathrm{SR}}(t, k^{\star}(r, w), w). \tag{2.3}$$

An obvious advantage of this method is that the short-run equilibrium is of interest in itself. Also, the short-run calculations can be very simple, as in this example. For comparison, to calculate the long-run equilibrium directly requires timing the capacity charges so that they are borne entirely by the resulting demand peaks—i.e., it requires finding a density function $\gamma \geq 0$ such that

$$\int_0^T \gamma(t)\,\mathrm{d}t = 1 \quad \text{and if } \gamma(t) > 0 \text{ then } y(t) = \sup_{\tau} y(\tau) \tag{2.4}$$

$$\text{where:} \quad y(t) = D_t(p(t)) \text{ and } \quad p(t) = w + r\gamma(t).$$

This poses a fixed-point problem that, unlike the short-run approach, is not much simplified by cross-price independence of demands.[16]

Since the operating profit is $\Pi_{\mathrm{SR}}(p,k,w) = k\int_0^T (p(t)-w)^+ \mathrm{d}t$, the break-even condition (2.2) can be rewritten as $r = \partial\Pi_{\mathrm{SR}}/\partial k$, i.e., it can be viewed as equating the capital input's price to its profit-imputed marginal value. This is, with any convex technology, the first-order necessary and sufficient condition for a profit-maximizing choice of investment k: together with a choice of output y that maximizes the short-run profit (SRP), such a choice of k maximizes the long-run profit (LRP), and thus turns the short-run equilibrium into the long-run equilibrium.

Furthermore, with any technology and any number of capital inputs, $r = \nabla_k\Pi_{\mathrm{SR}}$ if and only if r is the unique solution to the dual of the SRP maximization programme (and there is no duality gap): this is the derivative property of the optimal value Π_{SR} as a function of the primal parameter k. This identity is useful when, with a more complex technology, the SRP programme has to be solved by a duality method, i.e., together with its the dual. It means that the dual solution $\hat{r}(p,k,w)$, evaluated at the short-run equilibrium output price system $p^\star_{\mathrm{SR}}(k,w)$, can be equated to the capital inputs' *given* prices r to determine their long-run equilibrium quantities $k^\star$.

When the producer is a public utility, competitive profit maximization usually takes the form of marginal-cost pricing. In this context, the equality $r = \partial\Pi_{\mathrm{SR}}/\partial k$, or $r = \nabla_k\Pi_{\mathrm{SR}}$ when there is more than one type of capacity, guarantees that an SRMC price system is actually an LRMC. The result applies to any convex technology—even when the costs are nondifferentiable, and the marginal cost has to be defined by using the subdifferential as a generalized, multi-valued derivative. This is so in the above example of capacity pricing, since the long-run cost

$$C_{\mathrm{LR}}(y(\cdot),r,w) = w\int_0^T y(t)\,\mathrm{d}t + r\sup_{t\in[0,T]} y(t) \tag{2.5}$$

is nondifferentiable if the output y has multiple peaks: indeed, for every γ satisfying (2.4), the function $p = w + r\gamma$ represents a subgradient of C_{LR} with respect to y (w.r.t. y). And multiple peaks are more of a rule than an exception in equilibrium (note the peak output plateau in Figure 1d here, and see [26] for an extension to the case of cross-price dependent demands). Similarly, the short-run cost

$$C_{\mathrm{SR}}(y(\cdot),k,w) = \begin{cases} w\int_0^T y(t)\,\mathrm{d}t & \text{if } 0 \le y \le k \\ +\infty & \text{otherwise} \end{cases} \tag{2.6}$$

is nondifferentiable if $\sup_t y(t) = k$. In Figure 1a, the nondifferentiability shows in the (infinite) vertical interval $[w,+\infty)$ that represents the multi-valued instantaneous SRMC at $\mathsf{y} = k$.[17] In Figure 1c, it shows as a kink, at $\mathsf{y} = k$, in the graph of the instantaneous

[16] In terms of the subdifferential, ∂C, of the long-run cost (2.5) as a function of output, the fixed-point problem is to find a function p such that $p \in \partial C_{\mathrm{LR}}(D(p))$, where $D(p)(t) = D_t(p(t))$ if demands are cross-price independent.

[17] The SRMC and the short-run supply correspondences are inverse to each other, i.e., have the same graph: in Figure 1a, the broken line is both the supply curve and the SRMC curve.

cost function

$$c_{\mathrm{SR}}(\mathsf{y}) = \begin{cases} w\mathsf{y} & \text{if } 0 \le \mathsf{y} \le k \\ +\infty & \text{otherwise} \end{cases} \tag{2.7}$$

(which gives $C_{\mathrm{SR}}(y)$ as $\int_0^T c_{\mathrm{SR}}(y(t))\,\mathrm{d}t$, so that a TOU price p is an SRMC at y if and only if $p(t)$ is an instantaneous SRMC at $y(t)$ for each t). With this technology, C_{SR} is therefore nondifferentiable whenever k is the cost-minimizing capital input for the required output y: cost-optimality of k means merely that it provides just enough capacity, i.e., that $k = \mathrm{Sup}(y)$. This condition, being quite unrelated to the input prices r and w, obviously cannot ensure that an SRMC price system is an LRMC. To guarantee this, one must strengthen it to the condition that $r = \int_0^T (p(t) - w)^+ \mathrm{d}t$ in this example or, generally, that $r = \nabla_k \Pi_{\mathrm{SR}}$ (or that r belongs to the supergradient set $\widehat{\partial}_k \Pi_{\mathrm{SR}}(p, k, w)$, should Π_{SR} be nondifferentiable in k).[18] The capital's cost-optimality would suffice for the SRMC to be the LRMC if the costs were differentiable; this is the Wong-Viner Envelope Theorem. The preceding remarks show how to reformulate it to free it from differentiability assumptions. This is detailed in Section 11.

Cross-price independent demand arises from price-taking optimization by consumers and industrial users with additively separable utility and production functions. In this case, the short-run equilibrium prices can readily be given in terms of the marginal utility of the differentiated good (and its productivity if there are industrial users). For the simplest illustration, all demand is assumed to come from a single household maximizing the utility function

$$U(x(\cdot), m) = m + \int_0^T u(t, x(t))\,\mathrm{d}t$$

over $x(\cdot) \ge 0$ and $m \ge 0$, the consumptions of the nonstorable good and the numeraire, subject to the budget constraint

$$m + \int_0^T p(t)\, x(t)\,\mathrm{d}t \le M$$

where M is the income and $p(\cdot)$ is a TOU price in terms of the numeraire (which represents all the other goods and thus closes the model). For each t, the instantaneous utility $u(t, \mathsf{x})$ is taken to be a strictly concave, increasing and differentiable function of the consumption rate $\mathsf{x} \in \mathbb{R}_+$, with $(\partial u / \partial \mathsf{x})(t, 0) > w$ (to ensure that, in a short-run equilibrium with $k > 0$, consumption is positive at every t). The income consists of an endowment of the numeraire (m^{En}) and the pure profit from electricity sales, i.e.,

$$M = m^{\mathrm{En}} + k \int_0^T (p(t) - w)^+ \,\mathrm{d}t - rk.$$

To guarantee a positive demand for the numeraire, assume that $m^{\mathrm{En}} > (Tw + r)\,k$. Then, at any time t, demand (for the good) depends only on the current price $p(t)$, and it is determined from the equation

$$\frac{\partial u}{\partial \mathsf{x}}(t, x(t)) = p(t).$$

[18] This condition ($r = \nabla_k \Pi_{\mathrm{SR}}$) is stronger than cost-optimality of the fixed inputs when p is an SRMC.

In other words, $D_t(\mathsf{p}) = ((\partial u/\partial \mathsf{x})(t,\cdot))^{-1}(\mathsf{p})$. When $w < (\partial u/\partial \mathsf{x})(t,k)$, this value of $\partial u/\partial \mathsf{x}$ is the price needed to equate demand to k. So the short-run equilibrium price is

$$p^{\star}_{\mathrm{SR}}(t,k,w) = w + \left(\frac{\partial u}{\partial \mathsf{x}}(t,k) - w\right)^{+}. \tag{2.8}$$

By (2.2) and (2.3), the long-run equilibrium capacity $k^{\star}(r,w)$ is determined from

$$r = \int_0^T \left(\frac{\partial u}{\partial \mathsf{x}}(t,k) - w\right)^{+} \mathrm{d}t$$

and the long-run equilibrium price is, in terms of $k^{\star}$,

$$p^{\star}_{\mathrm{LR}}(t,r,w) = w + \left(\frac{\partial u}{\partial \mathsf{x}}(t,k^{\star}(r,w)) - w\right)^{+}. \tag{2.9}$$

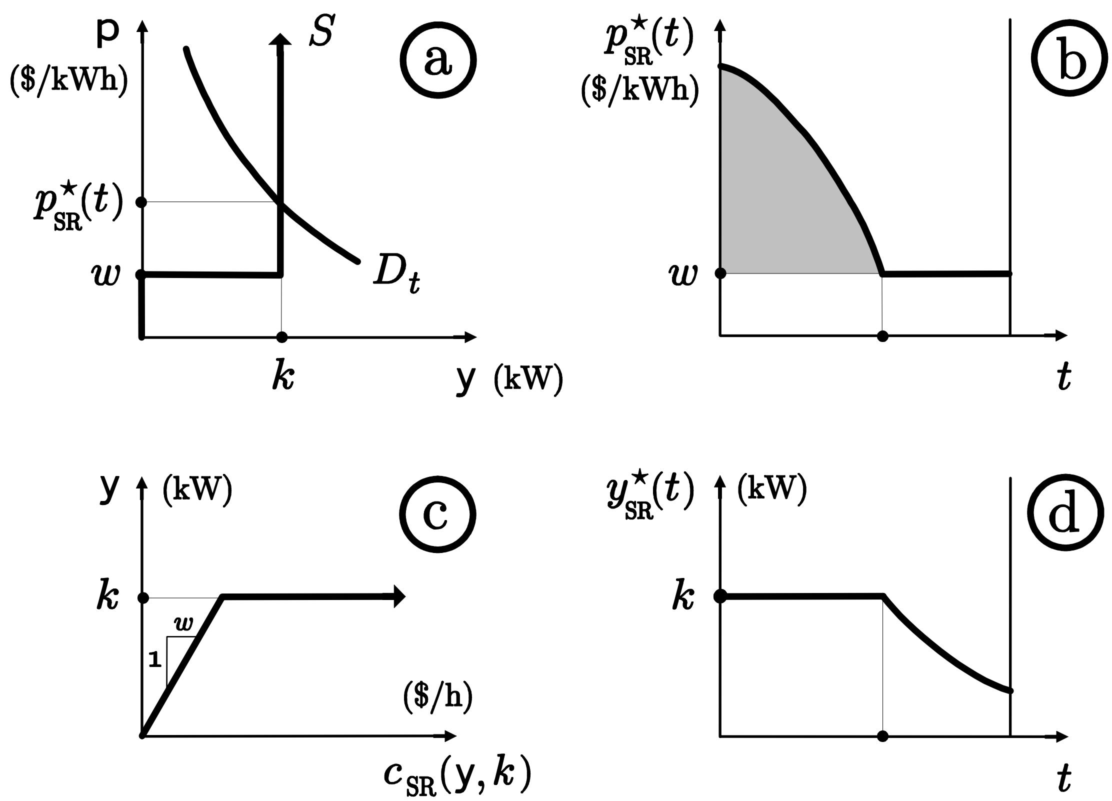

FIGURE 1. Short-run approach to long-run equilibrium of supply and (cross-price independent) demand for thermally generated electricity: (a) determination of the SR equilibrium price and output for each instant t, given a capacity k; (b) and (d) trajectories of the SR price and output; (c) the SR cost curve. When k is such that the shaded area in (b) equals r, the SR equilibrium is the LR equilibrium.

3. Cost and profit as values of programmes with quantity decisions

Costs and profits of a price-taking producer are, by definition, the optimal values of programmes with quantities as decision variables. With several variables, it can be much easier to solve the mathematical problem in stages, by fixing some variables and dealing with the resulting subproblem first. The subproblem may also be of independent interest, especially if it corresponds to a stage in a practical implementation of a complete solution. In production, the decision on plant operation (with fixed investment) corresponds to short-run profit maximization as a subproblem of long-run profit maximization: although operation is usually planned along with investment, the producer is still free to make operating decisions after constructing the plant. In other words, his final choices of the outputs y and the variable inputs v are made only after fixing the capital inputs k. Such a multi-stage problem can be solved in the reverse order: this means that the decisions to be implemented last are determined first but are made contingent on the decisions to be implemented earlier, and the complete solution is put together by back substitution. For the producer, this means first choosing y and v to maximize short-run profit, given an *arbitrary* k as well as the prices, p and w, for the variable commodities. Even within the confines of the purely periodic (or static) problems considered here, this approach has a couple of analytical advantages. First, in addition to being of independent interest, the short-run equilibrium (given k) can be much easier to find than the long-run equilibrium, as in Section 2. Second, when there is a number of technologies, the short-run equilibrium is usually much easier to find by solving the profit maximization programmes (to determine the total short-run supply and equate it to demand) than by solving the duals of cost minimization programmes (to determine the SRMCs, which would have to be equated both to one another and to the inverse demand). This *profit approach* is simpler than the cost approach in two ways, viz., by giving unique solutions to the producer problem with its dual, and by reducing the number of unknowns in the subsequent equilibrium problem: see Section 12.

A third advantage of the short-run approach emerges only when the framework, unlike this one, takes account of non-periodic demand and price uncertainty. The prices for the variable commodities (p, w), or their probability distribution in a stochastic model, will change in unforeseen ways between the planning and the completion of plants, and will also keep shifting thereafter. As a result, both the plant mix and the design of individual plants will become suboptimal. But whether a plant is optimal or not, it should be optimally operated, and a solution to this problem is part of the short-run approach.

It is the above considerations that make short-run profit maximization the subproblem of central interest to us. It, too, may be solved in two stages, though this time the order in which the decision variables (y and v) are determined is only a matter of convenience: it is usually best to start with the simpler subproblem. Here, it is assumed that short-run cost minimization (finding v given k and y) is easier than revenue maximization (finding y given k and v). The solution sequence (first v, then y and finally k) corresponds to a chain of three problems: (i) the "small" one of short-run cost minimization (with k and y as data, v as a decision), (ii) an "intermediate" problem of short-run profit maximization (with k as a datum,and y and v as decisions), and (iii) the "large" problem of long-run profit maximization (with k, y and v as decision variables).

A fourth problem, another intermediate one, is that of long-run cost minimization (with y as a datum, k and v as decision variables). It is in terms of this problem and its value function that that public utilities usually formulate their welfare-promoting principles of meeting the demand at a minimum operating cost, optimizing their capital stocks, and pricing their outputs at LRMC. Together, these policies result in long-run profit maximization and competitive equilibrium in the products' markets. Although the separate aims are stated in terms of long-run costs (as LRMC pricing and LRC minimization), their combination is best achieved through short-run calculations—for the reasons outlined above and detailed in Section 12.

Each of the four problems, when formulated as one of optimization constrained by a convex *production set* $\mathbb{Y}$, has a linear objective function.[19] This has several implications. One is that each problem (SRC or LRC minimization, or SRP or LRP maximization) can be formulated as a linear programme (LP), by representing $\mathbb{Y}$ as the intersection of a finite or infinite set of half-spaces; this is discussed further in Section 14. What matters for now is that in passing to a subproblem, once a decision variable has become a datum (like k in passing from long to short run), the corresponding term of the linear optimand ($r \cdot k$) can be dropped, since it is fixed. Its coefficient (r) can then be removed from the subproblem's data (which include k).[20]

The *commodity spaces* for outputs, fixed inputs and variable inputs are denoted by Y, K and V, respectively. These are paired with *price spaces* P, R and W by bilinear forms (a.k.a. scalar products) denoted by $\langle p \mid y \rangle$, etc.; the alternative notation $p \cdot y$ is employed to mean $p^{\mathrm{T}} y$ when both P and Y are equal to the finite-dimensional space $\mathbb{R}^n$ (where p^{T} is the row vector obtained by transposing a column p). Unless specified, the range of a decision variable (say y) is the whole space (Y).

With p, r and w denoting the prices for outputs, fixed inputs and variable inputs (y, k and v, respectively), the *long-run profit maximization* programme is:

$$\text{(3.1)} \qquad \text{Given } (p, r, w)\text{, maximize } \langle p \mid y \rangle - \langle r \mid k \rangle - \langle w \mid v \rangle \text{ over } (y, k, v)$$

$$\text{(3.2)} \qquad \text{subject to } (y, -k, -v) \in \mathbb{Y}.$$

Its optimal value, the maximum LRP as a function of the data, is denoted by $\Pi_{\mathrm{LR}}(p, r, w)$. By definition, (y, k, v) solves (3.1)–(3.2) if and only if

$$\text{(3.3)} \qquad (y, -k, -v) \in \mathbb{Y} \text{ and } \langle p, r, w \mid y, -k, -v \rangle = \Pi_{\mathrm{LR}}(p, r, w).$$

In the central case of constant returns to scale (c.r.t.s.), the production set $\mathbb{Y}$ is a cone, and Π_{LR} is the 0-∞ indicator of the *polar cone*

$$\text{(3.4)} \qquad \mathbb{Y}^{\circ} = \{(p, r, w) : \forall\, (y, -k, -v) \in \mathbb{Y} \;\; \langle p \mid y \rangle - \langle r \mid k \rangle - \langle w \mid v \rangle \leq 0\}$$

i.e., $\Pi_{\mathrm{LR}}(p, r, w)$ is 0 if $(p, r, w) \in \mathbb{Y}^{\circ}$, and it is $+\infty$ otherwise. Condition (3.3) is then equivalent to the conjunction of technological feasibility, price consistency and break-even

[19] Even if the objective were nonlinear, it could always be replaced by a linear one with an extra scalar variable, subject to an extra nonlinear constraint: as is noted in [12, p. 48], minimization of $f(y)$ over y is equivalent to minimization of ρ over y and ρ subject to $\rho \geq f(y)$ in addition to any original constraints on y.

[20] More generally, this is so whenever the optimand separates into a function of (r, k) plus terms independent of r and k.

conditions, which make up the *Complementarity Conditions*

$$(y,-k,-v)\in\mathbb{Y},\ (p,r,w)\in\mathbb{Y}^{\circ}\text{ and }\langle p,r,w\,|\,y,-k,-v\rangle=0. \tag{3.5}$$

One subprogramme of (3.1)–(3.2) is short-run profit maximization, i.e.,

$$\text{Given } (p,k,w)\,,\text{ maximize } \langle p\,|\,y\rangle-\langle w\,|\,v\rangle \text{ over } (y,v) \tag{3.6}$$

$$\text{subject to } (y,-k,-v)\in\mathbb{Y}. \tag{3.7}$$

Its optimal value is $\Pi_{\mathrm{SR}}(p,k,w)$, the maximum SRP.

Another subprogramme of (3.1)–(3.2) is long-run cost minimization, i.e.,

$$\text{Given } (y,r,w)\,,\text{ minimize } \langle r\,|\,k\rangle+\langle w\,|\,v\rangle \text{ over } (k,v) \tag{3.8}$$

$$\text{subject to } (y,-k,-v)\in\mathbb{Y}. \tag{3.9}$$

Its optimal value is $C_{\mathrm{LR}}(y,r,w)$, the minimum LRC.

The common subprogramme (of all three of the above) is short-run cost minimization, i.e.,

$$\text{Given } (y,k,w)\,,\text{ minimize } \langle w\,|\,v\rangle \text{ over } v \tag{3.10}$$

$$\text{subject to } (y,-k,-v)\in\mathbb{Y}. \tag{3.11}$$

Its optimal value is $C_{\mathrm{SR}}(y,k,w)$, the minimum SRC.

Partial conjugacy relationships between the value functions (Π_{LR}, Π_{SR}, C_{LR}, C_{SR}) are summarized in the following diagram:

$$\boxed{\begin{array}{ccccccccc}
 & & & & w & & & & \\
 & & & & \Pi_{\mathrm{LR}} & & & & \\
 & & & r & & p & & & \\
 & & \nearrow & & & & \nwarrow & & \\
 & & k & & & & y & & \\
w & \Pi_{\mathrm{SR}} & & & & & & C_{\mathrm{LR}} & w \\
 & & p & & & & r & & \\
 & & \nwarrow & & & & \nearrow & & \\
 & & & y & & k & & & \\
 & & & & C_{\mathrm{SR}} & & & & \\
 & & & & w & & & &
\end{array}}\,. \tag{3.12}$$

For example, the arrow from the y next to C_{SR} to the p next to Π_{SR} indicates that Π_{SR} is, as a function of p, the Fenchel-Legendre convex conjugate of C_{SR} as a function of y, with (k,w) fixed; i.e., by definition,

$$\Pi_{\mathrm{SR}}(p,k,w)=\sup_{y}\left\{\langle p\,|\,y\rangle-C_{\mathrm{SR}}(y,k,w)\right\}. \tag{3.13}$$

Similarly, $-\Pi_{\mathrm{LR}}$ is, as a function of r, the concave conjugate of Π_{SR} as a function of k, with (p,w) fixed; i.e.,

$$\Pi_{\mathrm{LR}}(p,r,w) = \sup_k \{\Pi_{\mathrm{SR}}(p,r,w) - \langle r \mid k \rangle\}. \tag{3.14}$$

The right half of the diagram (3.12) represents similar links between C_{LR} and C_{SR} or Π_{LR}. Details such as the signs and convexity or concavity are omitted.

As is spelt out next, those y and k which yield the suprema in (3.13) and (3.14) are parts of an input-output bundle that maximizes the long-run profit.

4. A primal-dual optimization system for the short-run approach

A joint programme for two or more decision variables can be split by optimizing in stages: first over a subset of the variables (keeping the rest fixed), then over the other variables (the optimand comprising the value function from the first stage) to obtain the complete solution by back substitution. The method can be applied to solve the LRP maximization programme (3.1)–(3.2) for (y,k,v) by:

(1) first minimizing $\langle w \mid v \rangle$ over v to find the solution set $\check{V}(y,k,w)$, or *the* solution $\check{v}(y,k,w)$ if it is indeed unique, and the minimum value $C_{\mathrm{SR}}(y,k,w)$, which is $\langle w \mid \check{v} \rangle$;

(2) then maximizing $\langle p \mid y \rangle - C_{\mathrm{SR}}(y,k,w)$ over y to find the solution set $\hat{Y}(p,k,w)$, or *the* solution $\hat{y}(p,k,w)$ if it is unique, and the maximum value $\Pi_{\mathrm{SR}}(p,k,w)$, which is $\langle p \mid \hat{y} \rangle - C_{\mathrm{SR}}(\hat{y})$;

(3) and finally, maximizing $\Pi_{\mathrm{SR}}(p,k,w) - \langle r \mid k \rangle$ over k to find the solution set $\hat{K}(p,r,w)$, or *the* solution $\hat{k}(p,r,w)$, should it be unique (which it obviously cannot be if returns to scale are constant, in the long run).

Every complete solution can then be given, in terms of p, r and w, as a triple $(y,-k,-v)$ such that: $k \in \hat{K}(p,r,w)$, $y \in \hat{Y}(p,k,w)$ and $v \in \check{V}(y,k,w)$. With decreasing returns to scale, if the solution is unique, it is the triple: $\hat{k}(p,r,w)$, $\hat{y}\left(p,\hat{k}(p,r,w),w\right)$ and $\check{v}\left(\hat{y}\left(p,\hat{k}(p,r,w),w\right),\hat{k}(p,r,w),w\right)$.

In other words, the LRP programme (3.1)–(3.2) for (y,k,v) can be reduced to an investment programme, for k alone, by first solving the SRP programme (3.6)–(3.7) for (y,v) and substituting its optimal value (Π_{SR}) for the term $\langle p \mid y \rangle - \langle w \mid v \rangle$ in (3.1). The SRP programme for (y,v) can, in turn, be reduced to a programme for y alone by solving the SRC programme (3.10)–(3.11) and substituting its value (C_{SR}) for the term $\langle w \mid v \rangle$ in (3.6).

So an input-output bundle $(y,-k,-v)$ maximizes long-run profit at prices (p,r,w) if and only if both

$$k \text{ maximizes } \Pi_{\mathrm{SR}}(p,\cdot,w) - \langle r \mid \cdot \rangle \text{ on } K \text{ (given } p, r \text{ and } w) \tag{4.1}$$

and the bundle $(y,-v)$ maximizes short-run profit (given k) at prices (p,w) or, equivalently,

$$y \text{ maximizes } \langle p \mid \cdot \rangle - C_{\mathrm{SR}}(\cdot,k,w) \text{ on } Y \text{ (given } p, k \text{ and } w) \tag{4.2}$$

$$v \text{ minimizes } \langle w \mid \cdot \rangle \text{ on } \{v \in V : (y,-k,-v) \in \mathbb{Y}\} \text{ (given } y, k \text{ and } w). \tag{4.3}$$

We call (4.1)–(4.3) the *split LRP optimization system*. Its SRC subprogramme for v in (4.3) is taken to be readily soluble. By contrast, the reduced SRP programme for y in (4.2) may require a duality approach. This consists in pricing the constraining parameters and solving the dual programme of valuation together with the primal (when there is no duality gap). For the SRP programme as the primal, this means valuing the fixed inputs k: a dual solution (with no gap) is a shadow-price system r such that

$$r \text{ minimizes } \langle \cdot \mid k \rangle + \Pi_{\mathrm{LR}}(p, \cdot, w) \text{ on } R \text{ (given } p, k \text{ and } w) \tag{4.4}$$

$$\text{and the minimum value, } \langle r \mid k \rangle + \Pi_{\mathrm{LR}}(p, r, w)\text{, equals } \Pi_{\mathrm{SR}}(p, k, w)\,. \tag{4.5}$$

Under c.r.t.s., Conditions (4.4)–(4.5) become

$$r \text{ minimizes } \langle \cdot \mid k \rangle \text{ on } \{r \in R : (p, r, w) \in \mathbb{Y}^{\circ}\} \text{ (given } p, k \text{ and } w) \tag{4.6}$$

$$\text{and the minimum value, } \langle r \mid k \rangle\text{, equals } \Pi_{\mathrm{SR}}(p, k, w)\,. \tag{4.7}$$

The duality scheme that produces the programme in (4.6) or (4.4) as the dual to SRP maximization is set out in Section 5.

As well as helping solve the operation problem, the dual solution can be used to check the investment for optimality, i.e., (4.1) is equivalent to (4.4)–(4.5). Formally, this follows from the definitional conjugacy relationship (3.14) between Π_{SR} and Π_{LR} (as functions of k and r) by using the first-order condition (C.24) and the Inversion Rule (C.32), given in Appendix C. The system (4.2)–(4.5) is therefore equivalent to (4.1)–(4.3), and hence also to LRP maximization (3.3), and to Complementarity (3.5) under c.r.t.s. It is, however, put entirely in terms of solutions to the SRP programme for (y, v) and its dual for r, with the primal split into the SRC programme (for v) and the reduced SRP programme (for y). We therefore call (4.2)–(4.5) the *split SRP optimization system*. It is likely to be the best basis for the short-run approach when the technology is specified by means of a production set. Alternative systems are presented in Sections 6 to 8 and 10.

5. Cost and profit as values of programmes with price decisions

Unless there are duality gaps, short-run and long-run cost and profit are also the optimal values of programmes that are dual to those of Section 3. The scheme producing the duals is an application of the usual duality framework for convex programmes (CPs), expounded in, e.g., [44] and [36, Chapter 7]. However, ours starts not from a single programme but from a family of programmes that depend on a set of data, whose particular values complete the programme's specification. One way to perturb the programme is simply to add an increment to its data point, thus "shifting" it within the given family. Some, possibly all, of the scheme's primal perturbations are therefore increments to some—though typically not all—of the data. The same goes for the dual perturbations.

Before applying the duality scheme to the profit and cost programmes, we discuss it briefly and illustrate it in the framework of linear programming. A central idea is that the dual programme depends on the choice of perturbations of the primal programme; different perturbation schemes produce different duals. Theoretical expositions usually start from a programme without any data variables whose increments might serve as primal perturbations: say, $f(y)$ is to be maximized over y subject to $G(y) \leq 0$. In such a case, any perturbations must first be introduced, and the standard choice is to add $\epsilon = (\epsilon_1, \epsilon_2, \ldots)$ to the zeros on the r.h.s.'s, thus perturbing the original constraints

$G(y) \leq 0$ to $G(y) \leq \epsilon$. The original programme has no data other than the functions f and G themselves, and the increments Δf and ΔG (which change the programme to maximization of $(f + \Delta f)(y)$ over y subject to $(G + \Delta G)(y) \leq 0$) could never serve as primal perturbations—not even if they were taken to be linear, i.e., if f and G were a vector and a matrix of coefficients of the primal variables, $y = (y_1, y_2, \ldots)$. This is because the perturbed constrained maximand must be *jointly* concave in the decision variables *and* the perturbations,[21] but the bilinear form $f \cdot y$ is not concave (or convex) in f and y jointly.[22]

But in applications, increments to some of the programme's data can commonly serve as primal perturbations. We call those data the *intrinsic primal parameters*; some or all of the other data will turn out to be dual parameters. For example, in SRP maximization (3.6)–(3.7), the fixed-input bundle k is a primal parameter because, since the production set $\mathbb{Y}$ is convex, the constrained maximand is a concave function of (y, k, v): it is

$$\langle p \,|\, y \rangle - \langle w \,|\, v \rangle - \delta(y, -k, -v \mid \mathbb{Y})$$

where $\delta(\cdot, \cdot, \cdot \mid \mathbb{Y})$ denotes the 0-∞ indicator of $\mathbb{Y}$ (i.e., it equals 0 on $\mathbb{Y}$ and $+\infty$ outside of $\mathbb{Y}$). By contrast, the coefficient (say, p) of a primal variable (y) is *not* a primal parameter (i.e., its increment Δp cannot be a primal perturbation) because the bilinear form $\langle p \,|\, y \rangle$ is *not jointly* concave in p and y. For these reasons, all of the quantity data, but no price data, are primal parameters for the profit or cost optimization programmes of Section 3. As for the production set, it cannot itself serve as a parameter because convex sets do not form a vector space to begin with. However, once the technological constraint $(y, -k, -v) \in \mathbb{Y}$ has been represented in the form $Ay - Bk - Cv \leq 0$ (under c.r.t.s.), the matrices or, more generally, the linear operations A, B and C *are* vectorial data. But none can be a primal parameter, for lack of joint convexity of Ay in A and y, etc. Nor can A, B or C be a dual parameter (for a similar reason). Such data variables, which are neither primal nor dual parameters, and hence play no role in the duality scheme, we call *tertial* parameters.

It can be analytically useful, or indeed necessary, to introduce other primal perturbations, i.e., perturbations that are not increments to any of the data (which are listed after "Given" in the original programme). This amounts to introducing additional parameters, which we call *extrinsic*; their original, unperturbed values can be set as zeros, as in [44]. When the constraint set is represented by a system of inequalities and equalities, the standard "right-hand side" parameters are always available for this purpose (unless they are all intrinsic, but this is so only when the r.h.s. of each constraint is a separate datum of the programme and can therefore be varied independently of the other r.h. sides). In Section 14, we show how to relate the marginal effects of any other, "nonstandard" perturbations to those of the standard ones—i.e., how to express any "nonstandard" dual variables in terms of the usual Lagrange multipliers for the constraints. This is useful in

[21] This is equivalent to joint convexity of the constrained minimand, which is the sum of the minimand and the 0-∞ indicator function of the constraint set. In [44] it is called "the minimand" for brevity.

[22] After a linear change of variables, it becomes a saddle function: $4f \cdot y = (f + y) \cdot (f + y) - (f - y) \cdot (f - y)$ is convex in $f + y$ and concave in $f - y$.

the problems of plant operation and valuation, including those that arise in peak-load pricing (Section 16).[23]

Once a primal perturbation scheme has been fully defined, the framework is completed automatically (except for the choice of topologies and the continuous-dual spaces in the infinite-dimensional case): dual variables are introduced and paired to the specified primal perturbations (both the intrinsic and any extrinsic ones). The corresponding dual match is set up in reverse: to be paired with the primal variables, dual perturbations are introduced. Some or possibly all of these perturb the dual just like increments to some of the original programme's data—which are thus identified as the *intrinsic dual parameters*. Any other dual perturbations are called extrinsic, and these can be thought of as increments to *extrinsic dual parameters* (whose original, unperturbed values are set as zeros). However, in the profit or cost programmes, all the dual parameters are price data (and are therefore intrinsic).

In the *reduced* formulations of the profit or cost problems, some of the price data are *not* dual parameters because the corresponding quantities have been solved for in the reduction process, and have thus ceased to be decision variables: e.g., the variable-input price w is *not* a dual parameter of the reduced SRP programme in (4.2) because the corresponding input bundle v has been found in SRC minimization (and so it is no longer a decision variable). But in the *full* (i.e., non-reduced) formulations, all the price data are dual parameters, and thus the programme's data (other than the technology itself) are partitioned into the primal parameters (the quantity data) and dual parameters (the price data).

The primal and dual optimal values can differ at some "degenerate" parameter points (see Appendix A), but such *duality gaps* are exceptional, and they do not occur when the primal or dual value is semicontinuous in, respectively, the primal or dual parameters (Section 20). Note that both optimal values, primal and dual, depends on the data, which are the same for both programmes. So, in this scheme, each of the optimal values (primal and dual) is a function of both the primal and the dual parameters), and can have two varieties of continuity and differentiability properties:

(1) Properties of Type One are those of the primal value with respect to the primal parameters, and of the dual value w.r.t. the dual parameters.
(2) Properties of Type Two are those of the primal value w.r.t. the dual parameters, and of the dual value w.r.t. the primal parameters.

This distinction cannot be articulated when, as in [44] and [36], the primal and dual values are considered only as functions of either the primal or the dual parameters, respectively.

[23] In this as in other contexts, it can be convenient to think of extrinsic perturbations either as complementing the intrinsic perturbations (which are increments to the fixed inputs) by varying some aspects of the technology (such as nonnegativity constraints), or as replacing the intrinsic perturbations with finer, more varied increments (to the fixed inputs). For example, the time-constant capacity k_θ in (16.3) is an intrinsic primal parameter. The corresponding perturbation is a constant increment Δk_θ, and this can be refined to a time-varying increment $\Delta k_\theta(\cdot)$. This perturbation (Δk_θ or $\Delta k_\theta(\cdot)$) is complemented by the increment $\Delta n_\theta(\cdot)$ to the zero floor for the output rate $y_\theta(\cdot)$ in (16.3). The same goes for all the occurrences of Δk and Δn in the context of pumped storage and hydro, where $\Delta\zeta$ is another complementary extrinsic perturbation.

Comments (parameters and their marginal values, dual programme and FFE Conditions, the Lagrangian and Kuhn-Tucker Conditions for LPs):

(1) Let the primal linear programme be: Given any vectors p and s (and a matrix A), maximize $p \cdot y$ over y subject to $Ay \le s$. Here, the only intrinsic primal parameter is the standard parameter s. There is no obviously useful candidate for an extrinsic primal parameter, and if none is introduced, then the dual is the *standard dual* LP: Given p and s (and A), minimize $\sigma \cdot s$ over $\sigma \ge 0$ subject to $A^{\mathrm{T}}\sigma = p$, where A^{T} is the transpose of A.[24] The only dual parameter is p. If both programmes have unique solutions, $\hat{y}(s, p, A)$ and $\hat{\sigma}(s, p, A)$, with equal values $\mathcal{V}(s, p, A) := p \cdot \hat{y} = \hat{\sigma} \cdot s =: \overline{\mathcal{V}}(s, p, A)$, then the marginal values of all the parameters, including the tertial (non-primal, non-dual) parameter A, exist as ordinary derivatives. Namely: (i) $\nabla_s \mathcal{V} = \nabla_s \overline{\mathcal{V}} = \hat{\sigma}$, (ii) $\nabla_p \mathcal{V} = \nabla_p \overline{\mathcal{V}} = \hat{y}$, and (iii) $\nabla_A \mathcal{V} = \nabla_A \overline{\mathcal{V}} = -\hat{\sigma} \otimes \hat{y} = -\hat{\sigma}\hat{y}^{\mathrm{T}}$ (the matrix product of a column and a row, in this order, i.e., the tensor product), where ∇_A is arranged in a matrix like A (i.e., $\partial \mathcal{V}/\partial_{A_{ij}} = -\hat{\sigma}_i \hat{y}_j$ for each i and j). The first two formulae (for $\nabla_s \mathcal{V}$ and $\nabla_p \mathcal{V}$) are cases of a general derivative property of the optimal value in convex programming: see, e.g., [44, Theorem 16: (b) and (a)] or [32, 7.3: Theorem 1']. Heuristically, the third formula follows from each of the first two by comparing the marginal effect of A with that of either s or p on the constraints (primal or dual). It can also be proved formally by applying the Generalized Envelope Theorem for smooth optimization [1, (10.8)],[25] whereby each marginal value ($\nabla_s \mathcal{V}$, $\nabla_p \mathcal{V}$ and $\nabla_A \mathcal{V}$) is equal to the corresponding partial derivative of the Lagrangian, which is here

$$\mathcal{L}(y, \sigma; p, s; A) := \begin{cases} p \cdot y + \sigma^{\mathrm{T}}(s - Ay) & \text{if } \sigma \ge 0 \\ +\infty & \text{if } \sigma \not\ge 0 \end{cases} . \tag{5.1}$$

(2) The Kuhn-Tucker Conditions form the system

$$\sigma \ge 0, \ Ay \le s, \ \sigma^{\mathrm{T}}(Ay - s) = 0 \quad \text{and} \quad p^{\mathrm{T}} = \sigma^{\mathrm{T}} A \tag{5.2}$$

which, because of the quadratic term $\sigma^{\mathrm{T}} A y$, is nonlinear in the decision variables (y and σ).

(3) But the FFE Conditions (primal feasibility, dual feasibility and equality of the primal and dual objectives) form the system

$$Ay \le s, \ \sigma \ge 0, \ p^{\mathrm{T}} = \sigma^{\mathrm{T}} A \quad \text{and} \quad p \cdot y = \sigma \cdot s \tag{5.3}$$

which *is* linear (in y and σ). This makes it easier to solve than the Kuhn-Tucker system (5.2). For an LP, the FFE system is effective because the dual programme can be worked out from the primal explicitly.

(4) For a general CP, the dual cannot be given explicitly (i.e., without leaving an unevaluated extremum in the formula for the dual constrained objective function

[24] The dual constraint must be changed to $A^{\mathrm{T}}\sigma \ge p$ if $y \ge 0$ is adjoined as another primal constraint. (In that case, the primal LP may be interpreted as, e.g., revenue maximization given a resource bundle s, an output-price system p and a Leontief technology defined by an input-coefficient matrix A.)

[25] Without a proof of value differentiability, the Generalized Envelope Theorem is also given in, e.g., [47, 1.F.b].

in terms of the Lagrangian).[26] That is why the Kuhn-Tucker system is better as a general solution method than the FFE system, although the latter is simpler in some specific cases (such as linear programming). The FFE system requires forming the dual from the primal to start with, but the Kuhn-Tucker system requires only the Lagrangian. It offers a workable method of solving the programme pair, and this matters more than an explicit expression for the dual programme. However, as with an LP, the FFE system can be simpler with a specific CP that has an explicit dual.

The duality scheme is next applied to all four of the profit and cost programmes of Section 3; the one of most importance for our applications is the dual to SRP maximization. The duals are shown to consist in shadow-pricing the given quantities, so their subprogramme relationship is the reverse of that between the primals: the more quantities that are fixed, the more commodities there are to shadow-price. (In other words, the fewer primal variables, the more primal parameters, and hence the more dual variables.) For this reason, the duals are listed, below, in the reverse order to that of the primals (listed in Section 3). See also Figure 2, in which the large single arrows point from primal programmes to their subprogrammes, and the double arrows point from the dual programmes to their subprogrammes. Each of the four middle boxes gives the data for the pair of programmes represented by the two adjacent boxes (the outer box for the primal and the inner box for the dual); the data are partitioned into the primal parameters (the given quantities) and the dual parameters (the given prices). There are no other parameters in this scheme (i.e., it has no extrinsic parameters).

In the SRC minimization programme (3.10)–(3.11), only y and k can serve as primal parameters;[27] and perturbation by both increments, Δy and Δk, yields the following dual programme for shadow-pricing both the outputs and the fixed inputs:

$$\text{Given } (y,k,w)\,,\ \text{maximize } \langle p\,|\,y\rangle - \langle r\,|\,k\rangle - \Pi_{\mathrm{LR}}(p,r,w) \text{ over } (p,r)\,. \tag{5.4}$$

Its optimal value is denoted by $\underline{C}_{\mathrm{SR}}(y,k,w) \leq C_{\mathrm{SR}}(y,k,w)$, with equality when Section 21 applies. The dual parameter is w.

In the LRC minimization programme (3.8)–(3.9), only y can serve as a primal parameter; and perturbation by the increment Δy yields the following dual programme for shadow-pricing the outputs:

$$\text{Given } (y,r,w)\,,\ \text{maximize } \langle p\,|\,y\rangle - \Pi_{\mathrm{LR}}(p,r,w) \text{ over } p. \tag{5.5}$$

Its optimal value is denoted by $\underline{C}_{\mathrm{LR}}(y,r,w) \leq C_{\mathrm{LR}}(y,r,w)$. The dual parameters are r and w.

In the SRP maximization programme (3.6)–(3.7), only k can serve as a primal parameter; and perturbation by the increment Δk yields the following dual programme for

[26] The standard dual to the ordinary CP of maximizing a concave function $f(y)$ over y subject to $G(y) \leq s$ (where G_1, G_2, etc., are convex functions) is to minimize $\sup_y \mathcal{L}(y,\sigma) := \sup_y (f(y) + \sigma \cdot (s - G(y)))$ over $\sigma \geq 0$ (the standard dual variables, which are the Lagrange multipliers for the primal constraints): see, e.g., [44, (5.1)]. And $\sup_y \mathcal{L}$ (the Lagrangian's supremum over the primal variables) cannot be evaluated without assuming a specific form for f and G (the primal objective and constraint functions).

[27] Since the minimand $\langle w\,|\,v\rangle$ is not jointly convex in (w,v), w cannot serve as a primal parameter.

shadow-pricing the fixed inputs:

$$\text{Given } (p,k,w)\,,\ \text{minimize } \langle r \mid k\rangle + \Pi_{\mathrm{LR}}(p,r,w) \ \text{ over } r. \tag{5.6}$$

Its optimal value is denoted by $\overline{\Pi}_{\mathrm{SR}}(p,k,w) \geq \Pi_{\mathrm{SR}}(p,k,w)$, with equality when Section 21 or 23 applies.[28] The dual parameters are p and w.

The same programme for r—viz., (5.6) or (5.13)–(5.14) under c.r.t.s.—is also the dual of the *reduced* SRP programme in (4.2), again with k as the primal parameter. That is, the reduced and the full primal programmes have the same primal parameters and the same dual programme. Of course, the duality relationships cannot be exactly the same because the dual parameterizations are different: as has already been pointed out, the reduced primal programme has fewer variables, and hence fewer dual parameters, than the full programme (all of whose data are its primal and dual parameters). Since both programmes have the same data, the reduced one has therefore a datum that is neither a primal nor a dual parameter. In the case of the reduced SRP programme in (4.2), w is such a datum: the only primal parameter is k, and the only dual parameter is p (since y is the only primal variable). For comparison, in the full SRP programme (3.6)–(3.7) both p and w are dual parameters (paired to the primal variables y and v).[29]

The LRP maximization programme (3.1)–(3.2) is, in this context, unusual because all its data (p, r and w) are dual parameters: no datum can serve as a primal parameter. This means that the dual has no decision variable; formally, it is: given (p,r,w), minimize $\Pi_{\mathrm{LR}}(p,r,w)$. Having no variable, the dual minimand is a constant, and it equals the primal value (Π_{LR}): since the dual is trivial, there can be no question of a duality gap in this case.

By contrast, the other programme pairs *can* have duality gaps, especially when the spaces are infinite-dimensional. But even then a gap can appear only at an exceptional data point: the primal and dual values are always equal under the generalized Slater's Condition of [44, (8.12)] or the compactness-and-continuity conditions of [44, Example 4' after (5.13)] and [44, Theorem 18' (d) or (e)]. In the problem of profit-maximizing operation of a plant with capacity constraints, Slater's Condition requires only that the capacities be strictly positive, i.e., that $k \gg 0$ (Sections 24 and 27). In other words, it is always met unless the plant k lacks a component. See Lemma 23.1, Proposition 27.2 and Appendix A for further explanation.

Partial conjugacy relationships between the dual value functions ($\underline{C}_{\mathrm{SR}}$, $\underline{C}_{\mathrm{LR}}$, $\overline{\Pi}_{\mathrm{SR}}$, and $\overline{\Pi}_{\mathrm{LR}} = \Pi_{\mathrm{LR}}$) can be summarized in a diagram like (3.12) but with the arrows reversed

[28] As our notation indicates, we think of $\overline{\Pi}$ and $\underline{C}$ mainly as dual expressions for Π and C (although duality of programmes is fully symmetric).

[29] A similar remark applies to the full and the reduced shadow-pricing programmes, (5.4) for (p,r) and that in (6.10) for p alone. Both are parameterized by w and have the same dual, viz., the SRC programme (3.10)–(3.11). All the vector data (y, k, w) are primal or dual parameters of the full programme (5.4) programme. But the datum k is not a primal or dual parameter of the reduced programme in (6.10).

and with bars added to the symbols Π and C):

$$\boxed{\begin{array}{ccccccc}
 & & & w & & & \\
 & & & \Pi_{\mathrm{LR}} & & & \\
 & & r & & p & & \\
 & \swarrow & & & & \searrow & \\
 & k & & & & y & \\
w\ \ \overline{\Pi}_{\mathrm{SR}} & & & & & & \underline{C}_{\mathrm{LR}}\ \ w \\
 & p & & & & r & \\
 & & \searrow & & \swarrow & & \\
 & & y & & k & & \\
 & & & \underline{C}_{\mathrm{SR}} & & & \\
 & & & w & & &
\end{array}}\;. \tag{5.7}$$

For example, the arrow from the p next to $\overline{\Pi}_{\mathrm{SR}}$ to the y next to $\underline{C}_{\mathrm{SR}}$ indicates that $\underline{C}_{\mathrm{SR}}$ is, as a function of y, the convex conjugate of $\overline{\Pi}_{\mathrm{SR}}$ as a function of p (with k and w fixed): i.e., by definition,

$$\underline{C}_{\mathrm{SR}}\,(y,k,w) = \sup_{p} \left\{ \langle p \,|\, y \rangle - \overline{\Pi}_{\mathrm{SR}}\,(p,k,w) \right\}. \tag{5.8}$$

In any specific case, formation of the primal-dual programme pair requires formulae for both $\mathbb{Y}$ and Π_{LR}. When the technology is given by a production set ($\mathbb{Y}$), this requires working out its support function (Π_{LR}). The task simplifies under c.r.t.s.: Π_{LR} is then $\delta\,(\cdot \mid \mathbb{Y}^{\circ})$, the 0-$\infty$ indicator of the production cone's polar (3.4). In other words, $\mathbb{Y}^{\circ}$ is the implicit dual constraint set and, by making the constraint explicit, the dual programmes can be cast in the same form as the primals. For each primal, the general form of the dual is specialized to the case of c.r.t.s. in the same way, viz., by adjoining the constraint $(p,r,w) \in \mathbb{Y}^{\circ}$ and deleting the now-vanishing term Π_{LR} from (5.4), etc. So the dual programme is to impute optimal values to the given quantities by pricing them in a way consistent with the other, given prices, i.e., so that the entire price system lies in $\mathbb{Y}^{\circ}$.

Spelt out, under c.r.t.s., the dual to SRC minimization is the following programme of maximizing the total output value less fixed-input value (OFIV) by shadow-pricing both the outputs and the fixed inputs:

$$\text{Given } (y,k,w)\,, \text{ maximize } \langle p \,|\, y \rangle - \langle r \,|\, k \rangle \text{ over } (p,r) \tag{5.9}$$

$$\text{subject to } (p,r,w) \in \mathbb{Y}^{\circ}. \tag{5.10}$$

The dual to LRC minimization is (with c.r.t.s.) the following programme of maximizing the total output value (OV) by shadow-pricing the outputs:

$$\text{Given } (y,r,w)\,, \text{ maximize } \langle p \,|\, y \rangle \text{ over } p \tag{5.11}$$

$$\text{subject to } (p,r,w) \in \mathbb{Y}^{\circ}. \tag{5.12}$$

The dual to SRP maximization is (under c.r.t.s.) the following programme of minimizing the total fixed-input value (FIV) by shadow-pricing the fixed inputs:

$$\text{Given } (p, k, w), \text{ minimize } \langle r \mid k \rangle \text{ over } r \tag{5.13}$$

$$\text{subject to } (p, r, w) \in \mathbb{Y}^{\circ}. \tag{5.14}$$

The dual to LRP maximization has no decision variable, and, with c.r.t.s., it may be thought of as a *price consistency check*: its value is 0 if $(p, r, w) \in \mathbb{Y}^{\circ}$, and $+\infty$ otherwise. Formally, the dual is:

$$\text{Given } (p, r, w), \text{ minimize } 0 \text{ subject to } (p, r, w) \in \mathbb{Y}^{\circ}. \tag{5.15}$$

Thus, with c.r.t.s., the dual objectives are "automatic", and formation of the dual programmes boils down to working out $\mathbb{Y}^{\circ}$ from a specific cone $\mathbb{Y}$. Two frameworks for this are provided in Sections 14 and 25.

Like the primals, the dual programmes are henceforth named after their objectives, OFIV, OV and FIV. Strictly speaking, this terminology fits only the case of c.r.t.s. for the long run (i.e., the case of a production cone). But it is used also when c.r.t.s. are not assumed (e.g., in Figure 2, Section 6 and Tables 1 and 2).

Comments:

(1) The dual can be similarly spelt out for a programme of a more general form, with a parametric primal maximand

$$\langle p \mid y \rangle - I(y, -k) \tag{5.16}$$

where I: $Y \times K \rightarrow \mathbb{R} \cup \{+\infty\}$ is a bivariate convex function, y is the primal variable, p and k are the data, of which k is the primal parameter. There is no explicit constraint, but there is the implicit constraint $(y, -k) \in \operatorname{dom} I$. The dual minimand is then

$$\langle r \mid k \rangle + I^{\#}(p, r) \tag{5.17}$$

where $I^{\#}$: $Y \times K \rightarrow \mathbb{R} \cup \{+\infty\}$ is the total (bivariate) convex conjugate of I, r is the dual variable, and p is the dual parameter. (So the dual and the primal parameters are the coefficients of the primal and the dual decision variables.)

(2) If I is the 0-∞ indicator of a convex set $\mathbb{Y}$, then $I^{\#}$ is the support function of $\mathbb{Y}$. If additionally $\mathbb{Y}$ is a cone then $I^{\#}$ is the indicator of $\mathbb{Y}^{\circ}$, and the programme of minimizing $\langle r \mid k \rangle$ over r subject to $(p, r) \in \mathbb{Y}^{\circ}$ is dual to the primal programme of maximizing $\langle p \mid y \rangle$ over y subject to $(y, -k) \in \mathbb{Y}$ (parameterized by k). This is spelt out in the Proof of Proposition 18.1 (where (p, w) and $(y, -v)$ take place of the above p and y).

(3) The case of a finite LP, parameterized in the standard way, is obtained when

$$\mathbb{Y} = \{(y, -k) \in \mathbb{R}^n \times \mathbb{R}^m : Ay \leq k\} \text{ with } \mathbb{Y}^{\circ} = \left\{(p, r) \in \mathbb{R}^n \times \mathbb{R}^m : p = A^{\mathrm{T}} r,\ r \geq 0\right\}$$

where A is an $m \times n$ matrix. With general spaces, A: $Y \rightarrow K$ is a linear operation with the adjoint A^*: $R \rightarrow P$, defined by $\langle A^* r \mid y \rangle := \langle r \mid Ay \rangle$, which replaces A^{T}. In other words, minimization of $\langle r \mid k \rangle$ over r subject to $p = A^* r$ and $r \geq 0$ is dual to maximization of $\langle p \mid y \rangle$ over y subject to $Ay \leq k$, with k as the primal parameter vector.

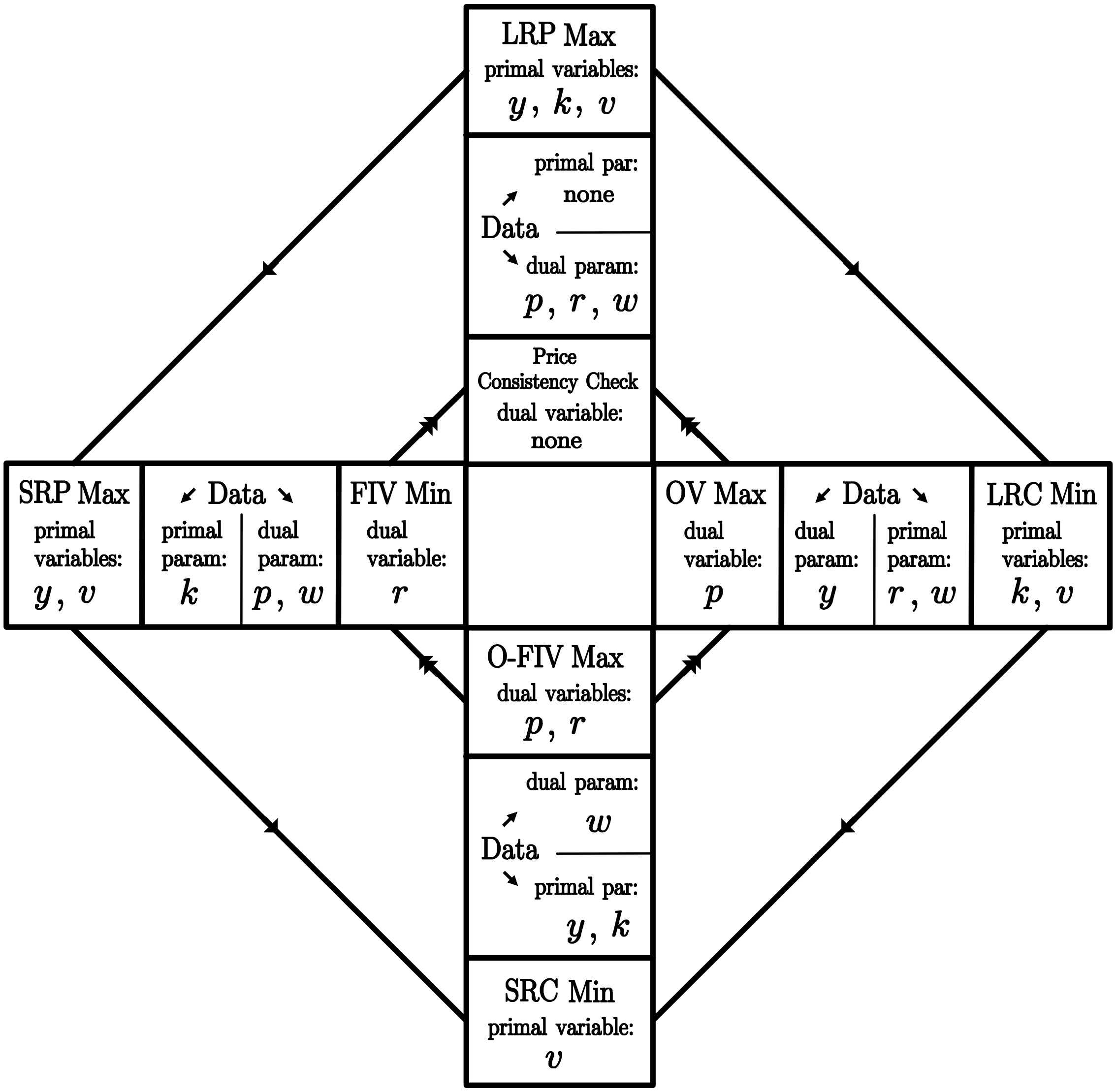

FIGURE 2. Decision variables and parameters for primal programmes (optimization of long-run profit, short-run profit, long-run cost, short-run cost) and for dual programmes (price consistency check, optimization of: fixed-input value, output value, output value less fixed-input value). In each programme pair, the same prices and quantities—(p, y) for outputs, (r, k) for fixed inputs, and (w, v) for variable inputs—are differently partitioned into decision variables and data (which are subdivided into primal and dual parameters). Arrows lead from programmes to subprogrammes.

6. Other primal-dual optimization systems

Our use of conjugate functions gives a characterization of the profit-maximizing investment in terms of its imputed values, i.e., it reformulates the investment-optimality condition (4.1) as the valuation condition (4.4). The valuation programme in (4.4) is subsequently identified as the dual (5.6), or (5.13)–(5.14) under c.r.t.s., to the short-run profit maximization programme (3.6)–(3.7), which appears in (4.2)–(4.3) in a split form. Thus the use of conjugacy produces the system (4.2)–(4.5) of optimality conditions on y, v and r; and the use of duality shows that this system means that (y, v) and r form a pair of solutions to the SRP programme and its dual.[30] Similar arguments lead to characterizations of optimality in terms of the LRC or SRC programme with its dual, i.e., each of the following three systems of conditions is equivalent to maximization of long-run profit at prices (p, r, w) by an input-output bundle $(y, -k, -v)$:

(1) $(y, -v)$ maximizes the short-run profit at prices (p, w), and r minimizes the value of the fixed-input k (plus maximum LRP if r.t.s. are decreasing), and the two optimal values are equal (i.e., under c.r.t.s., maximum SRP equals minimum FIV). Formally:

(6.1) (y, v) solves the primal SRP programme (3.6)–(3.7), given (p, k, w).

(6.2) r solves the dual (5.6), which is (5.13)–(5.14) under c.r.t.s., given (p, k, w).

(6.3) $\overline{\Pi}_{\mathrm{SR}} (p, k, w) = \Pi_{\mathrm{SR}} (p, k, w)$.

(2) (k, v) minimizes the long-run cost at prices (r, w), and p maximizes the value of output y (less maximum LRP under d.r.t.s.), and the two optimal values are equal (i.e., under c.r.t.s., minimum LRC equals maximum OV). Formally:

(6.4) (k, v) solves the primal LRC programme (3.8)–(3.9), given (y, r, w).

(6.5) p solves the dual (5.5), which is (5.11)–(5.12) under c.r.t.s., given (y, r, w).

(6.6) $\underline{C}_{\mathrm{LR}} (y, r, w) = C_{\mathrm{LR}} (y, r, w)$.

(3) v minimizes the short-run cost at price w, and (p, r) maximizes the value of output y less that of fixed-input k (and less maximum LRP under d.r.t.s.), and the two optimal values are equal (i.e., under c.r.t.s., minimum SRC equals maximum OFIV). Formally:

(6.7) (p, r) solves the dual (5.4), a.k.a. (5.9)–(5.10) under c.r.t.s., given (y, k, w).

(6.8) v solves the primal SRC programme (3.10)–(3.11), given (y, k, w).

(6.9) $\underline{C}_{\mathrm{SR}} (y, k, w) = C_{\mathrm{SR}} (y, k, w)$.

[30] These arguments exploit the subprogramme concept as well as that of duality, i.e., Π_{SR} is viewed in two ways: (i) as the value of a subprogramme, and (ii) as the primal value. Both contexts give rise to the conjugacy between Π_{SR} and Π_{LR}—and that is why there are two ways of deriving the valuation programme in (4.4). In detail, since Π_{SR} is the value of the subprogramme of LRP maximization obtained by fixing k, its (concave) conjugate w.r.t. k is $-\Pi_{\mathrm{LR}}$ as a function of r: this is (3.14). It follows, by (C.24) and (C.32), that k solves the "conjugacy programme" in (4.1) if and only if r solves the "reverse" one in (4.4) and (4.5) holds. The same programme for r can be derived independently as the dual to SRP maximization parameterized by k, as is done in Proposition 18.1 (which also identifies p and w as the dual parameters). Alternatively, it can be identified as the dual by using the conjugacy between Π_{SR} and Π_{LR}: it is a foundation of duality for CPs that the (concave) conjugate of the primal maximum value (as a function of the primal parameter) plus the primal parameter times the dual variable (here, $\Pi_{\mathrm{LR}} (r) + \langle r \mid k \rangle$) is the dual minimand. See, e.g., [44, Theorem 7], which here must be applied to the function $\Delta k \mapsto \Pi_{\mathrm{SR}} (k + \Delta k)$ as Rockafellar's primal value (his is a function of the parameter increment, rather than of the parameter point like ours, and this shifts the argument by k and adds the term $\langle r \mid k \rangle$ to the conjugate).

Additionally, one can split each of the joint programmes for two decision variables: just as (3.6)–(3.7) has been split into (4.2) and (4.3), so the joint programme (5.4) for (p, r) can be replaced by two programmes for p and r separately. Condition (6.7) is therefore equivalent to:[31]

(6.10) $\quad p$ maximizes $\langle \cdot \mid y \rangle - \overline{\Pi}_{\mathrm{SR}}(\cdot, k, w)$ on P (given y, k and w)

(6.11) $\quad r$ solves (5.6), given (p, k, w).

Thus the joint shadow-pricing programme (5.4) for (p, r) is reduced to an output-pricing programme, for p alone, by first solving the fixed-input shadow-pricing programme (5.6) for r and substituting its optimal value ($\overline{\Pi}_{\mathrm{SR}}$) for the term $\langle r \mid k \rangle + \Pi_{\mathrm{LR}}(p, r, w)$ in (5.4). In other words, two-stage solving means in this case:

(1) first minimizing $\langle r \mid k \rangle$ over r to find the solution set $\hat{R}(p, k, w)$, or *the* solution $\hat{r}(p, k, w)$ if it is indeed unique, and the minimum value $\overline{\Pi}_{\mathrm{SR}}(p, k, w)$, which is $\langle \hat{r} \mid k \rangle$;

(2) then maximizing $\langle p \mid y \rangle - \overline{\Pi}_{\mathrm{SR}}(p, k, w)$ over p to find the solution set $\check{P}(y, k, w)$, or *the* solution $\check{p}(y, k, w)$, should it be unique. This gives every complete solution (in terms of y, k and w) as a (p, r) such that $p \in \check{P}(y, k, w)$ and $r \in \hat{R}(p, k, w)$. Should the solution be unique, it is the pair $\check{p}(y, k, w)$ and $\hat{r}(\check{p}(y, k, w), k, w)$.

Another proof of the equivalence, to LRP maximization, of the three systems (6.1)–(6.3), (6.4)–(6.6) and (6.7)–(6.9) follows from a general inequality between the values of a programme pair (taking for granted that (5.4) to (5.6) are indeed the relevant duals, as is stated and proved in Sections 5 and 18). What is to be shown is that each system is equivalent to (3.3), or to the Complementarity Conditions (3.5) in the case of c.r.t.s. For each of the three programme pairs, (3.3) or (3.5) means: (i) primal feasibility, of either (y, v) or (k, v) or v, (ii) dual feasibility, of either r or p or (p, r), and (iii) equality of the primal maximand to the dual minimand, at the two points in question. So it suffices to note that these *FFE Conditions* (which have already appeared as (5.3) in the LP context) fully characterize a pair of solutions with equal values because the primal maximand never exceeds the dual minimand (at feasible points).

Thus the data (p, r, w) and the solution $(y, -k, -v)$ of the LRP programme (3.1)–(3.2) can be permuted to form the data and solutions to any of the three subprogrammes with its dual (when there is no duality gap). In each case, a pair of solutions gives three of the six variables—one from each of the three price-quantity pairs (viz., (p, y) for outputs, (r, k) for fixed inputs, and (w, v) for variable inputs)—in terms of the other three (which are parameters, not decision variables).

The three systems (6.1)–(6.3), (6.4)–(6.6) and (6.7)–(6.9) can be called the *SRP*, *LRC*, and *SRC optimization systems*, since each is put entirely in terms of solutions to the named programme and its dual. Also, each system contains a joint programme, which can be split to produce the corresponding *split optimization system*, viz.: (4.2)–(4.5), or (11.11)–(11.14) spelt out in Section 11, or (6.8)–(6.11). We have chosen to introduce the first of these, the split SRP system (4.2)–(4.5), before the programme for r in (4.4) could be formally identified as the dual of the SRP programme (in Section 5). In (6.2), the

[31] The maximum value in (6.10) is $\underline{C}_{\mathrm{SR}}(y, k, w)$, by the definitions of $\overline{\Pi}_{\mathrm{SR}}$ and $\underline{C}_{\mathrm{SR}}$ as the optimal values of (5.6) and (5.4).

same programme is referred to as the dual. So the split SRP optimization system can be restated as the conjunction of (4.2)–(4.3) and (6.2)–(6.3).

Comment (alternative ways of applying LP algorithms): Under c.r.t.s., once *both* the production cone $\mathbb{Y}$ *and* its polar polar $\mathbb{Y}^{\circ}$ have been represented as intersections of half-spaces,[32] each profit or cost programme and its dual become LPs, which are finite if $\mathbb{Y}$ is a polyhedral cone in a finite-dimensional commodity space. Then the FFE Conditions, which are the Complementarity Conditions (3.5), become a finite system of linear inequalities and equalities in finitely many variables. Like any such system, it can be solved either directly by Fourier-Motzkin elimination (which gives all the solutions) or indirectly by converting it into an auxiliary LP and applying the simplex method (or another algorithm) to find at least one solution, and thus also the value of the original profit or cost programme and its dual (any other solutions can then also be found). However, it seems somewhat better to deal with the original LP than to solve the Complementarity Conditions by either method. First, the Fourier-Motzkin elimination is far less efficient than the simplex method (applied to the auxiliary LP); this is noted in, e.g., [11, p. 242]. Second, the original LP is smaller in size than the auxiliary LP.[33]

7. A saddle differential system for the short-run approach

In convex programming, optimality is fully expressed by the first-order condition. Furthermore, by combining the FOC with the Inversion Rule for the derivative of a conjugate function, the optimal solution can be interpreted as a marginal value. This *derivative property* of the optimal-value function extends to the case of nonunique solutions. The value is then nondifferentiable in the ordinary way, but it has a generalized, multi-valued derivative. For a convex function, this is the *subdifferential* (a.k.a. the subgradient set),

[32]This requires switching from parametric equations to inequalities in coordinates: whenever the one cone, say $\mathbb{Y}$, is given by a system of homogeneous linear inequalities in coordinates, its polar ($\mathbb{Y}^{\circ}$) is readily given by a system of parametric equations, but these must be converted to coordinate inequalities. Geometrically, $\mathbb{Y}$ is the intersection of a finite number of half-spaces if and only if their outer normal vectors generate $\mathbb{Y}^{\circ}$ (i.e., give it as the set of their nonnegative linear combinations); this is Farkas's Lemma. In symbols, $\mathbb{Y}$ consists of all those $(y, -k)$ with $Ay - Bk \leq 0$ if and only if $\mathbb{Y}^{\circ}$ consists of all those (p, r) with $p^{\mathrm{T}} = \sigma^{\mathrm{T}} A$ and $r^{\mathrm{T}} = \sigma^{\mathrm{T}} B$ for some $\sigma \geq 0$. But what is needed is an equivalent system of the form $\left[p^{\mathrm{T}}, r^{\mathrm{T}}\right] M \leq 0$. Such conversions (from a parametric to a coordinate description of a cone or more generally a polyhedron, and *vice versa*) can be done in the way described in, e.g., [11, Chapters 16 and 18]. In the case considered later in this Comment, the conversion is immediate because $B = \mathrm{I}$, and so the representing parameter σ can be replaced by r (so the polar is given by the coordinate inequalities and equalities $p^{\mathrm{T}} = r^{\mathrm{T}} A$ and $r \geq 0$).

[33]To see this (i.e., that the original is smaller than the auxiliary LP), let the original primal LP be to maximize $p \cdot y$ over $y \in \mathbb{R}^n$ subject to $Ay \leq k$, given arbitrary vectors $p \in \mathbb{R}^n$ and $k \in \mathbb{R}^m$, and an $m \times n$ matrix A. The dual LP is to minimize $r \cdot k$ over $r \geq 0$ subject to $r^{\mathrm{T}} A = p^{\mathrm{T}}$. So the FFE (Complementarity) Conditions on (y, r) are: $Ay \leq k$, $r \geq 0$, $r^{\mathrm{T}} A = p^{\mathrm{T}}$ and $p \cdot y \geq r \cdot k$ (or, equivalently, $p \cdot y = r \cdot k$). This is a system with $n + m$ variables and $2m + 2n + 1$ inequalities (counting an equality as two inequalities). Its auxiliary LP has $n + m + 1$ decision variables, (viz., y, r and an artificial variable, say $z \geq 0$, as the minimand, whose minimum value is zero if and only if the FFE system is soluble) and $2(m + n + 1)$ inequality constraints (viz., $z \geq 0$ and all the complementarity inequalities but with z subtracted from the lesser side, i.e., $p \cdot y \geq r \cdot k - z$, etc.): see [11, (16.2), p. 240]. So the auxiliary LP has one more variable and one more constraint than the original primal and dual LPs *together*. Solving the auxiliary LP by a primal-dual algorithm gives a solution to the original LP "in duplicate".

defined by (C.11) and denoted by ∂. The *superdifferential* of a concave function, denoted here by $\widehat{\partial}$, is defined by (C.23). Each of the functions Π_{SR}, C_{SR} and C_{LR} is either convex or concave jointly in two of its three variables, and it is concave or convex in the other variable. For example, $\Pi_{\mathrm{SR}}(p,k,w)$ is jointly convex in (p,w), and concave in k (as is $\overline{\Pi}_{\mathrm{SR}}$).

The split LRP optimization system (4.1)–(4.3) is thus transformed into the partial subdifferential system that consists of the FOCs for (4.1) and (4.2) and of the derivative property of C_{SR} as the optimal value of (4.3), i.e., into the system

$$r \in \widehat{\partial}_k \Pi_{\mathrm{SR}}(p,k,w) \tag{7.1}$$

$$p \in \partial_y C_{\mathrm{SR}}(y,k,w) \tag{7.2}$$

$$v \in \widehat{\partial}_w C_{\mathrm{SR}}(y,k,w)\,. \tag{7.3}$$

We call it the *SRC-P saddle differential system*, since it uses $\partial_y C_{\mathrm{SR}}$ and $\widehat{\partial}_w C_{\mathrm{SR}}$, the partial sub/super-differentials of C_{SR} as a saddle (convex-concave) function of (y,w), in addition to using $\widehat{\partial}_k \Pi_{\mathrm{SR}}$. A similar use of C_{SR}, as a saddle function of (k,w), arises later in the L-SRC system (11.8)–(11.10): the affices "L" and "P" in these names stand for "long-run" and "profit".

Comments (use of a differential condition to absorb a no-gap condition):

1. The system (7.1)–(7.3) can be derived also from the split SRP optimization system (4.2)–(4.5). The FOC for (4.2) and the derivative property of C_{SR} as the value function for (4.3) are used just as before. But, instead of the FOC for (4.1), this time the third condition is the derivative property of $\overline{\Pi}_{\mathrm{SR}}$ as the value function for (4.4) or (5.6), i.e., that $r \in \widehat{\partial}_k \overline{\Pi}_{\mathrm{SR}}(p,k,w)$. Taken together, this and (4.5) mean exactly that $r \in \widehat{\partial}_k \Pi_{\mathrm{SR}}$, since (4.5) means that $\overline{\Pi}_{\mathrm{SR}} = \Pi_{\mathrm{SR}}$, at (p,k,w).
2. The last argument is a case of absorbing a no-gap condition in a subdifferential condition by changing the derivative from Type Two (here, $\widehat{\partial}_k \overline{\Pi}_{\mathrm{SR}}$) to Type One ($\widehat{\partial}_k \Pi_{\mathrm{SR}}$). This is done by changing the value function either from dual to primal (if the parameter in question is primal like the k here), or *vice versa*. The optimal solution is *always* equal to the marginal value of the programme being solved; this is a derivative of Type Two. It is actually of Type One—i.e., it is the marginal value of the programme *dual* to that being solved—if there is no duality gap. But if there is a gap, the Type One derivative does not exist. In the above case of fixed-input valuation, the set of solutions, for r, of (4.4) or (5.6) is always identical to $\widehat{\partial}_k \overline{\Pi}_{\mathrm{SR}}$ (which is a derivative of Type Two). It equals $\widehat{\partial}_k \Pi$ (a derivative of Type One) if $\Pi_{\mathrm{SR}} = \overline{\Pi}_{\mathrm{SR}}$ at the given (p,k,w). But if $\Pi_{\mathrm{SR}} \neq \overline{\Pi}_{\mathrm{SR}}$ then $\widehat{\partial}_k \Pi = \emptyset$ (the empty set); so if $r \in \widehat{\partial}_k \Pi_{\mathrm{SR}}$ then $\Pi_{\mathrm{SR}} = \overline{\Pi}_{\mathrm{SR}}$ (at the given p, k and w).

8. Other subdifferential systems

Applied to the split SRC optimization system (6.8)–(6.11), the same methods—viz., the FOC for (6.10) and the derivative properties of C_{SR} and $\overline{\Pi}_{\mathrm{SR}}$ as the value functions for (6.8) and (6.11), followed by changing $\partial_w C_{\mathrm{SR}}$ to $\partial_w \underline{C}_{\mathrm{SR}}$ to absorb the no-gap condition

(6.9)—yield the partial subdifferential system

$$r \in \widehat{\partial}_k \overline{\Pi}_{\mathrm{SR}}(p,k,w) \tag{8.1}$$

$$y \in \partial_p \overline{\Pi}_{\mathrm{SR}}(p,k,w) \tag{8.2}$$

$$v \in \widehat{\partial}_w \underline{C}_{\mathrm{SR}}(y,k,w)\,. \tag{8.3}$$

We call it the *FIV saddle differential system*, since it uses $\partial_p \overline{\Pi}_{\mathrm{SR}}$ and $\widehat{\partial}_k \overline{\Pi}_{\mathrm{SR}}$, the partial sub/super-differentials of $\overline{\Pi}_{\mathrm{SR}}$ as a saddle function of (p,k), in addition to using $\widehat{\partial}_w \underline{C}_{\mathrm{SR}}$. Thus it uses only the dual value functions ($\overline{\Pi}_{\mathrm{SR}}$ and $\underline{C}_{\mathrm{SR}}$), whilst the system (7.1)–(7.3) uses only the primal value functions (Π_{SR} and C_{SR}).

The derivative property of the optimal value can also be used to transform the "unsplit" optimization systems of Section 6. For example, by the derivative property applied twice, the SRP optimization system (6.1)–(6.3) is equivalent to:

$$(y,-v) \in \partial_{p,w} \Pi_{\mathrm{SR}}(p,k,w)\,,\ r \in \widehat{\partial}_k \overline{\Pi}_{\mathrm{SR}}(p,k,w) \text{ and } \overline{\Pi}_{\mathrm{SR}}(p,k,w) = \Pi_{\mathrm{SR}}(p,k,w)\,.$$

The no-gap condition can be absorbed in either subdifferential condition by changing $\overline{\Pi}_{\mathrm{SR}}$ to Π_{SR} or *vice versa*. This produces the *SRP subdifferential system*

$$(y,-v) \in \partial_{p,w} \Pi_{\mathrm{SR}}(p,k,w) \tag{8.4}$$

$$r \in \widehat{\partial}_k \Pi_{\mathrm{SR}}(p,k,w) \tag{8.5}$$

and the *FIV subdifferential system*

$$(y,-v) \in \partial_{p,w} \overline{\Pi}_{\mathrm{SR}}(p,k,w) \tag{8.6}$$

$$r \in \widehat{\partial}_k \overline{\Pi}_{\mathrm{SR}}(p,k,w)\,. \tag{8.7}$$

Similarly, the LRC optimization system (6.4)–(6.6) is equivalent to

$$(k,v) \in \widehat{\partial}_{r,w} C_{\mathrm{LR}}(y,r,w)\,,\ p \in \partial_y \underline{C}_{\mathrm{LR}}(y,r,w) \text{ and } \underline{C}_{\mathrm{LR}}(y,r,w) = C_{\mathrm{LR}}(y,r,w)$$

and, hence, also to the *LRC subdifferential system*

$$(k,v) \in \widehat{\partial}_{r,w} C_{\mathrm{LR}}(y,r,w) \tag{8.8}$$

$$p \in \partial_y C_{\mathrm{LR}}(y,r,w) \tag{8.9}$$

as well as to the *OV subdifferential system*

$$(k,v) \in \widehat{\partial}_{r,w} \underline{C}_{\mathrm{LR}}(y,r,w) \tag{8.10}$$

$$p \in \partial_y \underline{C}_{\mathrm{LR}}(y,r,w) \tag{8.11}$$

Finally, the SRC optimization system (6.7)–(6.9) is equivalent to

$$v \in \widehat{\partial}_w C_{\mathrm{SR}}(y,k,w)\,,\ (p,-r) \in \partial_{y,k} \underline{C}_{\mathrm{SR}}(y,k,w) \text{ and } \underline{C}_{\mathrm{SR}}(y,k,w) = C_{\mathrm{SR}}(y,k,w)$$

and, hence, also to the *SRC subdifferential system*

$$v \in \widehat{\partial}_w C_{\mathrm{SR}}(y,k,w) \tag{8.12}$$

$$(p,-r) \in \partial_{y,k} C_{\mathrm{SR}}(y,k,w) \tag{8.13}$$

as well as to the *OFIV subdifferential system*

$$v \in \widehat{\partial}_w \underline{C}_{\mathrm{SR}}(y,k,w) \tag{8.14}$$

$$(p, -r) \in \partial_{y,k} \underline{C}_{\mathrm{SR}} (y, k, w) . \tag{8.15}$$

Comments (on the terminology):

(1) Like the names of valuation programmes, the name "FIV/OV/OFIV system" is used only for brevity, i.e., without actually assuming c.r.t.s.

(2) The derivative properties of profit and cost as functions of prices—i.e., characterizations of optimality such as (8.4), (8.10) and (8.12)—are known as the Shephard-Hotelling Lemmas; their proofs are detailed in Section 19. Similarly, long-run profit maximization is equivalent to: $(y, -k, -v) \in \partial_{p,r,w} \Pi_{\mathrm{LR}} (p, r, w)$.

9. Transformations of differential systems by the Subdifferential Sections Lemma

So far, all the differential systems have been derived from optimization systems—and this has to be so in convex analysis because it uses the FOC for maximization as the very definition of the subdifferential: see (C.12). But this definition can be used to transform one subdifferential condition into another. Once formulated, such results can be applied to transform the differential systems into one another "directly", i.e., without going explicitly through the FOCs. In particular, partial subdifferential systems can be derived from systems with joint subdifferentials: a condition involving a subdifferential taken jointly in two groups of variables—such as $\partial_{y,k} C_{\mathrm{SR}}$ in (8.13) or $\partial_{p,w} \overline{\Pi}_{\mathrm{SR}}$ in (8.6)—can be recast in terms of partial subdifferentials (∂_y, ∂_k, ∂_p, ∂_w). This cannot, however, be achieved simply by splitting the joint derivative into the partials (as in the differentiable case) because a joint subdifferential does not usually factorize into the Cartesian product of the partials: it is a general convex set, not a product set. In other words, the obvious inclusions[34]

$$\partial_{y,k} C_{\mathrm{SR}} (y, k) \subseteq \partial_y C_{\mathrm{SR}} (y, k) \times \partial_k C_{\mathrm{SR}} (y, k) \tag{9.1}$$

$$\partial_{p,w} \overline{\Pi}_{\mathrm{SR}} (p, w) \subseteq \partial_p \overline{\Pi}_{\mathrm{SR}} (p, w) \times \partial_w \overline{\Pi}_{\mathrm{SR}} (p, w) \tag{9.2}$$

are usually strict: see Appendix B for further explanation and examples. But the two variables of differentiation can be split from each other in a way that parallels, and derives from, the staged approach to optimization (introduced in Section 4). First, the joint subdifferential is used to formulate a FOC for simultaneous optimization over the two variables; this programme is then split into two successive optimization programmes for one variable each, and each of these has a separate FOC that uses a partial subdifferential. In the case of $\partial_{y,k} C_{\mathrm{SR}}$, this argument consists in stating the FOCs for maximizing the LRP over y and k simultaneously as well as for maximizing it over y and k successively. The FOC for a maximum of $\langle p \mid y \rangle - \langle r \mid k \rangle - C_{\mathrm{SR}}$ over (y, k) is that $(p, -r) \in \partial_{y,k} C_{\mathrm{SR}}$. The FOC for a maximum of $\langle p \mid y \rangle - C_{\mathrm{SR}} (y, k, w)$ over y is that $p \in \partial_y C_{\mathrm{SR}}$; the maximum value is Π_{SR}, and the FOC for a maximum of $\Pi_{\mathrm{SR}} (p, k, w) - \langle r \mid k \rangle$ over k is that $r \in \widehat{\partial}_k \Pi_{\mathrm{SR}}$. Since the "joint" FOC is equivalent to the two "partial" FOCs together,[35]

$$(p, -r) \in \partial_{y,k} C_{\mathrm{SR}} (y, k, w) \Leftrightarrow \Big(p \in \partial_y C_{\mathrm{SR}} (y, k, w) \text{ and } r \in \widehat{\partial}_k \Pi_{\mathrm{SR}} (p, k, w) \Big) . \tag{9.3}$$

[34] Being fixed, the third variable (k or w) is suppressed here.

[35] Dually, (8.6) is equivalent to (8.2)–(8.3), i.e., $(y, -v) \in \partial_{p,w} \overline{\Pi}_{\mathrm{SR}}$ if and only if both $y \in \partial_p \overline{\Pi}_{\mathrm{SR}}$ and $v \in \widehat{\partial}_w \underline{C}_{\mathrm{SR}}$.

This is the Subdifferential Sections Lemma (SSL) for this context; it requires bringing in another function (Π_{SR}), which is linked to the original function (C_{SR}) by partial conjugacy. This result is fully formalized in Appendix C (Lemma C.5).

The SSL is the basic tool for "splitting" joint subdifferentials, but there is also a couple of derived techniques, viz., the Partial Inversion Rule and its dual variant (PIR and DPIR, i.e., Corollaries C.6 and C.8). Each of these can be applied to the joint subdifferentials of Section 8:

(1) With k fixed, the DPIR applies to $C_{\mathrm{SR}}(\cdot, k, \cdot)$ as a saddle function on $Y \times W$ which is a partial conjugate of the 0-∞ indicator of the short-run production set $\mathbb{Y}_{\mathrm{SR}}(k)$, defined formally by (21.1). The indicator is a convex function on $Y \times V$, and its total conjugate is $\Pi_{\mathrm{SR}}(\cdot, k, \cdot)$ on $P \times W$. It follows that the condition $(y, -v) \in \partial_{p,w} \Pi_{\mathrm{SR}}$ can be replaced by: $p \in \partial_y C_{\mathrm{SR}}$ and $v \in \widehat{\partial}_w C_{\mathrm{SR}}$. Thus the SRP subdifferential system (8.4)–(8.5) can be transformed into the SRC-P saddle differential system (7.1)–(7.3).

The PIR gives the same result, though it requires an additional, preliminary step, viz., using the SRP system's implication that C_{SR} is l.s.c. at y to invert the conjugacy relationship (3.13), i.e., to show that the saddle function $C_{\mathrm{SR}}(\cdot, k, \cdot)$ is indeed a partial conjugate of the convex function $\Pi_{\mathrm{SR}}(\cdot, k, \cdot)$.

The same saddle system, (7.1)–(7.3), can also be derived from the SRC subdifferential system (8.12)–(8.13). This is what (9.3) shows: with w fixed, the SSL applies to $\Pi_{\mathrm{SR}}(\cdot, \cdot, w)$ as a saddle function on $P \times K$ which is (by definition) a partial conjugate of $C_{\mathrm{SR}}(\cdot, \cdot, w)$, a convex function on $Y \times K$. So the condition $(p, -r) \in \partial_{y,k} C_{\mathrm{SR}}$ can be replaced by: $p \in \partial_y C_{\mathrm{SR}}$ and $r \in \widehat{\partial}_k \Pi_{\mathrm{SR}}$.

(2) Similarly, with w fixed, the DPIR applies to $\overline{\Pi}_{\mathrm{SR}}(\cdot, \cdot, w)$ as a saddle function on $P \times K$ which is a partial conjugate of $\Pi_{\mathrm{LR}}(\cdot, \cdot, w)$. When $\mathbb{Y}$ is a cone, the latter function is the indicator of $\mathbb{Y}^{\circ}_{w}$, the section of $\mathbb{Y}^{\circ}$ through w. In any case, it is a convex function on $P \times R$, and its total conjugate is $\underline{C}_{\mathrm{SR}}(\cdot, \cdot, w)$ on $Y \times K$. This shows that the condition $(p, -r) \in \partial_{y,k} \underline{C}_{\mathrm{SR}}(y, k, w)$ can be replaced by: $y \in \partial_p \overline{\Pi}_{\mathrm{SR}}$ and $r \in \widehat{\partial}_k \overline{\Pi}_{\mathrm{SR}}$. Thus the OFIV subdifferential system (8.14)–(8.15) can be transformed into the FIV saddle differential system (8.1)–(8.3).

The PIR gives the same result, though it requires an additional, preliminary step, viz., using the OFIV system's implication that $\overline{\Pi}_{\mathrm{SR}}$ is l.s.c. at p to invert the conjugacy relationship (5.8), i.e., to show that the saddle function $\overline{\Pi}_{\mathrm{SR}}(\cdot, \cdot, w)$ is indeed a partial conjugate of the convex function $\underline{C}_{\mathrm{SR}}(\cdot, \cdot, w)$.

The same saddle system, (8.1)–(8.3), can also be derived from the FIV subdifferential system (8.6)–(8.7). This is because, with k fixed, the SSL applies to $\underline{C}_{\mathrm{SR}}(\cdot, k, \cdot)$ as a saddle function on $Y \times W$ which is (by definition) a partial conjugate of $\overline{\Pi}_{\mathrm{SR}}(\cdot, k, \cdot)$, a convex function on $P \times W$. So the condition $(y, -v) \in \partial_{p,w} \overline{\Pi}_{\mathrm{SR}}$ can be replaced by: $y \in \partial_p \overline{\Pi}_{\mathrm{SR}}$ and $v \in \widehat{\partial}_w \underline{C}_{\mathrm{SR}}$.

10. Summary of systems characterizing a long-run producer optimum

Tables 1 and 2 summarize ten duality-based systems (and proofs of their equivalence, which are detailed in Section 19). Since the top right entry of the one table is identical to the bottom right of the other, the twelve entries include two repetitions. The ten distinct entries are all but three of the systems given so far—all except for the three that use the LRC programme and its dual or their value functions. Those three, and four more systems to appear in Section 11, are mirror images of the systems shown in the two tables, from which they can be obtained by replacing $\Pi_{\mathrm{SR}}(p,k)$ with $C_{\mathrm{LR}}(y,r)$ and changing signs where needed. Thus the three systems, viz.: (8.8)–(8.9), (6.4)–(6.6) and (8.10)–(8.11), correspond to those on the left in Table 1, and the four systems of Section 11 come from the distinct systems on the right in Tables 1 and 2.[36] In other words, Tables 1 and 2 deal explicitly with the values and programmes in the left halves of the conjugacy diagrams (3.12) and (5.7), but the analysis applies equally to the right halves.

In differential systems, the Type One derivatives that exclude duality gaps are identified. In optimization systems, the various duals are referred to as "optimization of the fixed quantities' value", although this name fits only the case of c.r.t.s. (which need not be assumed). The constraint sets ($\mathbb{Y}$ and $\mathbb{Y}^{\circ}$, under c.r.t.s.) are not shown.

Comments (partition into a short-run subsystem and a supplementary condition):

(1) All but three of the ten systems in Tables 1 and 2—all except for (6.7)–(6.9), (8.12)–(8.13) and (8.14)–(8.15), which appear on the left in Table 2—contain a condition on r and (p,k,w) that is either exactly or at least nearly equivalent to k being a profit-maximizing investment at prices (p,r,w), i.e., to (4.1). The condition in question is: $r \in \widehat{\partial}_k \Pi_{\mathrm{SR}}$, or $r \in \widehat{\partial}_k \overline{\Pi}_{\mathrm{SR}}$, or "$r$ minimizes FIV". Together, the system's other conditions—on p, y, w, v and k—are then essentially equivalent to (4.2)–(4.3), i.e., to $(y,-v)$ being a short-run profit-maximizing input-output bundle at prices (p,w), given capital inputs k. This *short-run subsystem* is to be solved for v and either y or p—given w and either p or y, as well as k. It may be so simple that, as in Section 2, it can be solved on its own, separately from the remaining *supplementary condition* (i.e., without recourse to duality). Apart from being handy in such simple cases, the system's partition (into a short-run subsystem and a supplementary condition that involves r and essentially means investment being at a profit maximum) is worth examining in detail to clarify the various ways in which the complete systems exclude duality gaps. Most do so within the subsystem, but some rely on the supplementary condition (when it is that $r \in \widehat{\partial}_k \overline{\Pi}_{\mathrm{SR}}$, which is a Type One derivative). Therefore, the various subsystems describe two "grades" of short-run profit maximum: the "plain" one and the one without a duality gap. Only the latter kind can be a long-run profit maximum (for some choice of capital-input prices).

(2) More formally, given (p,w) and k, a *potential long-run* profit maximizing bundle is a $(y,-v)$ such that $(y,-k,-v)$ maximizes long-run profit at (p,w) and some r. Obviously, every system can be turned into a characterization of potential long-run optimality by binding r with an existential quantifier. But in the three excepted systems ((6.7)–(6.9), (8.12)–(8.13) and (8.14)–(8.15)), the condition on r involves also y (as well as p, k, w), and it expresses the optimality not only of k

[36]The three systems on the left in Table 2 do not yield new ones (when Π_{SR} is replaced by C_{LR}) simply because they do not involve Π_{SR} at all. So there are not ten but seven (3 already given and 4 yet to come) of the "mirror images".

but also of y: e.g., (8.13) is exactly equivalent to (4.1) and (4.2) together. That is why these three systems cannot be partitioned by detaching an investment optimality condition. By contrast, in each of the other seven systems in Tables 1 and 2 the condition on r involves only p, k and w (apart from r itself). The subsystem consisting of all the other conditions describes either (i) a plain SRP maximum, in the case of (7.2)–(7.3) or (8.4), or (ii) an SRP maximum without a duality gap, in all the other five cases. A plain SRP maximum can have a duality gap (see Appendix A), in which case it is not a potential LRP maximum. Where a subsystem does exclude a gap between SRP maximization and its dual, it may do so explicitly by the condition that $\overline{\Pi}_{\mathrm{SR}} = \Pi_{\mathrm{SR}}$ at (p,k,w), or implicitly by the condition(s) involving one or two subdifferentials of Type One ($\partial_{p,w}\overline{\Pi}_{\mathrm{SR}}$, or $\partial_p\overline{\Pi}_{\mathrm{SR}}$ and $\widehat{\partial}_w\underline{C}_{\mathrm{SR}}$ together). In one case, only the entire subsystem, (6.8)–(6.10), excludes the gap.[37]

11. Extended Wong-Viner Theorem and other transcriptions from SRP to LRC

The preceding analysis can be re-applied to SRC minimization as a subprogramme of LRC minimization, instead of SRP maximization. As part of this, the Subdifferential Sections Lemma (Lemma C.5) can be applied to C_{SR} as the bivariate convex "parent" function of the saddle function C_{LR}, instead of the saddle function Π_{SR} as in (9.3). This shows that, with w fixed and suppressed from the notation,

$$\left.\begin{array}{c} p \in \partial_y C_{\mathrm{SR}}(y,k) \\ r \in \widehat{\partial}_k \Pi_{\mathrm{SR}}(p,k) \end{array}\right\} \Leftrightarrow (p,-r) \in \partial_{y,k} C_{\mathrm{SR}}(y,k) \Leftrightarrow \left\{\begin{array}{c} p \in \partial_y C_{\mathrm{LR}}(y,r) \\ r \in -\partial_k C_{\mathrm{SR}}(y,k) \end{array}\right. . \tag{11.1}$$

This is the *Extended Wong-Viner Theorem*. Note that the condition that $r \in -\partial_k C_{\mathrm{SR}}$ is the FOC for k to yield the infimum in the definitional formula

$$C_{\mathrm{LR}}(y,r,w) = \inf_k \{\langle r \mid k\rangle + C_{\mathrm{SR}}(y,k,w)\} \tag{11.2}$$

(which means that C_{LR} is, as a function of r, the concave conjugate of $-C_{\mathrm{SR}}$ as a function of k, with y and w fixed).

For comparison, the usual Wong-Viner Envelope Theorem for differentiable costs gives

$$\left.\begin{array}{c} p = \nabla_y C_{\mathrm{SR}}(y,k) \\ r = -\nabla_k C_{\mathrm{SR}}(y,k) \text{ i.e., } k \text{ yields the inf in (11.2)} \end{array}\right\} \Rightarrow p = \nabla_y C_{\mathrm{LR}}(y,r). \tag{11.3}$$

Comparisons with the two "outer" systems in (11.1) show that their equivalence is indeed an extension of (11.3). This is because

$$\widehat{\partial}_k \Pi_{\mathrm{SR}}(p,k) \subseteq -\partial_k C_{\mathrm{SR}}(y,k) \quad \text{when } p \in \partial_y C_{\mathrm{SR}}(y,k) \tag{11.4}$$

[37] The subsystem's condition that $\underline{C}_{\mathrm{SR}} = C_{\mathrm{SR}}$ at (y,k,w) excludes a *different* gap, and on its own it does not imply that $\overline{\Pi}_{\mathrm{SR}} = \Pi_{\mathrm{SR}}$ at (p,k,w) when y maximizes short-run profit at (p,k,w): see Appendix A for an example (in which $\underline{C}_{\mathrm{SR}} = C_{\mathrm{SR}}$ trivally because the technology has no variable input).

SRP subdiff. sys. (8.4)–(8.5)		SRC-P saddle diff. sys. (7.1)–(7.3)
$(y,-v) \in \partial_{p,w}\Pi_{\mathrm{SR}}$ (Type Two) $r \in \widehat{\partial}_k \Pi_{\mathrm{SR}}$ (Type One)	Dual Part. Inv. Rule $\Longleftrightarrow$	$p \in \partial_y C_{\mathrm{SR}}$ $v \in \widehat{\partial}_w C_{\mathrm{SR}}$ $r \in \widehat{\partial}_k \Pi_{\mathrm{SR}}$ (Type One)
$\Updownarrow$ Deriv. Prop. of Opt. Val. (twice) Absorption of No-Gap Cond.		First-Order Condition $\Updownarrow$ Deriv. Prop. of Opt. Val. (twice) Absorption of No-Gap Cond.
SRP opt. sys. (6.1)–(6.3)		**split SRP opt. sys. (4.2)–(4.5)**
(y,v) maxi'es short-run profit r minimizes fixed-input value $\overline{\Pi}_{\mathrm{SR}} = \Pi_{\mathrm{SR}}$ at (p,k,w)	Two-stage solving $\Longleftrightarrow$	y maximizes revenue less C_{SR} v minimizes short-run cost r minimizes fixed-input value $\overline{\Pi}_{\mathrm{SR}} = \Pi_{\mathrm{SR}}$ at (p,k,w)
$\Updownarrow$ Deriv. Prop. of Opt. Val. (twice) Absorption of No-Gap Cond.		
FIV subdiff. sys. (8.6)–(8.7)		**FIV saddle diff. sys. (8.1)–(8.3)**
$(y,-v) \in \partial_{p,w}\overline{\Pi}_{\mathrm{SR}}$ (Type One) $r \in \widehat{\partial}_k \overline{\Pi}_{\mathrm{SR}}$ (Type Two)	Subdiff. Sect. Lem. $\Longleftrightarrow$	$y \in \partial_p \overline{\Pi}_{\mathrm{SR}}$ $v \in \widehat{\partial}_w \underline{C}_{\mathrm{SR}}$ (Type One) $r \in \widehat{\partial}_k \overline{\Pi}_{\mathrm{SR}}$

Table 1. The SRP optimization system with its split form, and four derived differential systems (of which three follow directly by the DP and FOC, and one indirectly by the SSL).

i.e., when y yields the supremum in (3.13).[38] In the differentiable case, the inclusion (11.4) reduces to the equality $\nabla_k \Pi_{\mathrm{SR}} = -\nabla_k C_{\mathrm{SR}}$ (when $p = \nabla_y C_{\mathrm{SR}}$), and thus (11.1) becomes:

$$\text{if } r = -\nabla_k C_{\mathrm{SR}}(y,k) \text{ then } \left(p = \nabla_y C_{\mathrm{SR}}(y,k) \Leftrightarrow p = \nabla_y C_{\mathrm{LR}}(y,r)\right) \tag{11.5}$$

which is the usual Wong-Viner Theorem.

Comment (failure of naive extension): The Wong-Viner Theorem *cannot* be extended to the general, subdifferentiable case simply by transcribing the ∇'s to ∂'s in (11.5) or

[38] The inclusion (11.4) follows directly from (3.13) by Remark C.7 (applied to the saddle function Π_{SR} as a partial conjugate of C_{SR}).

OFIV subdiff. sys. (8.14)–(8.15)		FIV saddle diff. sys. (8.1)–(8.3)
$v \in \widehat{\partial}_w \underline{C}_{\mathrm{SR}}$ (Type One) $(p, -r) \in \partial_{y,k} \underline{C}_{\mathrm{SR}}$ (Type Two)	Dual Part. Inv. Rule $\Longleftrightarrow$	$y \in \partial_p \overline{\Pi}_{\mathrm{SR}}$ $v \in \widehat{\partial}_w \underline{C}_{\mathrm{SR}}$ (Type One) $r \in \widehat{\partial}_k \overline{\Pi}_{\mathrm{SR}}$
$\Updownarrow$ Deriv. Prop. of Opt. Val. (twice) Absorption of No-Gap Cond.		First-Order Condition $\Updownarrow$ Deriv. Prop. of Opt. Val. (twice) Absorption of No-Gap Cond.
SRC opt. sys. (6.7)–(6.9)		split SRC opt. sys. (6.8)–(6.11)
v minimizes short-run cost (p, r) maxs rev. $-$ fix.-inp. val. $\underline{C}_{\mathrm{SR}} = C_{\mathrm{SR}}$ at (y, k, w)	Two-stage solving $\Longleftrightarrow$	p maximizes revenue less $\overline{\Pi}_{\mathrm{SR}}$ v minimizes short-run cost r minimizes fixed-input value $\underline{C}_{\mathrm{SR}} = C_{\mathrm{SR}}$ at (y, k, w)
$\Updownarrow$ Deriv. Prop. of Opt. Val. (twice) Absorption of No-Gap Cond.		
SRC subdiff. sys. (8.12)–(8.13)		SRC-P saddle diff. sys. (7.1)–(7.3)
$v \in \widehat{\partial}_w C_{\mathrm{SR}}$ (Type Two) $(p, -r) \in \partial_{y,k} C_{\mathrm{SR}}$ (Type One)	Subdiff. Sect. Lem. $\Longleftrightarrow$	$p \in \partial_y C_{\mathrm{SR}}$ $v \in \widehat{\partial}_w C_{\mathrm{SR}}$ $r \in \widehat{\partial}_k \Pi_{\mathrm{SR}}$ (Type One)

TABLE 2. The SRC optimization system with its split form, and four derived differential systems (of which three follow directly by the DP and FOC, and one indirectly by the SSL).

(11.3) because, even when $r \in -\partial_k C_{\mathrm{SR}}(y, k)$,

$$p \in \partial_y C_{\mathrm{SR}}(y, k) \nRightarrow p \in \partial_y C_{\mathrm{LR}}(y, r) . \tag{11.6}$$

It is the *reverse* inclusion that always holds, i.e.,

$$\text{if } r \in -\partial_k C_{\mathrm{SR}}(y, k) \ \text{ then } \partial_y C_{\mathrm{LR}}(y, r) \subseteq \partial_y C_{\mathrm{SR}}(y, k) \tag{11.7}$$

but the inclusion is generally strict (i.e., $\partial_y C_{\mathrm{LR}} \neq \partial_y C_{\mathrm{SR}}$).[39] Our extension (11.1) succeeds because it strengthens the insufficient condition $r \in -\partial_k C_{\mathrm{SR}}$ in (11.6) to $r \in \widehat{\partial}_k \Pi_{\mathrm{SR}}$ (this is stronger because the inclusion in (11.4) is usually strict, when C_{SR} is nondifferentiable).

[39]The inclusion (11.7) follows directly from (11.2) by Remark C.7 (applied to the saddle function C_{LR} as a partial conjugate of C_{SR}).

The peak-load pricing example of Section 2 provides a simple, yet extreme, illustration: that $r \in -\partial_k C_{\mathrm{SR}}(y,k,w)$ says merely that $r \geq 0$, with $r = 0$ if $k > \sup_t y(t)$. By contrast, the condition $r = \partial \Pi_{\mathrm{SR}} / \partial k = \int (p(t) - w)^+ \, \mathrm{d}t$ specifies r and is therefore much stronger (if $p \in \partial_y C_{\mathrm{SR}}(y,k,w)$, i.e., if: $y(t) = k$ when $p(t) > w$, and $y(t) = 0$ when $p(t) < w$). That it is strong enough to ensure that $p \in \partial_y C_{\mathrm{SR}}(y,k) \Rightarrow p \in \partial_y C_{\mathrm{LR}}(y,r)$ can also, in that example, be checked by calculating both subdifferentials explicitly.

It follows from (11.1) that LRP maximisation, being equivalent to (7.1)–(7.3), is also equivalent to the system

$$p \in \partial_y C_{\mathrm{LR}}(y,r,w) \tag{11.8}$$

$$r \in -\partial_k C_{\mathrm{SR}}(y,k,w) \tag{11.9}$$

$$v \in \widehat{\partial}_w C_{\mathrm{SR}}(y,k,w)\,. \tag{11.10}$$

We call it the *L-SRC saddle differential system*, since it uses $\partial_k C_{\mathrm{SR}}$ and $\widehat{\partial}_w C_{\mathrm{SR}}$, the partial sub/super-differentials of C_{SR} as a saddle (convex-concave) function of (k,w), in addition to using $\partial_y C_{\mathrm{LR}}$. It is the "mirror image" of the SRC-P saddle differential system (7.1)–(7.3), so it can be obtained by re-applying the same arguments (with LRC instead of SRP). It can also be derived from (7.1)–(7.3), and also from the SRC subdifferential system (8.12)–(8.13), by using (11.1).

When the producer is a public utility, LRMC pricing and LRC minimization—i.e., Conditions (11.8) to (11.10)—are often taken as the definition of a long-run producer optimum. If the SRC function is simpler than the LRC function (as is usually the case), and the SRP function is also simple, then the Extended Wong-Viner Theorem (11.1) can facilitate the short-run approach by characterizing optimality in terms of the SRC and SRP functions. This has been used in the introductory peak-load pricing example of Section 2). In that problem, the cost-minimizing inputs were obvious, but the question was how to ensure, by a simple condition put in terms of a short-run value function, that an SRMC output price was actually an LRMC price, i.e., that it met (11.8). This was achieved by employing the special case (2.2) of (7.1), i.e., of the condition that $r \in \widehat{\partial}_k \Pi_{\mathrm{SR}}$. Thus the argument was a case of the Extended Wong-Viner Theorem or, in other words, of the equivalence of (7.1)–(7.3) to (11.8)–(11.10).

Like (7.1)–(7.3), the other split-optimization and partial-subdifferential systems of Sections 4 and 6–8 (shown on the right in Tables 1 and 2) can also be transcribed into equivalent characterizations of a long-run producer optimum by replacing the SRP with the LRC.[40] Just as (7.1)–(7.3) transcribes into (11.8)–(11.10), so the other three systems transcribe into:

(1) The *split LRC optimization system* (a transcription of (4.2)–(4.5)), which is

(11.11) k minimizes $\langle r \mid \cdot \rangle + C_{\mathrm{SR}}(y,\cdot,w)$ on K (given y, r and w).

(11.12) v solves (3.10)–(3.11), given (y,k,w).

(11.13) p solves (5.5), given (y,r,w).

[40] In detail, this is done by swapping p with $-r$ and y with k, and by replacing the function $(p,k) \mapsto \Pi_{\mathrm{SR}}(p,k)$ with $(y,-r) \mapsto C_{\mathrm{LR}}(y,r)$.

$$\underline{C}_{\mathrm{LR}}(y,r,w) = C_{\mathrm{LR}}(y,r,w)\,. \tag{11.14}$$

Here, two-stage solving means first minimizing $\langle w \mid v \rangle$ over v to find the solution $\check{v}$ and the minimum value $C_{\mathrm{SR}} = \langle w \mid \check{v} \rangle$ as functions of (y,k,w), and then minimizing $\langle r \mid k \rangle + C_{\mathrm{SR}}(y,k,w)$ over k to find the solution $\check{k}(y,r,w)$. This gives the complete solution (in terms of y, r and w) as the pair $\check{k}(y,r,w)$ and $\check{v}\left(y,\check{k}(y,r,w),w\right)$.

(2) The *OV saddle differential system* (a transcription of (8.1)–(8.3)), which is

$$p \in \partial_y \underline{C}_{\mathrm{LR}}(y,r,w) \tag{11.15}$$

$$k \in \widehat{\partial}_r \underline{C}_{\mathrm{LR}}(y,r,w) \tag{11.16}$$

$$v \in \widehat{\partial}_w \underline{C}_{\mathrm{SR}}(y,k,w)\,. \tag{11.17}$$

(3) The system

(11.18) v solves (3.10)–(3.11), given (y,k,w).

(11.19) r minimizes $\langle \cdot \mid k \rangle - \underline{C}_{\mathrm{LR}}(y,\cdot,w)$ on R (given y, k and w).

(11.20) p solves (5.5), given (y,r,w).

$$\underline{C}_{\mathrm{SR}}(y,k,w) = C_{\mathrm{SR}}(y,k,w)\,. \tag{11.21}$$

This may be called the *reverse-split SRC optimization system*, to distinguish it from (6.8)–(6.11), of which it is a transcription. (The two systems differ only in the order in which p and r are optimized when the joint programme (5.4) is split in two stages: in (6.8)–(6.11), the first stage is to find r in terms of p and calculate $\overline{\Pi}_{\mathrm{SR}}$, whereas in (11.18)–(11.21), the first stage is to find p in terms of r and calculate $\underline{C}_{\mathrm{LR}}$.)

12. Outline of the short-run approach to long-run general equilibrium

The preceding characterizations of long-run producer optimum can serve various purposes; ours is the short-run approach to long-run general equilibrium (LRGE). This means that the capital inputs k are kept fixed at the stage of calculating the equilibrium in the products' market. The variable-input prices w are assumed to be fixed throughout our analysis (although this is not at all essential, and w might instead be determined in equilibrium just like the output prices p). This leaves two alternative ways to handle the supply side of the short-run general equilibrium (SRGE) problem, and hence two varieties of the short-run approach to long-run producer optimum and general equilibrium:

(1) In the *short-run profit approach*, the output and variable-input quantities $\hat{y}$ and $\check{v}$, and the fixed-input values $\hat{r}$, are derived from any given p, k and w (usually by solving the SRP problem (3.6)–(3.7) and its dual (5.6) or (5.13)–(5.14) under c.r.t.s.). The supply $\hat{y}(p,k,w)$ is then equated to demand $\hat{x}(p)$ to determine the short-run equilibrium price system $p^{\star}_{\mathrm{SR}}(k)$, which depends also on w. This stage corresponds to the inner loop in Figure 3, if an iterative method is used

to solve the demand-supply equation for p.[41] The capital inputs' marginal values $\hat{r}\left(p^{\star}_{\mathrm{SR}}(k,w),k,w\right)$, imputed at the short-run equilibrium prices, are then equated to their *given*, fixed rental prices r^{F} to determine, by solving for k, the (long-run) equilibrium capacities $k^{\star}\left(r^{\mathrm{F}},w\right)$. This also gives the long-run equilibrium price system $p^{\star}_{\mathrm{LR}}\left(r^{\mathrm{F}},w\right) = p^{\star}_{\mathrm{SR}}\left(k^{\star}\left(r^{\mathrm{F}},w\right),w\right)$. This stage corresponds to the outer loop in Figure 3, if an iterative method is used to solve the price-value equation for k.

(2) In the short-run *cost approach*, the variable-input quantities $\check{v}$, and the shadow prices for outputs and fixed inputs—i.e., a typically nonunique $p \in \check{P}(y,k,w)$ with the associated, typically unique $\hat{r}(p,k,w)$—are derived from any given y, k and w (usually by solving the SRC problem (3.10)–(3.11) and its dual (5.4) or (5.9)–(5.10) under c.r.t.s.). To find the short-run equilibrium, inverse demand is then required to equal one of the typically nonunique output price systems that solve the short-run output-pricing programme in (6.10). This a subprogramme of (5.4); its solution set $\check{P}(y,k,w)$ consists essentially of SRMCs (see (12.3) for details). Finally, the long-run equilibrium capacities, and hence also the output prices, are found just as in the profit approach.

In principle, the duality theory of convex programming can be brought to bear however the commodities are divided into "variable" quantities with given prices and "fixed", unpriced quantities: in studying the producer optimum, the roles of prices and quantities are formally symmetric. At an abstract level, therefore, there is no reason to prefer any particular programme pair or the associated functional representation of the technology (by Π_{SR}, C_{LR} or C_{SR}, etc.). But the classification of commodities as "fixed in the short run" is not arbitrary and nominal but mostly real and objective: these are capital goods and natural resources. Their quantities (k) must be taken as known throughout the short-run analysis. Additionally, some of those quantities to be determined in the SRGE, such as the outputs (y), might also be taken as known at the earlier stage of finding the short-run producer optimum and the shadow prices: this would mean solving the SRC programme (for v) with its dual (for p and r). But this is disadvantageous analytically because, when the capital inputs (k) impose capacity constraints on a cyclic output (y), it results in dual solutions so indeterminate that they form an unbounded set: if not only r but also p are unknowns, then almost nothing can be said about capacity charges (which are terms of p, and give r as their total over the cycle). Another disadvantage of the SRC approach, which emerges only at the equilibrium stage, is that it entails working with the inverse supply maps ($\check{P}_{\theta}$) and "equating" each of these to inverse demand to find the SRGE output bundle ($y^{\star}_{\mathrm{SR}\,\theta}$) of each individual producer θ—from the inclusion a.k.a. "generalized equation" (12.2) below. This is usually much harder than simply adding up all the direct supply maps ($\hat{Y}_{\theta}$), equating their sum to demand, and solving (12.2) for the single market price system ($p^{\star}_{\mathrm{SR}}$)—which is what the SRP approach requires. In addition, unlike the multi-valued inverse supply map ($\check{P}_{\theta}$), the direct supply may well

[41] In finding $p^{\star}_{\mathrm{SR}}$ by Walrasian tatonnement, a manageable difficulty arises from discontinuity of supply when it is only an upper hemicontinuous correspondence (as in Figure 1a). With a continuous (single-valued) demand map, this is not much of a complication.

be a single-valued map $(\hat{y}_\theta)$, in which case the relevant inclusion (12.2) is an ordinary equation.

In summary, it is better not to fix any more quantities than is necessary—and this means using the SRP rather than the SRC approach. The profit approach is likely to be more workable because it has two advantages over the cost approach: (i) determinacy of solutions to the short-run producer problem and its dual, and (ii) reduction of the number of unknowns in the subsequent equilibrium problem. Both are detailed next.

The first advantage is simply the convenience of dealing mostly with single-valued maps rather than multi-valued correspondences. Solutions for (p,r) to the dual (5.4) of the SRC problem are typically nonunique: indeed, the set of optimal (p,r)'s is unbounded because, in pure SRC calculations, the capacity premium is completely indeterminate (except when it vanishes because there is excess capacity). But the r associated with a particular p may well be unique, and so may y and v (as we have tacitly assumed by using the notation $\hat{r}$ and $\hat{y}$ in describing the short-run approach). That is, solutions for r and (y,v) to the SRP problem (3.6)–(3.7) and its dual (5.6) can both be expected to be unique or, at the very least, to form bounded sets. This can be illustrated with an elementary but instructive example. Suppose for simplicity that there is no variable input, and that $\mathbb{Y}$ is a cone. A long-run producer optimum is then described by the Complementarity Conditions (3.5), i.e.,

$$(y,-k) \in \mathbb{Y},\ (p,r) \in \mathbb{Y}^\circ \text{ and } \langle p \,|\, y \rangle = \langle r \,|\, k \rangle .$$

In the profit approach (given p and k), both inclusions are useful in solving this system for y and r. But in the cost approach (given y and k), the first inclusion restricts only the data—so, when it is met, it is of no help at all in solving for p and r. The simplest example is $\mathbb{Y} = \{(y,-k) \in \mathbb{R}^2 : y = k\}$; then $\mathbb{Y}^\circ = \{(p,r) \in \mathbb{R}^2 : p = r\}$. In the cost approach the level of (p,r) is indeterminate, but in the profit approach both solutions are unique, viz., $(\hat{y},\hat{r}) = (k,p)$.[42] This principle is also borne out by more significant and complex examples such as peak-load pricing with storage, in which the optimum $\hat{r}(p,k,w)$ or $\hat{y}(p,k,w)$ is shown to be unique if the TOU tariff p is, respectively, a continuous or plateau-less function of time: see Section 16 here, or [21], [27, Sections 6 to 9] and [24].

The second, and more significant, advantage of the SRP approach over the SRC approach emerges, at the equilibrium stage, whenever there is a number of producers, with technologies $\mathbb{Y}_\theta$ for $\theta \in \Theta$. In the profit approach, the short-run equilibrium is found by equating the demand $\hat{x}(p)$ to the profit-maximizing total output $\sum_\theta \hat{y}_\theta(p,k_\theta,w)$ and solving for p; when the optimal output is nonunique, one solves for p the inclusion

$$\hat{x}(p) \in \sum_\theta \hat{Y}_\theta(p,k_\theta,w) \tag{12.1}$$

[42] When there *are* variable inputs whose cost-minimizing quantities $\check{v}$ are known functions of the data (y,k,w), the condition $(y,-k,-v) \in \mathbb{Y}$ in (3.5) boils down to $(y,-k,-\check{v}(y,k,w)) \in \mathbb{Y}$, which is again a pure restriction on the data with no information about the unknowns p and r. Of course, the profit approach would have a similar comparative weakness in the condition $(p,r,w) \in \mathbb{Y}^\circ$ if the fixed-input values $\hat{r}$ were easily calculated functions of the data (p,k,w). But the programme that we take to be readily soluble, without using duality, is the SRC programme for v, and not the dual of the SRP programme for r.

where $\hat{Y}_\theta$ is the solution set for the reduced SRP programme in (3.13) and (4.2). For comparison, the cost approach requires solving, for the output bundles (y_θ), the inclusion

$$\widetilde{p}\left(\sum_\theta y_\theta\right) \in \bigcap_\theta \check{P}_\theta\left(y_\theta, k_\theta, w\right) \tag{12.2}$$

where $\widetilde{p}$ is the inverse demand map and $\check{P}_\theta\left(y_\theta, k_\theta, w\right)$ is the solution set for the short-run output-pricing programme in (6.10), i.e., $\check{P}_\theta$ is essentially $\partial_y C_{\mathrm{SR}}^\theta$, the multi-valued SRMC of an individual plant. This route is likely to be more difficult because, with multiple producers, it means having to solve for a number of variables (y_θ) instead of the single variable p, as well having to intersect the price sets $\left(\check{P}_\theta\right)$ to start with. And these are large, unbounded sets if the fixed inputs impose capacity constraints.

Comments (the relative complexity of the cost approach):

(1) It is not even easy just to identify all those output allocations (y_θ) with $\bigcap_\theta \check{P}_\theta \neq \emptyset$ in (12.2), since this involves splitting the total output among the plants in a cost-minimizing way, which can be a difficult problem (known as optimal system despatch in the context of electricity generation). To see this in detail, note that

$$\text{if } \underline{C}_{\mathrm{SR}}^\theta = C_{\mathrm{SR}}^\theta \text{ at } \left(y_\theta, k_\theta, w\right) \text{ then } \check{P}_\theta \subseteq \partial_y C_{\mathrm{SR}}^\theta \text{ at } \left(y_\theta, k_\theta, w\right) \tag{12.3}$$

by Lemma 19.22.[43] So $\bigcap_\theta C_{\mathrm{SR}}^\theta$ is nonempty if $\bigcap_\theta \check{P}_\theta$ is. Furthermore, the industry's SRC as a function of its total output x is

$$\inf_{(y_\theta)_{\theta\in\Theta}} \left\{\sum_{\theta\in\Theta} C_{\mathrm{SR}}^\theta\left(y_\theta, k_\theta, w\right) : \sum_{\theta\in\Theta} y_\theta = x\right\} \tag{12.4}$$

i.e., it is the infimal convolution of the individual plants' operating cost functions $C_{\mathrm{SR}}^\theta\left(\cdot, k_\theta, w\right)$, abbreviated to C^θ. With $\triangle$ denoting the convolution operator, one has $p \in \bigcap_\theta \partial C^\theta\left(y_\theta\right)$ if and only if both $p \in \partial\left(\triangle_\theta C^\theta\right)\left(\sum_\theta y_\theta\right)$ and $\left(\triangle_\theta C^\theta\right)\left(\sum_\theta y_\theta\right) = \sum_\theta C^\theta\left(y_\theta\right)$: see, e.g., [36, 6.6.3 and 6.6.4]. The "only if" part shows that if $\bigcap_\theta \partial C^\theta\left(y_\theta\right) \neq \emptyset$, then (y_θ) is a cost-minimizing split of the total output $\sum_\theta y_\theta$ among the plants with the given capacities (k_θ) and technologies $(\mathbb{Y}_\theta)$. This means that competitive profit maximization, by the choice of outputs (y_θ) at a *common* output price p, leads to such an optimal allocation of the total output.

(2) Thus the decentralized, plant-by-plant derivation of the industry's total supply (given a common output price p) by-passes the problem of the cost-minimizing allocation of any *given* total output x, which is usually much more complex than the individual profit-maximizing operation problems. For example, cost-minimizing despatch of a hydro-thermal electricity-generating system necessitates a CP with no simple form for either the primal or the dual solution: see the policy construction in [35, pp. 201–219]. By contrast, profit-maximizing operation of a hydro plant (or a storage plant) is an LP whose solution has a relatively simple structure: see Section 16 here, [24] and [21] or [27, Section 5].

[43] Also, even when $\check{P}_\theta \subsetneq \partial_y C_{\mathrm{SR}}^\theta$ at $\left(y_\theta, k_\theta, w\right)$, the two sets have the same intersection with the set $\left\{p : \overline{\Pi}_{\mathrm{SR}}^\theta\left(p, k_\theta, w\right) = \Pi_{\mathrm{SR}}^\theta\left(p, k_\theta, w\right)\right\}$, by Corollary 19.23.

Our description of either variety, SRP or SRC, of the short-run approach assumes the use of either the SRP or the SRC optimization system (or its split form). Of the optimization systems, this is the one directly suited to the purpose; and when the technology is given by a production set (as in an engineering specification), there may be no tractable formulae for the value functions, and hence no usable alternative among differential systems. A differential system is likely to be useful only when each of the profit or cost functions it uses is either easy to calculate (by solving the relevant programme), or is simply given as a definition of the technology (as in econometric uses of duality). These remarks can be expanded as follows.

Comments (on choosing a system for a short-run approach):

(1) What defines a particular *approach* to the producer problem is which of its price and quantity variables are treated as *known* and which as *unknown*. With three groups of commodities, there are eight (2^3) possibilities: the knowns-unknowns patterns of the SRP approach is (p, k, w)-(y, r, v), whilst that of the SRC approach is (y, k, w)-(p, r, v). Either approach may use its "own" (SRP or SRC) optimization system, but it might also use the LRC system for the same purpose, viz., to determine r and v and either y or p from any known k, w and either p or y (thus solving not a long-run problem, but a short-run profit or cost problem with its dual). Indeed, either variety of the short-run approach may use whichever of the equivalent systems is most convenient: in principle, it need not matter whether a producer optimum is characterized in terms of short or long run, profit or cost, optima or marginal values.

(2) Within optimization systems, every choice leads to the same analysis if duality is used: all the systems lead to the same FFE Conditions (viz., Complementarity (3.5)), and also to the same Kuhn-Tucker Conditions (once the constraint sets $\mathbb{Y}$ and $\mathbb{Y}^{\circ}$ are represented by systems of inequalities).[44] When analyzed by either of these duality methods (Kuhn-Tucker's or FFE), all the optimization systems become therefore identical—but even so it simplifies the terminology to start from the approach's "own" system, i.e., the one whose programme data and decisions are, respectively, the knowns and the unknowns of the chosen approach. (In the short-run profit approach, this means using the SRP optimization system, as is done in Section 13.) Then "solving the programmes for their decisions" means exactly the same as "solving the system for the unknowns of the approach", which is what is to be done.

(3) If a different, "non-own" pair of programmes were solved—for *its* decisions in terms of *its* data—then the whole solution correspondence (data-to-decisions) would have to be obtained and part-inverted to express the unknowns in terms of the knowns (thus compensating for the mismatch between these and the data and decisions). This may be worthwhile, but only when a "non-own" programme is particularly easy to solve without using duality (since the use of duality leads from any programme pair to the same Kuhn-Tucker or FFE system).

(4) When there is such a readily soluble programme and its value function is easy not only to calculate but also to differentiate, it may be best to use the corresponding

[44]For c.f.c. techniques, the Kuhn-Tucker Conditions are spelt out in Section 24.

differential system. This may be a "non-own" system, i.e., one in which the arguments and the derivatives of the function do not correspond to the knowns and the unknowns of the approach. In such a case, after calculating the subdifferential correspondence, one must part-invert it as required. The method may be useful when there is no explicit formula for the chosen approach's "own" function (whose arguments and derivatives are, respectively, the knowns and unknowns of the approach), but there *is* a formula for another value function. For example, there is no general formula for the SRP of a c.f.c. technique, but the SRP approach might be based on a formula for the LRC (24.3) or the SRC (24.21). However, this is worthwhile only if the input requirement functions ($\check{k}$ and $\check{v}$) are simple enough. When they are not, it is better to use an optimization system.[45]

(5) It might seem that those (seven) systems are preferable which decompose in the way discussed in Comments in Section 10. When such a system is used for the SRP approach to LRGE, the calculation of SRGE requires only the subsystem but not the supplementary condition—i.e., this stage requires solving the SRP programme (3.6)–(3.7) for (y, v), but it need not include shadow-pricing the fixed inputs by solving the dual programme (5.6), or (5.13)–(5.14) under c.r.t.s., for r (or possibly by differentiating Π_{SR} or $\overline{\Pi}_{\mathrm{SR}}$ w.r.t. k). But this does not save on computation if, as is usual, the SRP programme has to be solved by a duality method: the dual is then being solved together with the primal anyway.

13. A framework for the short-run profit approach to long-run general equilibrium

The equilibrium framework set out next is designed to price a range of commodities with joint costs of production. The product range can be a single good differentiated over commodity characteristics, such as time. Such a differentiated good is usually produced by a variety of techniques; this is so in the motivating application to the peak-load pricing of electricity (Sections 15 to 17).

To concentrate on the issues of investment and pricing for the differentiated output of a particular Supply Industry (SI), we simplify the equilibrium model by aggregating commodities on the basis of some fixed relative prices. As a result, there are just two consumption goods apart from the differentiated good—viz., the numeraire (measured in \$) and a produced final good which is a homogeneous composite representing those commodities whose production requires an input of the differentiated good. The prices for most of the SI's inputs, including all the variable inputs, are also assumed to be given. But, to keep the equilibrium capacities (and the variable inputs) as explicit entries of the equilibrium allocation, we choose *not* to aggregate these inputs with the numeraire (despite their fixed prices).

The Supply Industry's technology consists of a finite number of production techniques, each of which uses a different set of input commodities to produce the same set of output commodities. For each technique $\theta \in \Theta$, its sets of the fixed and the variable inputs are

[45] This is how we choose to deal, in Sections 16 and 17, with the pumped-storage technique (15.4) because the subdifferential of the storage capacity requirement function (15.6), calculated in [21], is not particularly simple (even under the simplifying assumption of perfect energy conversion).

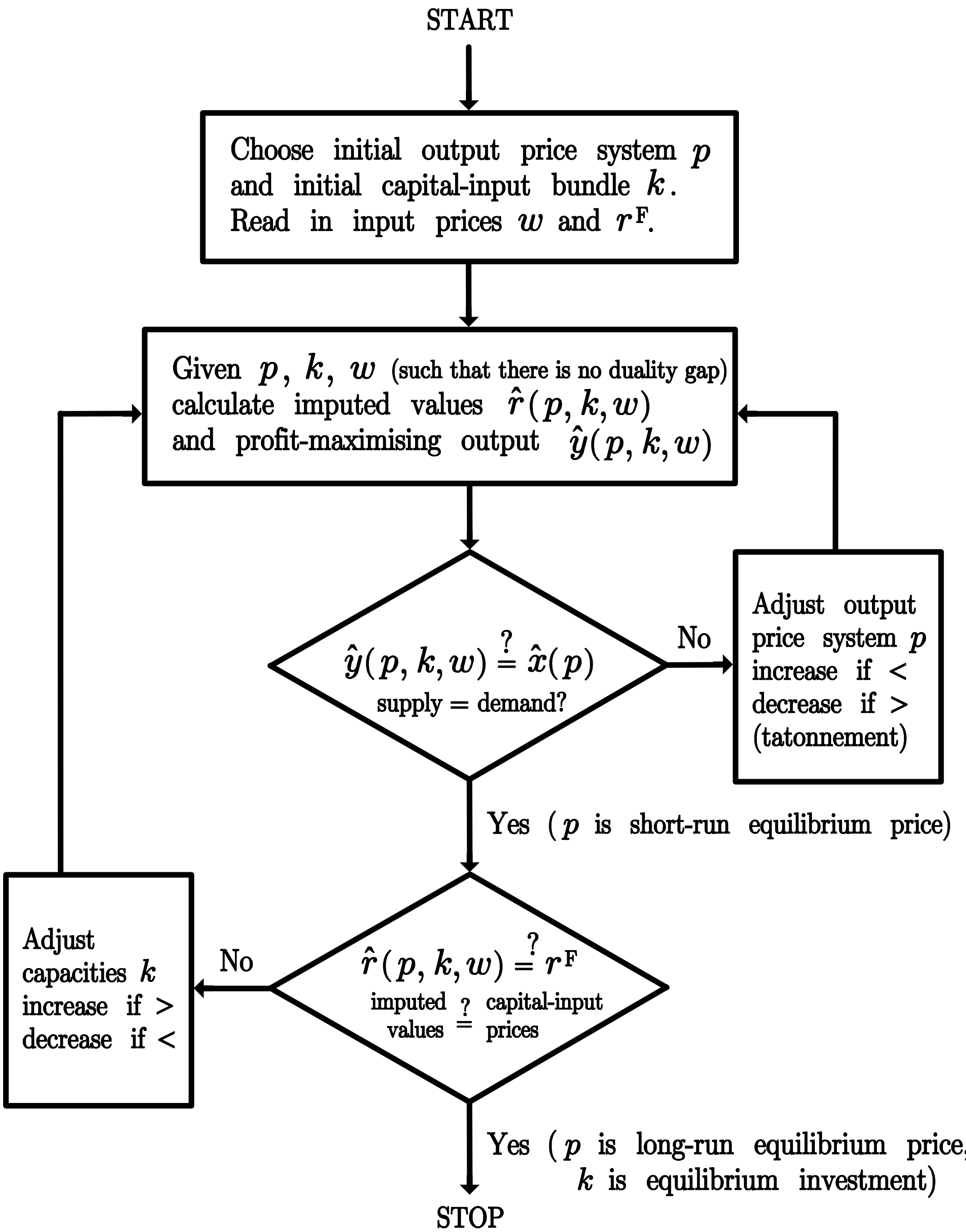

FIGURE 3. Flow chart for iterative implementation of SR profit approach to LR general equilibrium. For simplicity, all demand for outputs is taken to be consumer demand that is independent of profit income, and all input prices are fixed (in numeraire terms). Absence of duality gap and existence of optima ($\hat{r}$, $\hat{y}$) can be ensured by using the results of Sections 20 to 23.

denoted by Φ_θ and Ξ_θ; and its long-run production set is taken to be a convex cone

$$\mathbb{Y}_\theta \subset Y \times \mathbb{R}^{\Phi_\theta} \times \mathbb{R}^{\Xi_\theta}. \tag{13.1}$$

Thus $\mathbb{Y}_\theta$ lies in a space that depends on θ. To be formally regarded as a subset of the full commodity space, $\mathbb{Y}_\theta$ must be embedded in it as $\mathbb{Y}_\theta \times \{(0, 0, \ldots)\}$, i.e., by inserting zeros in the input-output bundle at the other positions.

Investment in technique θ is denoted by $k_\theta \in \mathbb{R}^{\Phi_\theta}$; so the SI's total investment in fixed input ϕ is

$$\sum_{\theta \colon \phi \in \Phi_\theta} k_{\theta\phi} \quad \text{for } \phi \in \Phi_\Theta := \bigcup_{\theta \in \Theta} \Phi_\theta \tag{13.2}$$

which is the SI's set of fixed inputs. When the sets Φ_θ are pairwise disjoint, the sum in (13.2) reduces to a single term (for each ϕ), and the notation can be simplified: see (13.20), etc.

The set of all the fixed inputs of the SI, Φ_Θ, is partitioned into two subsets: Φ_Θ^{F} consisting of those with given prices, and Φ_Θ^{E} consisting of those whose prices are determined only in long-run equilibrium. For a particular technique θ, its set of fixed inputs Φ_θ is thus partitioned into two subsets

$$\Phi_\theta^{\mathrm{E}} := \Phi_\Theta^{\mathrm{E}} \cap \Phi_\theta \quad \text{and} \quad \Phi_\theta^{\mathrm{F}} := \Phi_\Theta^{\mathrm{F}} \cap \Phi_\theta.$$

An input $\phi \in \Phi_\Theta^{\mathrm{F}} = \bigcup_{\theta \in \Theta} \Phi_\theta^{\mathrm{F}}$ is supplied at a fixed unit cost r_ϕ^{F} (in terms of the numeraire), so its total supply cost is linear. By contrast, the total supply cost of an input $\phi \in \Phi_\Theta^{\mathrm{E}} = \bigcup_{\theta \in \Theta} \Phi_\theta^{\mathrm{E}}$ is given by a convex function, G_ϕ, of the supplied quantity k_ϕ. Typically, G_ϕ is a strictly convex and increasing, finite function on an interval $\left[0, \overline{k}_\phi\right]$, with $G_\phi(0) = 0$. But the case of an input in a fixed supply $\overline{k}_\phi$ (without free disposal) is captured by setting $G_\phi(k_\phi)$ equal to 0 for $k_\phi = \overline{k}_\phi$ and to $+\infty$ otherwise (in which case the equilibrium condition that $r_\phi \in \partial G_\phi(k_\phi)$ means merely that $k_\phi = \overline{k}_\phi$). For examples in the electricity supply industry (ESI), see Section 17 here, or [21] and [24].

This classification of inputs will not always be clear-cut, but as a rough rule, for an industry supplying a good with a cyclical demand, its fixed inputs are those whose adjustment takes longer than one demand cycle: even if the cycle is a year, this is usually just a fraction of plant construction times. Variable inputs are those which can be adjusted quickly to the time-varying output rate $y_\theta(t)$. For example, fuel inputs are assumed to be instantaneously adjustable in our model of thermal electricity generation: see (15.1). The variable inputs are regarded as having fixed prices (w_ξ), e.g., by reason of being internationally traded. Likewise, a typical fix-priced capital input $\phi \in \Phi_\Theta^{\mathrm{F}}$ is internationally traded equipment, and its rental price r_ϕ^{F} is the annuity consisting of interest on the purchase price and depreciation.[46] By contrast, an equilibrium-priced capital input $\phi \in \Phi_\Theta^{\mathrm{E}}$—whose rental price r_ϕ^{E} is determined only in long-run equilibrium—is typically a factor which can only be supplied locally and at an increasing marginal cost, as a result of the fixity of some assets required for its supply (such as special sites or other natural resources). Constancy of returns to scale for the SI's technology need not

[46]Formally, the fixed prices r^{F} and w are built into the standard competitive equilibrium model by introducing a linear production set equal to the hyperplane perpendicular to the vector $\left(r^{\mathrm{F}}, w, 1\right)$ and passing through the origin in the space of the supplier's fix-priced inputs and the numeraire.

extend to its input supply, and in the application to peak-load pricing with storage the reservoir capacity has an increasing marginal cost (Section 17).

For simplicity, all input demand for the SI's products is taken to come from a single Industrial User (IU), who produces a final good from inputs of the differentiated good and the numeraire. The user's production function $F\colon Y_+ \times \mathbb{R}_+ \to \mathbb{R}$, assumed to be strictly concave and increasing, defines his production set

$$\mathbb{Y}_{\mathrm{IU}} = \{(-z; \varphi, -n) \in Y_- \times \mathbb{R} \times \mathbb{R}_- : F(z, n) \geq \varphi\} \tag{13.3}$$

where Y_+ is a convex cone that is P-closed (i.e., closed for some, and hence for every, locally convex topology on Y that yields P as the continuous dual space). When, as in superdifferentiation at the algebraic boundary points (non-core points) of $Y_+ \times \mathbb{R}_+$, the function F must be regarded as defined on the whole space $Y \times \mathbb{R}$, it is extended by setting its value to $-\infty$ outside of $Y_+ \times \mathbb{R}_+$.[47]

A complete commodity bundle, then, consists of: (i) the produced differentiated good, (ii) the Supply Industry's fixed and variable inputs, (iii) the Industrial User's product, and (iv) the numeraire. The quantities are always listed in this order; but those which are irrelevant in a particular context (and can be set equal to zero) are omitted for brevity, as in (13.1) and (13.3). A consumption bundle consists of quantities of the differentiated good, the IU's product and the numeraire; so it may be written as $(x; \varphi, m) \in Y \times \mathbb{R}^2$. A matching consumer price system is $(p; \varrho, 1) \in P \times \mathbb{R}^2$—whilst a complete price system is

$$\left(p; r^{\mathrm{E}}, r^{\mathrm{F}}; w, \varrho, 1\right) = \left(p; \left(r^{\mathrm{E}}_{\phi}\right)_{\phi \in \Phi^{\mathrm{E}}_{\Theta}}, \left(r^{\mathrm{F}}_{\phi}\right)_{\phi \in \Phi^{\mathrm{F}}_{\Theta}}; (w_\xi)_{\xi \in \Xi_\Theta}, \varrho, 1\right)$$

(where $\Xi_\Theta := \bigcup_{\theta \in \Theta} \Xi_\theta$). There is a finite set, Ho, of households; and for each $h \in \mathrm{Ho}$ its utility is a concave nondecreasing function U_h on the consumption set $Y_+ \times \mathbb{R}^2_+$.[48] It is assumed to be nonsatiated in each of the two homogeneous goods (the IU's product and the numeraire), i.e., $U_h(x; \varphi, m)$ is increasing in φ and in m; this guarantees that both prices are positive in equilibrium. Each household's initial endowment is a quantity $m^{\mathrm{En}}_h > 0$ of the numeraire only; and its share of profit from the supply of input $\phi \in \Phi^{\mathrm{E}}_\Theta$ is $\varsigma_{h\phi} \geq 0$, with $\sum_h \varsigma_{h\phi} = 1$. Similarly, $\varsigma_{h\,\mathrm{IU}}$ denotes household h's share in the User Industry's profit.

The Supply Industry's profit is zero in long-run equilibrium (because of c.r.t.s.), but an exact short-run analysis requires specifying the households' shares in the operating profits from the SI's plants—since the profit Π^θ_{SR} in (13.10) is only approximately offset by the liabilities $r^{\mathrm{EF}}_{|\theta} \cdot k_\theta$, which represents plant depreciation and interest (on the debt from which the plant is assumed to have been financed). A *plant* is specified by its type θ and by its capacities (or, more generally, its quantities of the fixed inputs) $k_{\theta\phi}$, for

[47] This matters in calculating $\widehat{\partial} F$ at a point that belongs to $Y_+ \times \mathbb{R}_+$ but *not* to its core (a.k.a. algebraic interior). To spell this out, assume that F, as a function on its effective domain $Y_+ \times \mathbb{R}_+$, has a Mackey continuous concave extension F^{Ex} defined on all of $Y \times \mathbb{R}$. Then $\widehat{\partial} F = \widehat{\partial} F^{\mathrm{Ex}}$ at any core points of $Y_+ \times \mathbb{R}_+$, but in general $\widehat{\partial} F(z, n) = \widehat{\partial} F^{\mathrm{Ex}}(z, n) + \{(\mu, \nu) \in P_+ \times \mathbb{R}_+ : \langle \mu \mid z \rangle + \nu n = 0\}$ because $F = F^{\mathrm{Ex}} - \delta(\cdot \mid Y_+ \times \mathbb{R}_+)$.

[48] Consumer preference can of course be regarded as defined on the orthant in the full commodity space $L := Y \times \mathbb{R}^{\Phi_\Theta} \times \mathbb{R}^{\Xi_\Theta} \times \mathbb{R}^2$ by positing that the consumer has no use for the Supply Industry's inputs k and v: this means regarding a utility U_h on $Y_+ \times \mathbb{R}^2_+$ as a function on $Y_+ \times \mathbb{R}^{\Phi_\Theta}_+ \times \mathbb{R}^{\Xi_\Theta}_+ \times \mathbb{R}^2_+$ defined by $(x; k, v; \varphi, m) \mapsto U_h(x; \varphi, m)$.

$\phi \in \Phi_\theta$. We assume that every plant of a particular type θ has the *same capacity ratios* $(k_{\theta 1} : k_{\theta 2} : \ldots)$; with c.r.t.s., this amounts to assuming that there is at most one plant of each type. Though this is rarely so in a real industry which has evolved over time, the condition *is* met in long-run equilibrium, the calculation of which is our main use for the short-run model. It makes sense, then, to speak of *profit shares in a technique*: denoted by $\varpi_{h\theta}$ (with $\sum_h \varpi_{h\theta} = 1$), household h's share in the operating profit from technique θ is

$$\varpi_{h\theta} := \sum_i \beta_{hi} \alpha_{i\theta}$$

where β_{hi} is h's share in producer i, and $\alpha_{i\theta}$ is i's share in the plant of type θ. (In other words, one can assume that all plants of a type are wholly owned by one and the same producer.)

Notation: The restriction, to Ξ_θ, of a w: $\Xi_\Theta \to \mathbb{R}$ is $w_{|\Xi_\theta}$, abbreviated to $w_{|\theta}$. Similarly, $r^{\mathrm{E}}_{|\theta}$ and $r^{\mathrm{F}}_{|\theta}$ mean the restrictions to Φ^{E}_θ and to Φ^{F}_θ of an r^{E}: $\Phi^{\mathrm{E}}_\Theta \to \mathbb{R}$ and an r^{F}: $\Phi^{\mathrm{F}}_\Theta \to \mathbb{R}$, respectively. The pair $\left(r^{\mathrm{E}}, r^{\mathrm{F}}\right)$ defines a case-function on $\Phi_\Theta := \Phi^{\mathrm{E}}_\Theta \cup \Phi^{\mathrm{F}}_\Theta$; it is occasionally denoted by r^{EF} for brevity.

By definition, given price systems $\left(r^{\mathrm{F}}, w\right)$ for the fix-priced capital inputs and the variable inputs, a *long-run competitive equilibrium* consists of:

- a system of prices $(p^\star, r^\star, \varrho^\star) \in P_+ \times \mathbb{R}_+^{\Phi^{\mathrm{E}}_\Theta} \times \mathbb{R}_{++}$ (all in terms of the numeraire) for: the Supply Industry's differentiated output good, the equilibrium-priced capital inputs, and the Industrial User's product
- an allocation made up of:
  - a consumption bundle $(x^\star_h, \varphi^\star_h, m^\star_h) \in Y \times \mathbb{R} \times \mathbb{R}$ for each household h
  - an input-output bundle of the Industrial User $(-z^\star, F(z^\star, n^\star), -n^\star) \in Y \times \mathbb{R} \times \mathbb{R}$
  - input-output bundles of the Supply Industry, $(y^\star_\theta, -k^\star_\theta, -v^\star_\theta) \in Y \times \mathbb{R}^{\Phi_\theta} \times \mathbb{R}^{\Xi_\theta}$ for each technique θ

that meet the following definitional conditions:

(1) *Producer optimum in Supply Industry*: For each θ,

$$(y^\star_\theta, -k^\star_\theta, -v^\star_\theta) \in \mathbb{Y}_\theta \quad \text{and} \quad \left(p^\star, \left(r^\star_{|\theta}, r^{\mathrm{F}}_{|\theta}\right), w_{|\theta}\right) \in \mathbb{Y}^\circ_\theta \tag{13.4}$$

$$\langle p^\star \mid y^\star_\theta \rangle = \left(r^\star_{|\theta}, r^{\mathrm{F}}_{|\theta}\right) \cdot k^\star_\theta + w_{|\theta} \cdot v^\star_\theta \tag{13.5}$$

i.e., the equilibrium quantities and prices meet the Complementarity Conditions (3.5), or any of the preceding equivalent systems of conditions. In other words, $(y^\star_\theta, -k^\star_\theta, -v^\star_\theta)$ maximizes (to zero) the long-run profit at prices $\left(p^\star, \left(r^\star_{|\theta}, r^{\mathrm{F}}_{|\theta}\right), w_{|\theta}\right)$.

(2) *Producer optimum in User Industry*: $(p^\star, 1) \in \varrho^\star \widehat{\partial} F(z^\star, n^\star)$.

(3) *Consumer utility maximization*: For each h, $(x^\star_h, \varphi^\star_h, m^\star_h)$ maximizes U_h on the budget set $B\left(p^\star, \varrho^\star, \hat{M}_{\mathrm{LR}\,h}(p^\star, r^\star, \varrho^\star)\right)$, where

$$B(p, \varrho, M) := \{(x, \varphi, m) \geq 0 : \langle p \mid x \rangle + \varrho\varphi + m \leq M\} \tag{13.6}$$

$$\Pi_\phi(\mathsf{r}) := \sup_{\mathsf{k}} (\mathsf{r}\mathsf{k} - G_\phi(\mathsf{k})) \quad \text{for } \mathsf{r} \in \mathbb{R} \tag{13.7}$$

$$\Pi_{\mathrm{IU}}(p,\varrho) := \sup_{z,n}\left(\varrho F(z,n) - \langle p \mid z\rangle - n\right) \tag{13.8}$$

$$\hat{M}_{\mathrm{LR}\,h}\left(p, r^{\mathrm{E}}, \varrho\right) := m_h^{\mathrm{En}} + \varsigma_{h\,\mathrm{IU}}\Pi_{\mathrm{IU}}(p,\varrho) + \sum_{\phi\in\Phi_\Theta^{\mathrm{E}}} \varsigma_{h\phi}\Pi_\phi\left(r_\phi^{\mathrm{E}}\right). \tag{13.9}$$

(4) *Market clearance*: $\sum_\theta y_\theta^\star = z^\star + \sum_h x_h^\star$ and $F(z^\star, n^\star) = \sum_h \varphi_h^\star$.
(5) *MC pricing of SI's fixed inputs*: $r_\phi^\star \in \partial G_\phi\left(\sum_\theta k_{\theta\phi}^\star\right)$ for each $\phi \in \Phi_\Theta^{\mathrm{E}}$.[49]

Comment: This is an instance of the usual equilibrium concept, except for being specialized to the case of nonzero prices ($\varrho^\star$ and 1) for the two composite goods (in particular, the above characterization of the IU's profit maximum, Condition 2, relies on the positivity of the output price $\varrho^\star$). The usual definition captures also the case of zero prices, but this cannot arise here because of our nonsatiation assumptions. In other words, price positivity is actually a property of an equilibrium (and not part of the concept itself).

The short-run profit approach to solving this system starts by fixing the SI's capital inputs $(k_\theta)_{\theta\in\Theta}$. Given these quantities as well as prices (p, w) for the SI's variable commodities, a suitably chosen system characterizing the long-run producer optimum is then solved for: the plants' outputs y_θ, their variable inputs v_θ and the values, r_θ, imputed to the fixed inputs in the plant of each type θ. The optimal outputs $\hat{y}_\theta\left(p, k_\theta, w_{|\theta}\right)$ are then equated to demand to find the short-run equilibrium price system $p_{\mathrm{SR}}^\star$, which depends on the k_θ's.[50] Finally, to determine the capacities k_θ, and the prices r^{E} of any equilibrium-priced capital inputs, the imputed value $\hat{r}_{\theta\phi}\left(p, k_\theta, w_{|\theta}\right)$ is equated either to the given price r_ϕ^{F} (for $\phi \in \Phi_\Theta^{\mathrm{F}}$) or to the marginal supply cost $\mathrm{d}G_\phi/\mathrm{d}k_\phi$ (for $\phi \in \Phi_\Theta^{\mathrm{E}}$). As part of this long-run equilibrium condition, if any input ϕ is used by two or more plant types θ' and θ'', i.e., $\phi \in \Phi_{\theta'} \cap \Phi_{\theta''}$, then its values imputed in the different uses, $\hat{r}_{\theta'\phi}$ and $\hat{r}_{\theta''\phi}$, are required to be equal. (In a short-run equilibrium, the values of the same capital input commodity in different uses may of course differ.) If done by iteration, the search for $p_{\mathrm{SR}}^\star$ corresponds to the inner loop in Figure 3, and the search for $k_\theta^\star$ corresponds to the outer loop in Figure 3.

Since the SI's technology is specified by production sets (rather than profit or cost functions), this approach generally uses, for a characterization of long-run producer optimum, the SRP optimization system (6.1)–(6.3) or its split form, which, with c.r.t.s., consists of (4.2)–(4.3) and (4.6)–(4.7). The split form can be convenient when the SRC programme is readily solved. The cases in which other systems may be equally workable are pointed to at the end of Section 12.

The two stages of calculating the long-run equilibrium are next described in detail. The first stage is to find the short-run equilibrium, given plants with arbitrary capacities $k = (k_\theta)_{\theta\in\Theta}$, and given arbitrary prices r^{E}, which complement the fixed prices r^{F} to a full capital-input price system $r^{\mathrm{EF}} = \left(r^{\mathrm{E}}, r^{\mathrm{F}}\right)$. At this stage, r^{EF} matters only in calculating

[49] The subdifferential ∂G_ϕ is an interval if the left and right derivatives of G_ϕ differ; this can be the case only on a countable subset of $\left(0, \overline{k}_\phi\right)$. Also, $\partial G_\phi(0) = [0, (\mathrm{d}G_\phi/\mathsf{d}k)(0+)]$ and $\partial G_\phi\left(\overline{k}_\phi\right) = \left[(\mathrm{d}G_\phi/\mathsf{d}k)\left(\overline{k}_\phi -\right), +\infty\right)$.

[50] The corresponding input demand, $\check{v}_\theta\left(\hat{y}_\theta\left(p, k_\theta, w_{|\theta}\right), k_\theta, w_{|\theta}\right)$, would similarly have to be equated to input supply, had the supply not been taken to be perfectly elastic (i.e., if the input prices w were not fixed, and had to be determined).

the total short-run income, which is

$$(13.10)\quad \hat{M}_{\mathrm{SR}\,h}\left(p; r^{\mathrm{E}}, r^{\mathrm{F}}; w, \varrho \,|\, k\right) := m_h^{\mathrm{En}} + \sum_{\theta \in \Theta} \varpi_{h\theta}\left(\Pi_{\mathrm{SR}}^{\theta}\left(p, k_\theta, w_{|\theta}\right) - r_{|\theta}^{\mathrm{EF}} \cdot k_\theta\right)$$
$$+ \sum_{\phi \in \Phi_{\Theta}^{\mathrm{E}}} \varsigma_{h\phi}\left(r_\phi^{\mathrm{E}} \sum_{\theta\colon \phi \in \Phi_\theta^{\mathrm{E}}} k_{\theta\phi} - G_\phi\left(\sum_{\theta\colon \phi \in \Phi_\theta^{\mathrm{E}}} k_{\theta\phi}\right)\right) + \varsigma_{h\,\mathrm{IU}} \Pi_{\mathrm{IU}}\left(p, \varrho\right).$$

Comment (on the composition of income in the short and long runs): The exact expression for the short-run income (13.10) can be approximated by simpler ones. The first sum over θ in (13.10) represents pure-profit income from the SI, and the sum over ϕ is the profit income from supplying any equilibrium-priced inputs to the SI. In the long run, these profits are competitively maximized over k_θ and, as a result, the SI's profit is zero.[51] The profit incomes from input supply usually remain positive in the long run, and their sum over ϕ is a term of $\hat{M}_{\mathrm{LR}\,h}$ in (13.9). For the purpose of calculating the long-run equilibrium by the short-run approach, one can therefore replace $\hat{M}_{\mathrm{SR}\,h}$ by the simpler expression $\hat{M}_{\mathrm{LR}\,h}$ in the *short*-run consumer problem (13.14). This would make the short-run consumer demand map identical to the long-run one. (The short-run equilibria so calculated would differ from the exact ones, but not by very much unless the short-run problem's capacities were far from long-run equilibrium.) Also, since the profit from input supply is likely to be relatively small in practice, it may be acceptable to disregard it in calculating consumer demand (thus taking the consumer's income to be $m_h^{\mathrm{En}} + \varsigma_{h\,\mathrm{IU}}\Pi_{\mathrm{IU}}$, instead of $\hat{M}_{\mathrm{SR}\,h}$ or $\hat{M}_{\mathrm{LR}\,h}$).

Given a $k = (k_\theta)_{\theta\in\Theta}$ as well as r^{E}, r^{F} and w, the short-run general equilibrium (SRGE) system to be solved consists of the following conditions on the other variables (viz., prices p paired with quantities y_θ, x_h and z, price ϱ paired with quantity φ_h, quantities v_θ, and amounts of numeraire m_h and n):

$$(13.11)\qquad y_\theta \text{ maximizes SRP, i.e., meets (4.2), for each } \theta$$

$$(13.12)\qquad v_\theta \text{ minimizes SRC, i.e., meets (4.3), for each } \theta$$

$$(13.13)\qquad (p, 1) \in \varrho \widehat{\partial} F\left(z, n\right)$$

$$(13.14)\qquad \left(x_h, \varphi_h, m_h\right) \text{ maximizes } U_h \text{ on } B\left(p, \varrho, \hat{M}_{\mathrm{SR}\,h}\left(p, r^{\mathrm{EF}}, w, \varrho \,|\, k\right)\right)$$

$$(13.15)\qquad \sum_{\theta\in\Theta} y_\theta = z + \sum_{h\in\mathrm{Ho}} x_h \quad \text{and} \quad F\left(z, n\right) = \sum_{h\in\mathrm{Ho}} \varphi_h.$$

The short-run equilibrium system (13.11)–(13.15) can be solved in steps:

(1) We take it to be easiest to start by solving the SRC programme in (4.3) to determine the short-run conditional demand of each plant type θ for its variable inputs. For a technology with conditionally fixed technical coefficients, i.e., for a

[51]Formally, this is because in long-run equilibrium $r_{|\theta}^{\mathrm{EF}} = \hat{r}_\theta$ as per (13.18), and because $\left(p, \hat{r}_\theta, w_{|\theta}\right) \in \mathbb{Y}_\theta^{\circ}$ for each θ by the dual constraint on r_θ. For the same reason, in calculating the long-run equilibrium one can restrict attention, already at the short-run stage, to those r^{EF}'s with $\left(p, r_{|\theta}^{\mathrm{EF}}, w_{|\theta}\right) \in \mathbb{Y}_\theta^{\circ}$ for each θ.

technology of the form (24.1), the conditional input demand $\check{v}_\theta(y_\theta)$ depends only on the plant's output y_θ. In general, it depends also on the fixed inputs k_θ and the variable-input prices $w_{|\theta}$.

(2) Since C^θ_{SR} is now a known function of $\left(y_\theta, k_\theta, w_{|\theta}\right)$—equal to $w_{|\theta} \cdot \check{v}_\theta$ if the SRC programme is feasible, and to $+\infty$ if not—the reduced SRP programme in (4.2) can be solved next; it is an LP if $\check{v}_\theta$ is linear in y_θ.[52] It generally has a multi-valued solution set, $\hat{Y}_\theta\left(p, k_\theta, w_{|\theta}\right)$.

(3) Consumer demands are found as functions $\left(\hat{x}_h, \hat{\varphi}_h\right)$ of $(p, \varrho; M)$, and the known value of $\Pi^\theta_{\mathrm{SR}}\left(p, k_\theta, w_{|\theta}\right)$—viz., $\langle p \mid y_\theta\rangle - C^\theta_{\mathrm{SR}}(y_\theta)$ for any $y_\theta \in \hat{Y}_\theta$—is used to calculate $\hat{M}_{\mathrm{SR}\,h}$ as per (13.10). Factor demands (of the User Industry) are found as functions $(\hat{z}, \hat{n})$ of $(p, \varrho) \in P_+ \times \mathbb{R}_{++}$, from (13.13).[53]

(4) Finally, the system

$$\hat{z}(p, \varrho) + \sum_{h \in \mathrm{Ho}} \hat{x}_h\left(p, \varrho; \hat{M}_{\mathrm{SR}\,h}\left(p; r^{\mathrm{E}}, r^{\mathrm{F}}; w, \varrho \,|\, k\right)\right) \in \sum_{\theta \in \Theta} \hat{Y}_\theta\left(p, k_\theta, w_{|\theta}\right) \tag{13.16}$$

$$\sum_{h \in \mathrm{Ho}} \hat{\varphi}_h\left(p, \varrho; \hat{M}_{\mathrm{SR}\,h}\left(p; r^{\mathrm{E}}, r^{\mathrm{F}}; w, \varrho \,|\, k\right)\right) = F\left(\hat{z}(p, \varrho), \hat{n}(p, \varrho)\right) \tag{13.17}$$

is solved for p and ϱ.

This gives the short-run equilibrium prices, $p^\star_{\mathrm{SR}}$ (for the Supply Industry's differentiated output good) and $\varrho^\star_{\mathrm{SR}}$ (for the Industrial User's product). It also gives, by back substitution, the short-run equilibrium quantities, viz.: (i) the outputs and demands for the differentiated good, with $\sum_\theta y^\star_{\mathrm{SR}\,\theta} = z^\star_{\mathrm{SR}} + \sum_h x^\star_{\mathrm{SR}\,h}$, (ii) the Supply Industry's variable inputs $v^\star_{\mathrm{SR}\,\theta}$, (iii) the User Industry's output $\varphi^\star_{\mathrm{SR}}$ and input $n^\star_{\mathrm{SR}}$, and (iv) consumption of the numeraire $\sum_h m^\star_{\mathrm{SR}\,h}$. Generally, all of these are functions of the short-run equilibrium problem's data k and r^{E} (as well as depending on the fixed prices r^{F} and w).[54]

The second stage is to determine the long-run equilibrium, i.e., the equilibrium capacities and the prices of any equilibrium-priced capital inputs (i.e., those in Φ^{E}_Θ). Optimality of investment k_θ in each technique is achieved by satisfying the rest of the split SRP optimization system, viz., (4.6)–(4.7). For this, the solution set $\hat{R}_\theta\left(p, k_\theta, w_{|\theta}\right)$ of the FIV minimization programme (5.13)–(5.14) with $\mathbb{Y}_\theta$ in place of $\mathbb{Y}$, or the solution $\hat{r}_\theta$ if it is unique, is calculated at $p = p^\star_{\mathrm{SR}}\left(k, r^{\mathrm{EF}}, w\right)$. Actually, $\hat{r}_\theta$ will usually have already been found as the dual solution in the process of solving the SRP programme for y_θ by a duality method, i.e., as a by-product of Step 2 in solving (13.11)–(13.15). Finally, the system of long-run equilibrium conditions

$$\left(r^{\mathrm{E}}_{|\theta}, r^{\mathrm{F}}_{|\theta}\right) \in \hat{R}_\theta\left(p^\star_{\mathrm{SR}}\left(k; r^{\mathrm{EF}}; w\right), k_\theta, w_{|\theta}\right) \text{ i.e., } r^{\mathrm{EF}}_{|\theta} \text{ meets (4.6) for each } \theta \in \Theta \tag{13.18}$$

$$r^{\mathrm{E}}_\phi \in \partial G_\phi\left(\sum_{\theta\colon \phi \in \Phi_\theta} k_{\theta\phi}\right) \quad \text{for each } \phi \in \Phi^{\mathrm{E}}_\Theta \tag{13.19}$$

---

[52]For example, in thermal electricity generation, $\check{v}_\theta(y_\theta) = \int y_\theta(t)\,\mathrm{d}t$ and so (16.1)–(16.3) is an LP.

[53]This step is independent of the preceding derivation of short-run supply.

[54]For simplicity, the short-run equilibrium is assumed to be unique.

is solved for $k = (k_\theta)_{\theta\in\Theta}$ and r^{E} (given r^{F} and w).[55] Any solution $(k^\star, r^\star)$ is a part of a long-run equilibrium—provided that there is no duality gap between the SRP programme and its dual (5.13)–(5.14) for any θ (i.e., if (4.7) or equivalently (13.5) holds). The rest of the long-run equilibrium follows by substituting $k^\star$ and $r^\star$ into the short-run equilibrium solution. In particular, in long-run equilibrium, consumer and factor demands for the differentiated good, its total output and its price system are:

$$\begin{aligned}\sum_h x^\star_{\mathrm{LR},h} &= \sum_h x^\star_{\mathrm{SR},h}\left(k^\star; r^\star, r^{\mathrm{F}}; w\right)\\ z^\star_{\mathrm{LR}} &= z^\star_{\mathrm{SR}}\left(k^\star; r^\star, r^{\mathrm{F}}; w\right)\\ \sum_\theta y^\star_{\mathrm{LR}\,\theta} &= \sum_\theta y^\star_{\mathrm{SR}\,\theta}\left(k^\star; r^\star, r^{\mathrm{F}}; w\right)\\ p^\star_{\mathrm{LR}} &= p^\star_{\mathrm{SR}}\left(k^\star; r^\star, r^{\mathrm{F}}; w\right).\end{aligned}$$

The SRGE system (13.11)–(13.15) together with the long-run conditions (13.18)–(13.19) can be called the *SRP programme-based LRGE system*.

Comments:

(1) The SRGE system simplifies when there is no income effect on the differentiated good (i.e., when $\hat{x}_h$ is independent of M, in the relevant range): the solution $(p^\star_{\mathrm{SR}}, \varrho^\star_{\mathrm{SR}})$ to (13.16)–(13.17) is then independent of r^{EF}, as in Section 2.

(2) A production technique can usually be identified by its set of fixed inputs, i.e., $\Phi_{\theta'} \neq \Phi_{\theta''}$ for $\theta' \neq \theta''$. Under the stronger assumption that different techniques use disjoint sets of fixed inputs, i.e., that

$$\Phi_{\theta'} \cap \Phi_{\theta''} = \emptyset \quad \text{for } \theta' \neq \theta'', \tag{13.20}$$

the SI's total investment in fixed input ϕ is simply $k_{\theta\phi}$ for the one θ such that $\Phi_\theta \ni \phi$. In other words, it is the case-function (of ϕ) defined, piecewise, as equal to the function k_θ on each Φ_θ. Thus it can be identified with $k = (k_\theta)_{\theta\in\Theta}$ itself. So, under (13.20), the total investment can be denoted by $k\colon \Phi_\Theta \to \mathbb{R}$. The investment in technique θ is then the restriction of k to Φ_θ, which is denoted by $k_{|\Phi_\theta}$, abbreviated to $k_{|\theta}$. This is so in our model of the ESI's technology (Section 15).

(3) Assume that: (i) each input-cost function, G_ϕ, is differentiable on $\mathbb{R}_{++} := \mathbb{R}_+ \setminus \{0\}$, (ii) the techniques use disjoint sets of capital inputs, i.e., (13.20) holds, and (iii) a unique shadow price system $\hat{r}_\theta\left(p, k_{|\theta}, w_{|\theta}\right)$ exists at every $k \gg 0$ and p in a subspace of P that is known to contain $p^\star_{\mathrm{SR}}$. (As we show in [28] for a class of problems that includes peak-load pricing with storage, this is so for the space of continuous real-valued functions $\mathcal{C}[0,T]$, as a price subspace of $P = L^1[0,T]$.) If a long-run equilibrium with $k^\star \gg 0$ is sought, then Conditions (13.18)–(13.19) on k reduce to the following equations for k (a strictly positive vector in $\mathbb{R}^{\Phi_\Theta}$):

[55] As a basic check, note that the number of "generalized equations" in this system (each d-dimensional vector inclusion counting as d "equations") is the same as the number of unknowns (viz., $\sum_{\theta\in\Theta} \operatorname{card}\Phi_\theta + \operatorname{card}\Phi^{\mathrm{E}}_\Theta$).

$$\hat{r}_{\theta\phi}\left(p^{\star}_{\mathrm{SR}}\left(k;\left\{\frac{\mathrm{d}G_{\phi}}{\mathrm{d}\mathsf{k}}\left(k_{\phi}\right)\right\}_{\phi\in\Phi^{\mathrm{E}}_{\Theta}},r^{\mathrm{F}};w\right),k_{|\theta},w_{|\theta}\right) = \begin{cases} \frac{\mathrm{d}G_{\phi}}{\mathrm{d}\mathsf{k}}\left(k_{\phi}\right) & \text{if } \phi\in\Phi^{\mathrm{E}}_{\theta} \\ r^{\mathrm{F}}_{\phi} & \text{if } \phi\in\Phi^{\mathrm{F}}_{\theta} \end{cases} \tag{13.21}$$

for each θ and $\phi\in\Phi_{\theta}$.

(4) This investment problem has a partial-equilibrium version in which a *given* p replaces the $p^{\star}_{\mathrm{SR}}$ in the system (13.21), for a particular production technique θ. We study it in [22], and in [27, Section 11] for the case of pumped storage.

(5) All of the SI's inputs have been assumed to be homogeneous goods, but in some cases an input is a differentiated good. If it is also an equilibrium-priced fixed input, then its supply cost G_{ϕ} is a joint-cost function of the commodity bundle $k_{\phi}\in K_{\phi}$. The short-run approach readily accommodates such inputs (the only difference is that ∂G_{ϕ} is not an interval of $\mathbb{R}$, but a convex subset of the price space R_{ϕ} paired with K_{ϕ}). An example is the river flow $e\in L^{\infty}[0,T]$ for hydroelectric generation in Theorem 17.2, but in that case Condition (13.19) imposes no restriction on the water price function ψ because e is fixed (even in the long run).

14. Duality for linear programmes with nonstandard parameters in constraints

Once the production set $\mathbb{Y}$ has been represented as an intersection of half-spaces, each of the profit or cost programmes of Section 3 becomes an LP, i.e., a programme of optimizing a linear function subject to linear inequality or equality constraints. It is a parametric LP, with the fixed quantities k as its primal parameters (Section 5). The fixed quantities need not, of course, be the standard "right-hand side" parameters. But the marginal effects of any nonstandard parameters can be expressed in terms of those of the standard parameters, i.e., in terms of the standard dual solution σ, which consists of the usual Lagrange multipliers for the constraints. This is done in (14.12) below.

To start with, this formula is given for the case of a finite LP, i.e., an LP with finite numbers of decision variables, parameters and constraints. We focus on the SRP programme of a production technique with c.r.t.s. To simplify the notation, we assume that there is no variable input (i.e., $\Xi=\emptyset$). As well as being met literally by some techniques (e.g., the storage techniques of Section 15), the assumption is not at all restrictive because the output bundle y can always be reinterpreted as the bundle of all the variable commodities (i.e., outputs and variable inputs).

For now, $\mathbb{Y}$ is therefore a polyhedral cone in the finite-dimensional space $Y\times K=\mathbb{R}^{T}\times\mathbb{R}^{\Phi}$, where T and Φ are the sets of output and fixed-input commodities. Its polar, $\mathbb{Y}^{\circ}$, is a finitely generated convex cone in the price space $P\times R=\mathbb{R}^{T}\times\mathbb{R}^{\Phi}$. It can be represented as the sum of a linear subspace spanned by a finite set $\mathcal{G}''$ and a line-free convex cone generated by a finite set $\mathcal{G}'$, i.e.,

$$\mathbb{Y}^{\circ}=\operatorname{cone}\operatorname{conv}\mathcal{G}'+\operatorname{span}\mathcal{G}''$$

for some positively independent, finite set $\mathcal{G}'$ and another finite set $\mathcal{G}''$ (which can be chosen to be linearly independent).[56] The generators $\mathcal{G}'$ and the spanning vectors $\mathcal{G}''$ can serve as the rows of partitioned matrices $[A' \, B']$ and $[A'' \, B'']$ that give[57]

$$\mathbb{Y} = \left\{ (y, -k) \in \mathbb{R}^T \times \mathbb{R}^\Phi : A'y - B'k \leq 0 \text{ and } A''y - B''k = 0 \right\}. \tag{14.1}$$

The primal LP (of short-run profit maximization) is: given $(p, k) \in \mathbb{R}^T \times \mathbb{R}^\Phi$,

$$\text{maximize } p \cdot y \text{ over } y \in \mathbb{R}^T \tag{14.2}$$

$$\text{subject to:} \quad A'y \leq B'k \tag{14.3}$$

$$A''y = B''k. \tag{14.4}$$

Its optimal value is $\Pi_{\mathrm{SR}}(p, k)$, abbreviated to $\Pi(p, k)$. As in Section 5, the vector k is called the *intrinsic* primal parameter, and its increment Δk is an *intrinsic perturbation* of (14.2)–(14.4).

The corresponding *standard parametric LP* has primal parameters s' and s'', ranging over $\mathbb{R}^{\mathcal{G}'}$ and $\mathbb{R}^{\mathcal{G}''}$, in place of the $B'k$ and $B''k$ in (14.3)–(14.4). Its optimal value is the *standard primal value*, denoted by $\widetilde{\Pi}(p, s)$, where $s = (s', s'')$. So by definition, for every (p, k),

$$\Pi(p, k) = \widetilde{\Pi}(p, Bk) \quad \text{where } B := \begin{bmatrix} B' \\ B'' \end{bmatrix}. \tag{14.5}$$

The *standard perturbation* consists in relaxing (or tightening) the inequality constraints by adding an arbitrary vector $\Delta s = (\Delta s', \Delta s'') \in \mathbb{R}^{\mathcal{G}'} \times \mathbb{R}^{\mathcal{G}''}$ to the r.h.s. of (14.3)–(14.4), i.e., it uses a separate scalar increment for each constraint. This produces the *standard dual* of (14.2)–(14.4), which is: given the same $(p, k) \in \mathbb{R}^T \times \mathbb{R}^\Phi$,

$$\text{minimize} \quad \sigma^{\mathrm{T}} Bk = \sigma'^{\mathrm{T}} B'k + \sigma''^{\mathrm{T}} B''k \quad \text{over } \sigma = (\sigma', \sigma'') \in \mathbb{R}^{\mathcal{G}'} \times \mathbb{R}^{\mathcal{G}''} \tag{14.6}$$

$$\text{subject to:} \quad \sigma' \geq 0 \tag{14.7}$$

$$p = A^{\mathrm{T}} \sigma := A'^{\mathrm{T}} \sigma' + A''^{\mathrm{T}} \sigma'' \tag{14.8}$$

where $\cdot^{\mathrm{T}}$ denotes transposition. The variable σ is paired with Δs (not Δk)—this is the dual of the standard primal LP, which is parametrized by s. It is only *after* forming the dual that Bk is substituted for s to give $\sigma^{\mathrm{T}} Bk$ in (14.6). The *standard dual value*, denoted by $\widetilde{\widetilde{\Pi}}(p, s)$, is the optimal value of the LP (14.6)–(14.8) with s instead of Bk, i.e.,

[56] Although it follows that $\mathbb{Y}^\circ$ is the convex cone generated by $\mathcal{G}' \cup \mathcal{G}'' \cup (-\mathcal{G}'')$, it is better to keep $\mathcal{G}'$ and $\mathcal{G}''$ separate when it comes to parameterizing the programme (14.2)–(14.4) in the standard way. For this purpose, an equality constraint should *not* be converted to a pair of opposite inequalities. To do so would always complicate the dual solution by making it nonunique and unbounded: a primal equality constraint (say $a \cdot y = 0$) may have a unique multiplier λ, but if it were replaced by a pair of inequalities ($a \cdot y \leq 0$ and $-a \cdot y \leq 0$), then a corresponding multiplier pair would be any $(\sigma_1, \sigma_2) \geq 0$ with $\sigma_1 - \sigma_2 = \lambda$, i.e., any point of a half-line. Its unboundedness expresses the fact that the programme would become immediately infeasible if one inequality constraint of the pair were tightened without relaxing the other by the same amount (i.e., if the constraints were perturbed to $a \cdot y \leq s_1$ and 0 and $-a \cdot y \leq s_2$ for $s_1 < -s_2$).

[57] Formally, A and B are the $\mathcal{G} \times T$ and $\mathcal{G} \times \Phi$ matrices with entries $A_{gt} = g_t$ and $B_{g\phi} = g_\phi$ for $t \in T$, $\phi \in \Phi$ and $g \in \mathcal{G} = \mathcal{G}' \cup \mathcal{G}''$.

before the substitution. Its solution, the *standard dual solution*, is denoted by $\hat{\sigma}(p,s)$ when it is unique; in general, the solutions form a set $\hat{\Sigma}(p,s)$. The solution set of (14.6)–(14.8) is therefore $\hat{\Sigma}(p,Bk)$; when unique, the solution is $\hat{\sigma}(p,Bk)$. Its value is $\overline{\overline{\Pi}}(p,Bk)$. This is always equal to the fixed-input value as calculated from (5.13)–(5.14), i.e.,[58]

$$\overline{\overline{\Pi}}(p,Bk) = \overline{\Pi}(p,k) \quad \text{for every } (p,k)\,. \tag{14.9}$$

In other words, the standard dual LP has the same value as the intrinsic dual; here, the two duals are (14.6)–(14.8) and (5.13)–(5.14). For their solution sets, $\hat{\Sigma}$ and $\hat{R}$, it follows that

$$\begin{aligned}\hat{R}(p,k) = \widehat{\partial}_k\overline{\Pi}(p,k) = B^{\mathrm{T}}\,\widehat{\partial}_s\overline{\overline{\Pi}}(p,s)\Big|_{s=Bk} &= B^{\mathrm{T}}\hat{\Sigma}(p,Bk) \\ &:= \left\{B^{\mathrm{T}}\sigma : \sigma \in \hat{\Sigma}(p,Bk)\right\}\end{aligned} \tag{14.10}$$

by applying the Chain Rule to (14.9),[59] and by using (twice) the identity of the dual solution and the marginal value of Type Two.[60] Thus the intrinsic dual solution ($\hat{R}$) is expressed as the linear image of the standard dual solution ($\hat{\Sigma}$) under the adjoint ($B^{\mathrm{T}}\cdot$) of the operation that maps the intrinsic to the standard primal parameters ($s = Bk$).

When $\Pi = \overline{\Pi}$ at (p,k), the marginal value is actually of Type One by Remark 19.8, i.e.,

$$\widehat{\partial}_k\Pi(p,k) = \widehat{\partial}_k\overline{\Pi}(p,k) = B^{\mathrm{T}}\hat{\Sigma}(p,Bk)\,. \tag{14.11}$$

This always applies to *finite* LPs because their primal and dual values are equal, unless both programmes are infeasible (in which case their values are oppositely infinite).[61] If additionally the dual solution is unique, then

$$\nabla_k\Pi(p,k) = B^{\mathrm{T}}\hat{\sigma}(p,Bk)\,. \tag{14.12}$$

This gives the marginal values of the generally nonstandard intrinsic parameters (k) in terms of the *standard* dual solution ($\hat{\sigma}$).

Comment (on standard and intrinsic perturbations): If B were the unit matrix I, the two perturbation schemes would obviously be the same (and Δs could be renamed to Δk). This *would* be so if the short-run production constraints corresponded, one-to-one, to the fixed inputs, i.e., if $\mathbb{Y}$ were defined by a system of inequalities (or equalities) of

[58]The identity (14.9) reduces to (14.5) when the primal and dual values are equal, i.e., when $\widetilde{\Pi} = \overline{\overline{\Pi}}$ and $\Pi = \overline{\Pi}$ at (p,k). This always applies to (feasible) finite LPs, but not always to infinite LPs. To prove (14.9) without relying on the absence of a duality gap, note that the constraint $(p,r) \in \mathbb{Y}^{\circ}$ in (5.14) means here that $A^{\mathrm{T}}\sigma = p$ and $B^{\mathrm{T}}\sigma = r$ for some $\sigma = (\sigma', \sigma'')$ with $\sigma' \geq 0$ (since the rows of $[A\,B]$ generate or span $\mathbb{Y}^{\circ}$). So the change of variable from r to $B^{\mathrm{T}}\sigma$ transforms (5.13)–(5.14) into (14.6)–(14.8). This argument extends to infinite LPs (and it applies also when there is a duality gap).

[59]For the Chain Rule for subdifferentials, see, e.g., [4, 4.3.6 a], [32, 4.2: Theorem 2], [42, 23.9] or [44, Theorem 19].

[60]First noted at the end of Section 7, the identity is detailed in Section 19 (Lemma 19.2).

[61]See, e.g., [11, 5.1 and 9.1] or [44, Example 1', p. 24] for proofs based on the simplex algorithm or on polyhedral convexity, respectively. This is *not* so with a pair of *infinite* LPs: both can be feasible without having the same value (i.e., the primal and dual values can both be finite but different). Appendix A gives an example.

the form $(Ay)_\phi \leq k_\phi$, one for each $\phi \in \Phi$. But such a correspondence generally fails to exist, for three reasons. First, two fixed inputs may appear in one constraint (say $a \cdot y \leq k_1 + k_2$). Second, a constraint may involve only the outputs ($a \cdot y \leq 0$, e.g., $y_t \geq 0$). Third, each fixed quantity k_ϕ may impose more than one constraint on y (say $(Ay)_1 \leq k_\phi$, $(Ay)_2 \leq k_\phi$, ...). Indeed, this is so whenever k_ϕ is a capacity: staying constant over a time period, it is a scalar but it imposes as many inequality constraints as there are time instants (e.g., $y_t \leq k_\phi$ for each t).[62] In such a case, B is a 0-1 matrix whose unit entries appear just once in a row, but more than once in a column. When additionally k is a scalar, B is the single column $\begin{bmatrix} 1 & \ldots & 1 \end{bmatrix}^{\mathrm{T}}$; and an intrinsic perturbation of the constraint system $Ay \leq \begin{bmatrix} k & k & \ldots \end{bmatrix}^{\mathrm{T}}$ relaxes all the constraints by the same amount, to $Ay \leq \begin{bmatrix} k + \Delta k & k + \Delta k & \ldots \end{bmatrix}^{\mathrm{T}}$. By contrast, a standard perturbation relaxes each constraint by a different amount, to $Ay \leq \begin{bmatrix} k + \Delta s_1 & k + \Delta s_2 & \ldots \end{bmatrix}^{\mathrm{T}}$. In this sense, the standard perturbation scheme is the finest; and, with this B, the intrinsic perturbation scheme is the coarsest. Once the scalar k is identified with the vector $(k, k, \ldots)$, the standard value function $\widetilde{\Pi}(p, \cdot)$ becomes an extension of the intrinsic value function $\Pi(p, \cdot)$ from the subspace of constant tuples to all of $\mathbb{R}^{\mathcal{G}' \cup \mathcal{G}''}$ (with $\mathcal{G}''$ empty if there is no equality constraint), and the intrinsic dual solution (a scalar) is simply the total sum of the standard dual solution, i.e., $\hat{r} = \begin{bmatrix} 1 & \ldots & 1 \end{bmatrix} \hat{\sigma} = \hat{\sigma}_1 + \hat{\sigma}_2 + \ldots$. In other words, the scalar parameter's marginal value is the sum of the marginal values of relaxing all the constraints in which it appears. This arises in the peak-load pricing application to give the total capacity values as the integrals of the rent flows over the period, in (16.10), (16.23)–(16.24), and (16.44)–(16.45). Also, since $\widetilde{\Pi}$ is an extension of Π, it can be convenient to use the same letter k as the second variable of both functions (i.e., to use k instead of the s in $\widetilde{\Pi}(p, s)$), provided that it is always made clear whether k is a scalar or a vector. We do so in the context of hydro and energy storage (where s signifies the water or energy stock and is not a parameter).

Finally, the standard dual can be reformulated by including the intrinsic dual variable r, which is paired with k and constrained to equal $B^{\mathrm{T}}\sigma$; thus r is wholly dependent on σ. The objective, $\sigma^{\mathrm{T}} B k$, may then be rewritten concisely as $r \cdot k$. This produces the following LP: given $(p, k) \in \mathbb{R}^T \times \mathbb{R}^\Phi$,

$$\text{minimize } r \cdot k \text{ over } r \in \mathbb{R}^\Phi \text{ and } \sigma = (\sigma', \sigma'') \in \mathbb{R}^{\mathcal{G}'} \times \mathbb{R}^{\mathcal{G}''} \tag{14.13}$$

$$\text{subject to: } \sigma' \geq 0,\ p = A^{\mathrm{T}}\sigma \text{ and } r = B^{\mathrm{T}}\sigma. \tag{14.14}$$

This may be called the *inclusive* standard dual—an LP for both r and σ. It is the dual that derives from *simultaneous* standard and intrinsic perturbations, i.e., from perturbing Bk on the r.h.s. of (14.2)–(14.4) to $\Delta s + B(k + \Delta k)$. Its solution gives both sets of marginal values explicitly ($\hat{\sigma}$ and $\hat{r}$), but it is in substance equivalent to the standard dual solution $\hat{\sigma}$. It can be more convenient to use a *partly inclusive* form of the standard dual,

[62]Also, a nonnegativity constraint on k_ϕ makes it appear a second time even if it imposes just one constraint on y (i.e., $0 \leq k_\phi$ in addition to $a \cdot y \leq k_\phi$ for some $a \neq 0$).

which includes only some of the intrinsic dual variables, leaving out those coordinates of r which correspond to "the simplest" columns of B—e.g., to the columns with 0-1 entries as in the Comment above. For example, the programme of valuing the hydro inputs (16.37)–(16.43) includes the TOU shadow price of water ψ but not the total capacity values r_{St} and r_{Tu}, which are simply the totals of the standard dual variables κ_{St} and κ_{Tu}.

Expressing general dual variables (r) in terms of the standard ones (σ) can be extended to infinite LPs. This requires using suitable cones in infinite-dimensional spaces of variables and parameters to formulate infinite systems of constraints on, generally, an infinity of variables. Such a framework is provided in, e.g., [12, 4.2], [36, 7.9] and [44, Examples 4, 4', 4"]. The assumptions we make here to adapt it are not the weakest possible; they are selected for simplicity and adequacy to our applications (Section 16). The output and fixed-input spaces, Y and K, are now taken to be general Banach spaces, i.e., complete normed spaces (instead of $\mathbb{R}^T$ and $\mathbb{R}^\Phi$). The norm-duals, Y^* and K^*, serve as the corresponding price spaces, P and R. For the primal programme of SRP maximization, Y is the primal-variable space paired with the dual parameter space P, and K is the primal-parameter space paired with the dual-variable space R. The production cone is given by (14.1) in terms of two norm-to-norm continuous linear operations: (i) $A'\colon Y \to L$ and $B'\colon K \to L$, whose common codomain L is a Banach lattice (with a vector order $\leq$ and the corresponding nonnegative cone L_+), and (ii) $A''\colon Y \to X$ and $B''\colon K \to X$, whose codomain X is a Banach space. The spaces L and X replace $\mathbb{R}^{\mathcal{G}'}$ and $\mathbb{R}^{\mathcal{G}''}$ as the spaces for standard perturbations ($\Delta s'$, $\Delta s''$). Their norm-duals, L^* and X^*, serve as the spaces for standard dual variables (σ', σ''). It is best to keep L and X small, but obviously L must contain the ranges of both A' on Y and B' on K (and similarly X must contain both $A''Y$ and $B''K$).

As for the choice of topologies, this must be consistent with the pairing of spaces. Furthermore, the norm topology has to be put on the primal parameter space L if the generalized Slater's Condition of [44, (8.12)] is to be met for the SRP programme (14.2)–(14.4), i.e., if a y is to exist such that $A'y - B'k \in -\operatorname{int}(L_+)$ and $A''y - B''k = 0_X$. Topologies on Y, K, L and X must make the maximand u.s.c. and the constraint relations closed; here, this means making $\langle p \mid \cdot \rangle$, A and B continuous. So the norm topologies on Y (the primal-variable space) and on K, L and X (the primal-parameter spaces) will do. On the dual-variable spaces K^*, L^* and X^*, the weak* topologies will do.[63] On Y^* (the dual parameter space), the Mackey topology $\mathrm{m}(Y^*, Y)$ is the best choice if continuity of the dual value function is sought. When Y has a Banach predual Y', it can also be useful to pair Y with Y' as a dual parameter space that is generally smaller than Y^*; the restriction of $\mathrm{m}(Y^*, Y)$ to Y' is the norm topology of Y'. The pairing of Y with Y' is adequate when $p \in Y'$, but *not* when $p \in Y^* \setminus Y'$.

There are at least two sources for the linear operations A and B that describe $\mathbb{Y}$ by (14.1). First, such a formula may be the original definition of $\mathbb{Y}$—in which case A and

[63] The weak topologies do not enter the analysis explicitly, but they make the adjoint operators continuous: see, e.g., [18, 16C].

B can simply be read off. This is so in our application to the ESI: the production sets (15.1), (15.4) or (15.9) are all of the form (14.1).[64]

Second, A' and B' (with no A'' or B'', i.e., with the zero space as X) can also be constructed from a weakly* compact convex base, Δ, for $\mathbb{Y}^\circ$, which exists if and only if $\mathbb{Y}$ is solid (i.e., has a nonempty interior) for the norm on $Y \times K$: see, e.g., [3, Theorem 3.16]. An interior point $\left(y^{\mathrm{S}}, -k^{\mathrm{S}}\right)$ defines the base

$$\Delta := \left\{(p,r) \in \mathbb{Y}^\circ : \left\langle p \,|\, y^{\mathrm{S}} \right\rangle - \left\langle r \,|\, k^{\mathrm{S}} \right\rangle = -1 \right\}. \tag{14.15}$$

Such a Δ can serve as a replacement for the finite set $\mathcal{G}'$ that generates $\mathbb{Y}^\circ$ when $\mathbb{Y}$ is a solid polyhedral cone in a finite-dimensional space. The Banach lattice of all weakly* continuous functions on Δ, denoted by $\mathcal{C}(\Delta)$, replaces $\mathbb{R}^{\mathcal{G}'}$ and serves as the codomain (L) for the operations A' and B'. These are specified by[65]

$$\left(A'y - B'k\right)(p,r) := \langle p \,|\, y \rangle - \langle r \,|\, k \rangle \quad \text{for } (p,r) \in \Delta. \tag{14.16}$$

So $\mathcal{C}(\Delta)$ is the space of standard perturbations, and the space of standard dual variables (the constraints' multipliers) is the space of all finite Borel measures $\mathcal{M}(\Delta) = \mathcal{C}^*(\Delta)$ by Riesz's Representation Theorem. Some points of Δ are convex combinations of others. This redundancy can be lessened by replacing Δ with any *closed*, and hence compact, subset $\mathcal{G}'$ such that $\operatorname{cl}\operatorname{conv}\mathcal{G}' = \Delta$. When the set of extreme points $\operatorname{ext}\Delta$ is closed, it is the best choice of $\mathcal{G}'$ (and all the redundancy is thus removed). But generally $\operatorname{ext}\Delta$ need not be closed, even if Δ is finite-dimensional.

Comments (on the construction of (A', B') from a base Δ for $\mathbb{Y}^\circ$):

(1) When $\mathbb{Y}$ is a solid polyhedral cone in a *finite*-dimensional space $\mathbb{R}^T \times \mathbb{R}^\Phi$, the operations A' and B' constructed from a base Δ for $\mathbb{Y}^\circ$ are at least as good as the A' and B' read off from any original formula for $\mathbb{Y}$. This is because $\operatorname{ext}\Delta$ is then a finite set generating $\mathbb{Y}^\circ$, and when its elements are put together as rows of a matrix $[A'\, B']$, it gives the simplest representation of $\mathbb{Y}$ in the form (14.1)—with the matrix $[A''\, B'']$ empty because $\mathbb{Y}$ is solid.

(2) But in the infinite case the original A' and B' can be simpler than those constructed from Δ, though the two can also turn out to be exactly the same. This can depend on the details of space specifications. For example, consider

$$\mathbb{Y} := \{(y, -k) \in Y \times \mathbb{R} : y \le k\} \tag{14.17}$$

[64]The output space is $Y = L^\infty[0,T]$, which has a predual $Y' = L^1[0,T]$. The fixed-input space K depends on the technique: it is either $\mathbb{R}$ for a thermal technique, or $\mathbb{R}^2$ for pumped storage, or $\mathbb{R}^2 \times L^\infty[0,T]$ for hydro. As for L (the space of standard perturbations of inequality constraints), it is either $L^\infty[0,T]$ or its Cartesian product with $\mathcal{C}[0,T]$ when, in the case of an energy storage technique, there are reservoir constraints in addition to generation constraints. And the balance constraint of a storage techniques has $\mathbb{R}$ as X (the space of standard perturbations of the equality constraint).

[65]Formula (14.16) adapts [12, p. 154, line 11 f.b.], where the construction is mistakenly proposed as a possible way of dealing with a non-solid cone (in such a case the polar cannot have a compact base, so the analysis does not apply). The construction can, however, be extended to the case that $\mathbb{Y}$ is relatively solid, i.e., has a nonempty interior in the linear subspace $\mathbb{Y} - \mathbb{Y}$ (assumed to be closed in $Y \times K$); the polar $\mathbb{Y}^\circ$ is then the sum of the annihilator $(\mathbb{Y} - \mathbb{Y})^\perp$ and a cone with a compact base Δ.

with either $\mathcal{C}[0,T]$ or $L^\infty[0,T]$ as the output space Y. (This is a simpler version of the technology (15.1), stripped of the variable input and without the nonnegativity constraint.)

(a) For now, $Y = \mathcal{C}[0,T]$. The original operations defining this $\mathbb{Y}$ by means of (14.1) are: the identity map $A'y = y$ for $y \in \mathcal{C}[0,T]$, and the embedding of $\mathbb{R}$ in $\mathcal{C}[0,T]$ by mapping scalars to constant functions, i.e., $B'k = k_{[0,T]} \in \mathcal{C}$ for $k \in \mathbb{R}$ (with no A'' or B'' because $\mathbb{Y}$ is solid). The interior point $\left(0_{[0,T]}, -1\right) \in \mathbb{Y}$ defines, by (14.15), the compact base

$$\Delta = \{p \in \mathcal{M}[0,T] : p \geq 0,\ p[0,T] = 1\} \times \{1\}. \tag{14.18}$$

Its set of extreme points is

$$\begin{aligned} \operatorname{ext}\Delta &= \operatorname{ext}\{p \in \mathcal{M}_+[0,T] : p[0,T] = 1\} \times \{1\} \\ &= \{\varepsilon_t : t \in [0,T]\} \times \{1\} \simeq [0,T] \end{aligned}$$

where ε_t is the Dirac measure at t (i.e., a unit mass concentrated at the single point t). Each ε_t is identified with t itself; and with Δ replaced by $\operatorname{ext}\Delta \simeq [0,T]$, Formula (14.16) reproduces the original operations A' and B' *exactly* by giving

$$(A'y - B'k)(t) \simeq (A'y - B'k)(\varepsilon_t, 1) := \langle \varepsilon_t \,|\, y \rangle - k = y(t) - k \quad \text{for } t \in [0,T].$$

(b) This is not quite so once the space Y in (14.17) is enlarged from $\mathcal{C}[0,T]$ to $L^\infty[0,T]$, the space of all essentially bounded functions. Although Δ is still the nonnegative part of the unit sphere, the sphere is now that of $L^{\infty *}$ rather than of $\mathcal{M}$ as in (14.18). In either case, its extreme points can be characterized as scalar-valued lattice-homomorphisms on L^∞ or $\mathcal{C}$ (into $\mathbb{R}$) of unit norm, and also as nonzero multiplicative linear functionals (i.e., scalar-valued algebra-homomorphisms) on L^∞ or $\mathcal{C}$: see, e.g., [2, 12.27] and [46, 11.32], respectively. More precisely, $\operatorname{ext}\Delta = H \times \{1\}$, where H is the set of all such homomorphisms on either L^∞ or $\mathcal{C}$. But the homomorphisms on $L^\infty[0,T]$ are not as simple as those on $\mathcal{C}[0,T]$, which are Dirac measures and thus correspond to points of $[0,T]$. In the case of L^∞, H is an extremally disconnected weakly* compact subset of $L^{\infty *}$, and $\mathcal{C}(H)$ is isomorphic (both as a normed lattice and as a normed vector algebra) to $L^\infty[0,T]$. In other words, the construction amounts to representing equivalence classes of bounded measurable functions on $[0,T]$ as continuous functions on another, much more complicated, compact set. The "*almost* everywhere" inequality constraint, $y(t) \leq k$ for a.e. t, is thus replaced by an infinite system of scalar inequalities, viz., $\langle p \,|\, y \rangle \leq k$ for *every* $p \in H$. But, since the indexing set H is far from simple, such a reformulation may not be worthwhile.

15. Technologies for electricity generation and energy storage

The rudimentary peak-load pricing example of Section 2 is next developed into a continuous-time equilibrium model of electricity pricing. This requires a fuller description of the industry's technology to start with. A typical electricity supply industry uses a combination of thermal generation, hydro, pumped energy storage, and other techniques.

A thermal plant can be classified by fuel type as, e.g., nuclear, coal-, oil- or gas-fired. A hydro plant can be classified by head height as high-, medium-, or low-head. A pumped-storage plant can be classified by its medium for energy storage as, e.g., a pumped-water or compressed-air plant (PWES or CAES plant), a superconducting magnetic coil (SMES plant) or a battery. Each type can be further subdivided by the relevant design characteristics, which all affect the plant's unit input costs as well as its technical performance parameters (such as response time and efficiency of energy conversion). But the structure of feasible input-output bundles is nearly the same for all the techniques within each of the three main types (thermal, hydro and pumped storage). To simplify these technology structures, we ignore some of the cost complexities and technical imperfections:

(1) A thermal plant is assumed to have a constant technical efficiency η, i.e., a constant heat rate (both incremental and average) of $1/\eta$.[66] So the plant has a constant unit running cost w (in \$/kWh, say) over the entire load range from zero to the plant's capacity.[67]
(2) A hydro plant is assumed to have a constant head, and a turbine-generator of a constant technical efficiency.[68]
(3) In a pumped-storage plant, the energy converter is taken to be perfectly efficient and symmetrically reversible (i.e., capable of converting both ways, and at the same rate).[69]
(4) All plant types are assumed to have no startup or shutdown costs or delays.[70]
(5) Like operation, investment is assumed to be divisible.

Some of these conditions—viz., perfect conversion in pumped storage and constant head in hydro—are imposed purely to simplify this presentation, and can be removed by using the results of [21] and [30]. As for indivisibility, it does not loom large in large-scale systems (nor does the sunk operating cost of a thermal plant, i.e., the no-load fuel cost

[66]A steam plant's efficiency is the product of the boiler's and turbine-generator's efficiencies, which is about $0.85 \times 0.45 \approx 38\%$ (i.e., the heat rate is about $1/0.38 \times 3600$ kJ/kWh≈ 9500 kJ/kWh).

[67]In reality, the minimum operating load is 10% to 25% of the maximum, and the incremental rate rises with load by up to 5% to 15%. Also, there is a no-load heat input (which is a sunk operating cost per unit time of being on line). See, e.g., [38, Figures 8.2 and 8.3, and Table 8.3].

[68]In reality, a turbine's efficiency varies with the load (from about 85% to 95% for movable-blade types, or 70% to 95% for fixed-blade types). Also, a plant's head varies with the water stock. The variation tends to be larger in lower-head plants, but it much depends on the particular plant: e.g., with a typical medium head (say about 150 m), the variation is 3% of the maximum in some plants, but over 30% in others. For a variable-head plant, we study the operation and valuation problems in [30].

[69]In reality, the round-trip conversion efficiency η_{Ro} is close to 1 in SMES (over 95%). In PWES and CAES, η_{Ro} is around 70% to 75% (i.e., 0.7 kWh of electricity is recovered from every kWh used up). The case of $\eta_{\mathrm{Ro}} < 1$ is included in our model of pumped storage [21], as are the cases of converter asymmetry or nonreversibility (although reversibility is usual, some high-head PWES plants do use nonreversible multi-stage pumps).

[70]In reality, startup times range from nearly zero for some energy storage plants (SMES coils and batteries can switch from charging to discharging in 4 to 20 miliseconds), through a few minutes (1–10 min) for other storage plants (PWES or CAES) as well as gas turbines and hydro plants, to hours for nuclear or fossil (coal, oil, gas) steam-plants (whose long startup times must of course be distinguished from the very much shorter loading times applicable to the spinning reserves): see, e.g., [38, Table 8.2] and [40] or [10].

of its being on line). Also, the model can be extended to include transmission costs and constraints.

The one restriction that *cannot* be relaxed without changing some of the model's mathematical foundations is the assumption of immediate startup at no cost. This condition means that the thermal operating cost is additively separable over time; it also means that both short-run and long-run thermal generation costs are symmetric (a.k.a. rearrangement-invariant) functions of the output trajectory over the cycle. These properties are fundamental to the integral formulae for the short-run and long-run thermal costs,[71] and hence also to our method of calculating the long-run marginal cost of thermal generation [19]. The symmetry property, and its weaker variants for other techniques, underlies also our time-continuity result for the equilibrium price function [28]. And price continuity is what guarantees that the two capacities of a pumped-storage plant (viz., the reservoir and the energy converter) have well-defined and separate profit-imputed marginal values, despite their "perfect complementarity": see [21] or [27]. In the case of a hydro plant, it also guarantees that the river flows have well-defined marginal values (as do the reservoir and turbine capacities): see [24].

But the assumption of no startup costs can be rather less distorting than it may seem. This is because the slow-starting plants tend to have low unit running costs, and the quick-starting plants tend to have high unit running costs. To minimize the operating cost, one allocates the base load to the lowest-cost plants, and the near-peak loads to the highest-cost plants. Thus the slowest starters end up serving mainly the constant load levels (the base load), and the quickest starters end up serving the most intermittent load levels (the near-peaks)—even if the differences in startup times are disregarded in the despatch policy.

The complete generating technology consists, then, of the various thermal, hydro and pumped-storage techniques, which form three sets: Θ_{Th}, Θ_{H} and Θ_{PS}. However, what we consider here is a smaller model with a number of thermal techniques and just one other, viz., either a pumped-storage technique or a hydro technique. So the single non-thermal technique can be denoted simply by PS or H, and the set of thermal techniques by $\{1, 2, \ldots, \Theta\}$, where Θ means the number of thermal techniques. In other words, the ESI's set of techniques is henceforth either $\{1, 2, \ldots, \Theta; \mathrm{PS}\}$ or $\{1, 2, \ldots, \Theta; \mathrm{H}\}$. It plays the role of the abstract set Θ of Sections 12 and 13.

The output space Y is here $L^\infty [0, T]$, which is the vector space of all essentially bounded real-valued functions on the interval $[0, T]$ that represents the cycle. Functions equal almost everywhere, w.r.t. the Lebesgue measure (meas), are identified with one another. With the usual order $\leq$ and the supremum norm

$$\|y\|_\infty := \operatorname{EssSup} |y| = \operatorname*{ess\,sup}_{t \in [0,T]} |y(t)|$$

L^∞ is a dual Banach lattice.[72] Its Banach predual is $L^1 [0, T]$, the space of all integrable functions. When, as here, it serves as the price space P, a TOU electricity price is a density function, i.e., a time-dependent rate $p(t)$ in \$/kWh. The price space $L^1 [0, T]$ is

[71]For a one-station technology, the thermal SRC and LRC are given by (2.5) and (2.6). The formulae are extended to a multi-station technology in, e.g., [23, (22)–(26)] and [24].

[72]For Banach-lattice theory, see, e.g., [2, Chapter 4], [8, XV.12] and [33, Chapter X].

sufficient in the case of interruptible demand because capacity charges are then spread out over a flattened peak: see [26]. A larger price space is needed to accommodate the instantaneous capacity charge that arises in the case of a firm, pointed peak.[73]

A *thermal* technique, θ, generates an output flow $y \in L_+^\infty [0,T]$ from two input quantities: k_θ (in kW) of generating capacity of type θ, and v_θ (in kWh) of fuel of the appropriate kind, $\widetilde{\xi}_\theta$. Its long-run production set is the convex cone

$$\mathbb{Y}_\theta := \left\{ (y; -k_\theta, -v_\theta) : y \leq k_\theta,\ \frac{1}{\eta_\theta} \int_0^T y(t)\, \mathrm{d}t \leq v_\theta,\ y \geq 0 \right\} \tag{15.1}$$

where the constant η_θ is the efficiency of energy conversion (the ratio of electricity output to heat input). Viewed as a subset of $L_+^\infty \times \mathbb{R}_-^2$, the set $\mathbb{Y}_\theta$ looks independent of θ (except for the coefficient η_θ), i.e., all thermal techniques have the same structure. But each uses its own input commodities: in terms of (13.1), $\Phi_\theta = \{\theta\}$, $\Xi_\theta = \left\{\widetilde{\xi}_\theta\right\}$, and $\mathbb{Y}_\theta$ is formally a subset of $L_+^\infty [0,T] \times \mathbb{R}_-^{\{\theta, \widetilde{\xi}_\theta\}}$, a space that depends on θ.[74]

The unit fuel cost $\widetilde{w}_\theta$ (in \$ per kWh of electricity output) is, for each plant type θ, its heat rate $1/\eta_\theta$ times its fuel's price $w_{\widetilde{\xi}_\theta}$ (in \$ per kWh of heat input). To simplify the notation, we assume that different types of plants use different fuels (i.e., that $\widetilde{\xi}_{\theta'} \neq \widetilde{\xi}_{\theta''}$ for $\theta' \neq \theta''$): fuel of kind $\widetilde{\xi}_\theta$ can then be unambiguously measured in kWh of electricity generated by plant type θ (instead of being measured as the heat input). Such measurement redefines the plant's efficiency η_θ as 1, and so it equates the plant's unit fuel cost $\widetilde{w}_\theta$ to its fuel's price $w_{\widetilde{\xi}_\theta}$, which can be abbreviated to w_θ. In the case of different types of plant using the same kind of fuel, the w_θ (in this and the next two sections), must be replaced by $\widetilde{w}_\theta = w_{\widetilde{\xi}_\theta}/\eta_\theta$, but no other change is needed.[75]

Henceforth, w_θ is actually taken to represent all of the unit running cost (a.k.a. operating or variable cost).[76] Also, the thermal techniques are numbered in the order of

[73] An instantaneous charge can be represented by a point measure; in the context of electricity pricing, this is a capacity charge in \$ per kW of *power* taken at the peak instant, and it is additional to the marginal fuel charge, which is a price density in \$ per kWh of *energy* at any time. A general singular measure can be interpreted as a concentrated charge. As we point out in [26, Sections 1 and 2], the Banach dual $L^{\infty *}$ can be useful in arriving at such a price representation when the equilibrium allocation lies actually in the space of continuous functions $\mathcal{C}[0,T] \subset L^\infty[0,T]$. This is because the restriction, to $\mathcal{C}$, of a linear functional $p \in L^{\infty *}$ has the Riesz representation by a (countably additive) measure $p_\mathcal{C} \in \mathcal{M} = \mathcal{C}^*$, which can have a singular part as well as a density part. The failure of $L^{\infty *}$ itself to have a tractable mathematical form is thus side-stepped without restricting the analysis to the case of price densities. (The alternative of working entirely within $\mathcal{C}$ and $\mathcal{M}$ as the commodity and price spaces is suitable when *all* demand is uninterruptible [20]. When all demand is harmlessly interruptible, the equilibrium price is a density [26].)

[74] As in Section 13, each $\mathbb{Y}_\theta$ is embedded in the full commodity space as $\mathbb{Y}_\theta \times \{(0,0,\ldots)\}$, by inserting zeros in the input-output bundle at all the positions other than θ, $\widetilde{\xi}_\theta$ and the t's.

[75] This is because our assumption of fixed fuel prices is equivalent to that of fixed unit fuel costs. In any analysis with variable fuel prices, the only implication of the same fuel being used by multiple plant types would be that their unit fuel costs could change only in a fixed proportion.

[76] The other components of unit running cost (extra maintenance, etc.) can be accounted for by a levy on fuel (i.e., by increasing the original w_θ).

increasing unit operating cost, i.e.,

$$w_1 \leq w_2 \leq \ldots \leq w_\Theta.$$

Known as the *merit order*, it is the main conceptual reason for including several thermal techniques in the model. By contrast, inclusion of several storage or hydro techniques would not add new features to the analysis.

Comment: Thermal generation is a technique with conditionally fixed coefficients, i.e., its conditional input demands depend on the output bundle y, but not on the input prices. Formally, $\mathbb{Y}_\theta$ is a case of (24.1) with $Y_0 = L_+^\infty [0,T]$ and with

$$\check{k}_\theta (y) = \mathrm{EssSup}(y) := \operatorname{ess\,sup}_{t\in[0,T]} y(t) \tag{15.2}$$

$$\check{v}_\theta (y) = \int_0^T y(t)\, \mathrm{d}t \tag{15.3}$$

which are the capacity and fuel requirement functions.

Pumped storage produces a signed output flow $y \in L^\infty [0,T]$ from the inputs of storage capacity k_{St} (in kWh) and conversion capacity k_{Co} (in kW). Energy is moved in and out of the reservoir with a converter, which is taken to be perfectly efficient and symmetrically reversible: this means that, in unit time, a unit converter can either turn a unit of electricity into a unit of the storable energy, or *vice versa*. So the output from storage, $y = y^+ - y^-$, equals the rate of energy flow of from the reservoir, $-\dot{s} = -\mathrm{d}s/\mathrm{d}t$ (where $s(t)$ is the energy stock at time t). Energy can be held in storage at no running cost (or loss of stock). The long-run production set is, therefore, the convex cone

$$\mathbb{Y}_{\mathrm{PS}} := \Big\{ (y; -k_{\mathrm{St}}, -k_{\mathrm{Co}}) \in L^\infty [0,T] \times \mathbb{R}_-^2 : |y| \leq k_{\mathrm{Co}},\ \int_0^T y(t)\, \mathrm{d}t = 0 \\ \text{and } \exists s_0 \in \mathbb{R}\ \forall t \in [0,T]\ \ 0 \leq s_0 - \int_0^t y(\tau)\, \mathrm{d}\tau \leq k_{\mathrm{St}} \Big\}. \tag{15.4}$$

Comment: This is also a technique with conditionally fixed coefficients, which means that $\mathbb{Y}_{\mathrm{PS}}$ has the form (24.1). In this case

$$Y_0 = L_0^\infty [0,T] := \left\{ y \in L^\infty : \int_0^T y(t)\, \mathrm{d}t = 0 \right\} \tag{15.5}$$

and the requirements for storage capacity and conversion capacity, when the (signed) output from storage is $y \in L_0^\infty$, are:

$$\check{k}_{\mathrm{St}} (y) = \max_{t\in[0,T]} \int_0^t y(t)\, \mathrm{d}t + \max_{t\in[0,T]} \int_t^T y(t)\, \mathrm{d}t \tag{15.6}$$

$$\check{k}_{\mathrm{Co}} (y) = \|y\|_\infty = \operatorname{ess\,sup}_{t\in[0,T]} |y(t)|. \tag{15.7}$$

In these terms, $(y, -k_{\mathrm{St}}, -k_{\mathrm{Co}}) \in \mathbb{Y}_{\mathrm{PS}}$ if and only if:

$$\int_0^T y(t)\, \mathrm{d}t = 0,\ \ \check{k}_{\mathrm{St}} (y) \leq k_{\mathrm{St}} \text{ and } \check{k}_{\mathrm{Co}} (y) \leq k_{\mathrm{Co}}. \tag{15.8}$$

Formula (15.6) is derived in [21].

Hydro generation produces an output flow $y \in L_+^\infty[0,T]$ from the inputs of storage capacity k_{St} (in kWh), turbine capacity k_{Tu} (in kW) and river flow $e \in L_+^\infty[0,T]$, whose rate $e(t)$ can also be measured in units of power (instead of volume per unit time). This is because the height at which water flows in and is stored, called the head, is taken to be constant. So the potential energy of water is in a constant proportion to its volume, and the energy can be referred to as "water". Since the turbine-generator's efficiency η_{Tu} is also taken to be constant, water can be measured as the output it actually yields on conversion (i.e., in kWh of electric energy). This redefines η_{Tu} as 1, i.e., in unit time, a unit turbine can convert a unit of stock into a unit of output.

A hydroelectric water storage policy generally consists of an output $y(t) \geq 0$ and a spillage $\sigma(t) \geq 0$. The resulting net outflow from the reservoir is $-\dot{s} = y - e + \sigma$ (where $s(t)$ is the water stock at time t, and $e(t)$ is the rate of river flow). Water can be held in storage at no running cost (or loss of stock). The long-run hydro production set is, therefore, the convex cone

$$
\begin{aligned}
(15.9)\quad \mathbb{Y}_{\mathrm{H}} := \Bigg\{ (y; -k_{\mathrm{St}}, k_{\mathrm{Tu}}; -e) \in L_+^\infty[0,T] \times \mathbb{R}_-^2 \times L_-^\infty[0,T] : 0 \leq y \leq k_{\mathrm{Tu}} \\
\text{and } \exists \sigma \in [0,e] \left(\int_0^T (y(t) - e(t) + \sigma(t))\,\mathrm{d}t = 0 \text{ and} \right. \\
\left. \exists s_0 \in \mathbb{R}\ \forall t \quad 0 \leq s_0 - \int_0^t (y(\tau) - e(\tau) + \sigma(\tau))\,\mathrm{d}\tau \leq k_{\mathrm{St}} \right) \Bigg\}.
\end{aligned}
$$

Comments (on hydro and pumped storage):

(1) If $k_{\mathrm{Tu}} \geq e$ then there is no need for spillage and, furthermore, it is feasible for the hydro plant to "coast", i.e., to generate at the rate $y(t) = e(t)$. In this case, all the incentive to use the reservoir comes from a time-dependent output price: if p were a constant, the plant might as well coast all the time.

(2) In both pumped storage and hydro generation, the flows to and from the reservoir are required to balance over the cycle ($\int_0^T \dot{s}\,\mathrm{d}t = 0$), i.e., the stock must be a periodic function of time. But its level at the beginning or end of a cycle is taken to be a costless decision variable, s_0. In other words, when it is first commissioned, the reservoir comes filled up to any required level at no extra cost, but its periodic operation thereafter is taken to be a technological constraint. For a brief comparison with the case of a *given* s_0, or a variable but costly s_0, see [21].

(3) In some ways, the hydro technology is analytically similar to pumped storage. But, unlike pumped storage, hydro is not a technique with conditionally fixed coefficients: although the conditional input demand for the turbine depends on nothing but the output (it is $\check{k}_{\mathrm{Tu}}(y) = \mathrm{EssSup}(y)$), various combinations of an inflow function and a reservoir capacity can yield the same output y (e.g., any e with $\int_0^T e\,\mathrm{d}t = \int_0^T y\,\mathrm{d}t$ and a high enough k_{St} will do).[77]

[77] Also, though this is only a technicality, the hydro technique has an infinity of input variables $(e(t))_{t\in[0,T]}$, unlike a c.f.c. technique as defined in Section 24.

16. Operation and valuation of electric power plants

For each of the plant types described in Section 15, the problem of profit-maximizing operation can be formulated as a doubly infinite linear programme for the output rate $y(t)$ at each time t (in kW), given a TOU electricity price rate $p(t)$ for each time t (in $\$$/kWh).

For a *thermal* plant of capacity k_θ with a unit running cost w_θ, the operation LP (reduced by working out the short-run cost as $w_\theta \int_0^T y\,\mathrm{d}t$) is:

$$\text{(16.1)} \qquad \text{Given } (p, k_\theta, w_\theta) \in L^1[0,T] \times \mathbb{R}_+ \times \mathbb{R}_+$$

$$\text{(16.2)} \qquad \text{maximize } \int_0^T (p(t) - w_\theta)\, y(t)\,\mathrm{d}t \text{ over } y \in L^\infty[0,T]$$

$$\text{(16.3)} \qquad \text{subject to: } \quad 0 \le y(t) \le k_\theta \quad \text{for a.e. } t.$$

Every optimal output is given by

$$\text{(16.4)} \qquad y(t) \in \begin{cases} \{0\} & \text{for } p(t) < w_\theta \\ [0, k_\theta] & \text{for } p(t) = w_\theta \\ \{k_\theta\} & \text{for } p(t) > w_\theta \end{cases}$$

i.e., measurable functions satisfying (16.4) form the solution set $\hat{Y}_\theta(p, k_\theta, w_\theta)$. So the plant's operating profit is $\Pi_{\mathrm{SR}}^\theta(p, k_\theta, w_\theta) = k_\theta \int_0^T (p(t) - w_\theta)^+\,\mathrm{d}t$, and its unit rental value (in $\$$/kW) is

$$\text{(16.5)} \qquad \hat{r}_\theta(p, k_\theta, w_\theta) = \frac{\partial \Pi_{\mathrm{SR}}^\theta}{\partial k_\theta}(p, k_\theta, w_\theta) = \int_0^T (p(t) - w_\theta)^+\,\mathrm{d}t \quad \text{if } k_\theta > 0.$$

Differentiation is the simplest way to value a unit of thermal capacity because the operation problem is so simple that its solution and value function can be calculated directly (i.e., without using a duality method). Of course, $\hat{r}_\theta$ can also be calculated by solving the dual problem of capacity valuation. The standard dual of the operation LP is the following programme for the flow of rent κ_θ (whose total for the cycle is r_θ), with ν_θ as the Lagrange multiplier for the nonnegativity constraint on y in (16.3):

$$\text{(16.6)} \qquad \text{Given } (p, k_\theta, w_\theta) \text{ as in (16.1)}$$

$$\text{(16.7)} \qquad \text{minimize } k_\theta \int_0^T \kappa_\theta(t)\,\mathrm{d}t \quad \text{over } \kappa_\theta \in L^1[0,T] \text{ and } \nu_\theta \in L^1[0,T]$$

$$\text{(16.8)} \qquad \text{subject to: } \quad \kappa_\theta \ge 0, \ \nu_\theta \ge 0$$

$$\text{(16.9)} \qquad p(t) - w_\theta = \kappa_\theta(t) - \nu_\theta(t) \quad \text{for a.e. } t.$$

The standard dual's inclusive form, introduced in (14.13)–(14.14), has also the dependent decision variable

$$\text{(16.10)} \qquad r_\theta = \int_0^T \kappa_\theta(t)\,\mathrm{d}t$$

which is the thermal plant's unit rental value. The standard dual solution, unique if $k_\theta > 0$, is

$$\hat{\kappa}_\theta = (p - w)^+ \quad \text{and} \quad \hat{\nu}_\theta = (p - w)^- \tag{16.11}$$

and hence, again,

$$\hat{r}_\theta = \int_0^T \hat{\kappa}_\theta (t) \, \mathrm{d}t = \int_0^T (p(t) - w)^+ \, \mathrm{d}t.$$

Comments (comparison of standard and intrinsic duals of the thermal plant operation programme):

(1) The standard perturbation of the primal LP (16.1)–(16.3), which produces the dual LP (16.6)–(16.9), consists in adding cyclically varying increments $(\Delta k_\theta (t), \Delta n_\theta (t))$ to the constants $(k_\theta, 0) \in \mathbb{R} \times \mathbb{R}$ in (16.3). The resource increments, $(\Delta k_\theta, -\Delta n_\theta) \in L^\infty \times L^\infty$, are paired with Lagrange multipliers $(\kappa_\theta, \nu_\theta) \in L^1 \times L^1$.

(2) By giving the unit rent's distribution over time, κ_θ—rather than only its total for the cycle, r_θ—the standard dual LP is the "fine" form of the valuation problem (in the sense of the first Comment in Section 14, with the integral $\kappa \mapsto \int \kappa (t) \, \mathrm{d}t$ as the adjoint operation $\sigma \mapsto B^{\mathrm{T}} \sigma$). The "coarse" form of valuation is a case of the intrinsic dual (5.13)–(5.14), which can be reformulated by substituting the input requirement functions (15.2) and (15.3) for $\check{k}$ and $\check{v}$, and L^∞_+ for Y_0, in either (24.12)–(24.15) or (28.6)–(28.9). The latter programme is then an LP for the single variable r_θ.

(3) In terms of our general duality scheme (Sections 5 and 14), r_θ is the intrinsic dual variable. Correspondence of notation between that scheme and its applications to the ESI is spelt out in Table 3.

For a *pumped-storage* plant with capacities $(k_{\mathrm{St}}, k_{\mathrm{Co}})$, the operation LP is:

$$\text{Given } (p; k_{\mathrm{St}}, k_{\mathrm{Co}}) \in L^1 [0, T] \times \mathbb{R}_+ \times \mathbb{R}_+ \tag{16.12}$$

$$\text{maximize } \int_0^T p(t) \, y(t) \, \mathrm{d}t \text{ over } y \in L^\infty [0, T] \text{ and } s_0 \in \mathbb{R} \tag{16.13}$$

$$\text{subject to:} \quad -k_{\mathrm{Co}} \le y(t) \le k_{\mathrm{Co}} \quad \text{for a.e. } t \tag{16.14}$$

$$\int_0^T y(t) \, \mathrm{d}t = 0 \tag{16.15}$$

$$0 \le s_0 - \int_0^t y(\tau) \, \mathrm{d}\tau \le k_{\mathrm{St}} \quad \text{for every } t. \tag{16.16}$$

Unlike the case of Π^θ_{SR}, there is no explicit formula for $\Pi^{\mathrm{PS}}_{\mathrm{SR}} (p, k_{\mathrm{St}}, k_{\mathrm{Co}})$, the operating profit of a pumped-storage plant; and both operation and rental valuation of a storage plant are best approached through the dual problem of capacity valuation. The standard dual of the operation LP is the following programme for: (i) the flow of reservoir's rent κ_{St}, and (ii) the flow of converter's rents $\kappa_{\mathrm{Co}} = \kappa_{\mathrm{Pu}} + \kappa_{\mathrm{Tu}}$, which it earns in its two modes of work, viz., charging the reservoir as a "pump" and discharging it as a "turbine". Their totals for the cycle are the unit rental values: (i) of the reservoir r_{St} (in $\$$/kWh), and (ii) of the converter r_{Co} (in $\$$/kW). The dual variables κ_{Pu} and κ_{Tu} range over $L^1 [0, T]$,

	intrinsic primal param.	intrinsic dual vbles	std primal param.	std dual vbles	relationship between intrinsic and standard dual vbles
GS	k [vect]	r [vect]	s [vect]	σ [vect]	$r = B^{\mathrm{T}}\sigma$
Th	k_θ [scal]	r_θ [scal]	$k_\theta(\cdot)$ $n_\theta(\cdot)$	$\kappa_\theta(\cdot)$ $\nu_\theta(\cdot)$	$r_\theta = \int \kappa_\theta \,\mathrm{d}t$
PS	k_{St} [scal] k_{Co} [scal]	r_{St} [scal] r_{Co} [scal]	$k_{\mathrm{St}}(\cdot)$ $n_{\mathrm{St}}(\cdot)$ $k_{\mathrm{Tu}}(\cdot)$ $k_{\mathrm{Pu}}(\cdot)$ ζ	$\kappa_{\mathrm{St}}(\mathrm{d}\cdot)$ $\nu_{\mathrm{St}}(\mathrm{d}\cdot)$ $\kappa_{\mathrm{Tu}}(\cdot)$ $\kappa_{\mathrm{Pu}}(\cdot)$ λ	$r_{\mathrm{St}} = \int \kappa_{\mathrm{St}}(\mathrm{d}t)$ $r_{\mathrm{Co}} = \int (\kappa_{\mathrm{Tu}} + \kappa_{\mathrm{Pu}})\,\mathrm{d}t$
Hy	k_{St} [scal] k_{Tu} [scal] $e(\cdot)$	r_{St} [scal] r_{Tu} [scal] $\psi(\cdot)$	$k_{\mathrm{St}}(\cdot)$ $n_{\mathrm{St}}(\cdot)$ $k_{\mathrm{Tu}}(\cdot)$ $n_{\mathrm{Tu}}(\cdot)$ ζ	$\kappa_{\mathrm{St}}(\mathrm{d}\cdot)$ $\nu_{\mathrm{St}}(\mathrm{d}\cdot)$ $\kappa_{\mathrm{Tu}}(\cdot)$ $\nu_{\mathrm{Tu}}(\cdot)$ λ	$r_{\mathrm{St}} = \int \kappa_{\mathrm{St}}(\mathrm{d}t)$ $r_{\mathrm{Tu}} = \int \kappa_{\mathrm{Tu}}\,\mathrm{d}t$ $\psi(t) = \lambda + (\kappa_{\mathrm{St}} - \nu_{\mathrm{St}})[0,t]$

TABLE 3. Correspondence of notation between the general duality scheme (Sections 5 and 15) and its applications to the ESI (Section 17). The abbreviations read: (i) in the leftmost column: GS = general scheme, Th = thermal generation, PS = pumped storage, Hy = hydro generation; (ii) elsewhere: St = storage reservoir, Co = converter, Pu/Tu = pump/turbine (two working modes of a reversible PS converter), Tu = hydro turbine. Functions of time are marked with a $(\cdot)$, and measures on the time interval are marked with a $(\mathrm{d}\cdot)$. In the general scheme, s and σ mean the standard parameters and Lagrange multipliers. But in the context of storage (both PS and Hy), s means the energy stock (and σ means spillage in Hy). Also, the intrinsic parameters and dual variables of the general scheme, r and k, correspond to (r, ψ) and (k, e) in the hydro problem.

like the κ_θ in (16.7). The space for κ_{St} is $\mathcal{M}[0,T]$, the space of Borel measures on $[0,T]$, which is the norm-dual of the space of continuous functions $\mathcal{C}[0,T]$. This is also the space for the multiplier ν_{St} for the nonnegativity constraint in (16.16). The multiplier for the balance constraint (16.15) is a scalar λ. So the LP of capacity valuation is:

(16.17) Given $(p; k_{\mathrm{St}}, k_{\mathrm{Co}})$ as in (16.12)

(16.18) $$\text{minimize } k_{\mathrm{St}} \int_{[0,T]} \kappa_{\mathrm{St}}(\mathrm{d}t) + k_{\mathrm{Co}} \int_0^T (\kappa_{\mathrm{Tu}} + \kappa_{\mathrm{Pu}})(t)\,\mathrm{d}t$$

(16.19) over $\lambda \in \mathbb{R}$ and $(\kappa_{\mathrm{St}}, \nu_{\mathrm{St}}, \kappa_{\mathrm{Pu}}, \kappa_{\mathrm{Tu}}) \in \mathcal{M} \times \mathcal{M} \times L^1 \times L^1$

(16.20) subject to: $(\kappa_{\mathrm{St}}, \nu_{\mathrm{St}}, \kappa_{\mathrm{Pu}}, \kappa_{\mathrm{Tu}}) \geq 0$

(16.21) $$\kappa_{\mathrm{St}}[0,T] = \nu_{\mathrm{St}}[0,T]$$

$$p(t) = \lambda + (\kappa_{\mathrm{St}} - \nu_{\mathrm{St}})[0,t] + \kappa_{\mathrm{Tu}}(t) - \kappa_{\mathrm{Pu}}(t) \quad \text{for a.e. } t. \tag{16.22}$$

The standard dual's inclusive form has also the dependent decision variables

$$r_{\mathrm{St}} = \int_0^T \kappa_{\mathrm{St}}(\mathrm{d}t) = \kappa_{\mathrm{St}}[0,T] \tag{16.23}$$

$$r_{\mathrm{Co}} = \int_0^T (\kappa_{\mathrm{Pu}}(t) + \kappa_{\mathrm{Tu}}(t))\,\mathrm{d}t \tag{16.24}$$

as per the last constraint of (14.14).

Comments (comparison of standard and intrinsic duals of the pumped-storage plant operation programme):

(1) The standard perturbation of the primal LP (16.12)–(16.16), which produces the dual LP (16.17)–(16.22), uses cyclically varying increments $(\Delta k_{\mathrm{St}}(t), \Delta n_{\mathrm{St}}(t))$ to the constants $(k_{\mathrm{St}}, 0)$ in (16.16). It also uses two separate increments $(\Delta k_{\mathrm{Pu}}(t), \Delta k_{\mathrm{Tu}}(t))$ to the two occurrences of k_{Co} in (16.14)—i.e., (16.14) is perturbed to:

$$-k_{\mathrm{Co}} - \Delta k_{\mathrm{Pu}}(t) \le y \le k_{\mathrm{Co}} + \Delta k_{\mathrm{Tu}}(t).$$

Additionally, a scalar $\Delta\zeta$ is used as an increment to the 0 on the r.h.s. of (16.15). The resource increments $\Delta k_{\mathrm{St}} \in \mathcal{C}$, $-\Delta n_{\mathrm{St}} \in \mathcal{C}$, $\Delta k_{\mathrm{Tu}} \in L^\infty$, $\Delta k_{\mathrm{Pu}} \in L^\infty$ and $\Delta\zeta \in \mathbb{R}$ are paired with the Lagrange multipliers $\kappa_{\mathrm{St}} \in \mathcal{M}$, $\nu_{\mathrm{St}} \in \mathcal{M}$, $\kappa_{\mathrm{Tu}} \in L^1$, $\kappa_{\mathrm{Pu}} \in L^1$ and $\lambda \in \mathbb{R}$. This perturbation scheme is described in detail in [21] and [27, Section 5].

(2) By giving the unit rents' distributions over time (and over the two conversion modes), κ_{St} and $\kappa_{\mathrm{Pu}} + \kappa_{\mathrm{Tu}}$—rather than only their totals for the cycle, r_{St} and r_{Co}—the standard dual LP is the "fine" form of the valuation problem (in the sense of the first Comment in Section 14). The "coarse" form of valuation is a case of the intrinsic dual (5.13)–(5.14) which can be reformulated by substituting the input requirement functions (15.6)–(15.7) for $\check{k}$, and (15.5) for Y_0, in either (24.12)–(24.15) or (28.6)–(28.9) with no $\check{v}$. The latter programme is then a semi-infinite LP for the variables r_{St} and r_{Co} (with an infinity of constraints).

The storage-plant valuation LP (16.17)–(16.22) can be transformed into an *unconstrained* convex programme by changing the variables from λ, $\kappa_{\mathrm{St}}(\mathrm{d}t)$ and $\nu_{\mathrm{St}}(\mathrm{d}t)$ to

$$\psi(t) = \lambda + (\kappa_{\mathrm{St}} - \nu_{\mathrm{St}})[0,t] \quad \text{for } t \in (0,T) \tag{16.25}$$

and by substituting $(p-\psi)^+$ and $(p-\psi)^-$ for κ_{Tu} and κ_{Pu} to eliminate these variables: see [21] or [27, Section 7] for details.[78] The new continuum of variables, ψ, is a function of bounded variation that can be interpreted as the TOU marginal value of the energy stock, i.e., its *shadow price*.

[78] This is done by using the constraints (16.21)–(16.22) and the disjointness conditions $\kappa_{\mathrm{Tu}} \wedge \kappa_{\mathrm{Pu}} = 0$ and $\kappa_{\mathrm{St}} \wedge \nu_{\mathrm{St}} = 0$, which are met by any solution to (16.17)–(16.22) if $k_{\mathrm{St}} > 0$ and $k_{\mathrm{Co}} > 0$; i.e., it is not optimal for the dual variables to overlap and partly cancel each other out in (16.22). Note that κ_{St} and ν_{St} are disjoint as measures on the circle, i.e., $\min\{\kappa_{\mathrm{St}}\{0,T\}, \nu_{\mathrm{St}}\{0,T\}\} = 0$ in addition to $\kappa_{\mathrm{St}} \wedge \nu_{\mathrm{St}} = 0$ in the lattice $\mathcal{M}[0,T]$.

Notation: The space $\mathrm{BV}(0,T)$ consists of all functions ψ of bounded variation on $(0,T)$ with $\psi(t)$ lying between the left and right limits, $\psi(t-)=\lim_{\tau\nearrow t}\psi(\tau)$ and $\psi(t+)=\lim_{\tau\searrow t}\psi(\tau)$.[79] A $\psi\in\mathrm{BV}(0,T)$ is extended by continuity to $[0,T]$; i.e., $\psi(0):=\psi(0+)$ and $\psi(T):=\psi(T-)$. The *cyclic positive variation* of ψ is

$$\mathrm{Var}_{\mathrm{c}}^{+}(\psi):=\mathrm{Var}^{+}(\psi)+(\psi(0)-\psi(T))^{+} \tag{16.26}$$

where $\mathrm{Var}^{+}(\psi)$ is the total *positive variation* (a.k.a. upper variation) of ψ, i.e., the supremum of $\sum_m(\psi(\overline{\tau}_m)-\psi(\underline{\tau}_m))^{+}$ over all finite sets of pairwise disjoint subintervals $(\underline{\tau}_m,\overline{\tau}_m)$ of $(0,T)$: see, e.g., [16, Section 8.1] for details.[80]

In these terms, the capacity valuation problem (for a pumped-storage plant) becomes the following programme for shadow-pricing the energy stock:

$$\text{Given } (p;k_{\mathrm{St}},k_{\mathrm{Co}})\in L^{1}[0,T]\times\mathbb{R}_{+}\times\mathbb{R}_{+} \tag{16.27}$$

$$\text{minimize } k_{\mathrm{St}}\,\mathrm{Var}_{\mathrm{c}}^{+}(\psi)+k_{\mathrm{Co}}\int_{0}^{T}|p(t)-\psi(t)|\,\mathrm{d}t \quad \text{over } \psi\in\mathrm{BV}(0,T). \tag{16.28}$$

Its main feature is the trade-off between minimizing the variation (which on its own would require setting ψ at a constant value) and minimizing the integral (which on its own would require setting ψ equal to p). This trade-off is what determines the extent to which local peaks of p should be "shaved off" and the troughs "filled in" to obtain the optimum shadow price $\hat{\psi}_{\mathrm{PS}}$, at least in the case of a piecewise strictly monotone p. The solution, shown in Figure 4a, is determined by constancy intervals of $\hat{\psi}_{\mathrm{PS}}$ around a local peak or trough of p. Unless $k_{\mathrm{St}}/k_{\mathrm{Co}}$ is relatively long, these intervals do not abut, and must all be of that length.[81] The optimal output has the "bang-coast-bang" form

$$\hat{y}_{\mathrm{PS}}(t)=k_{\mathrm{Co}}\operatorname{sgn}\left(p(t)-\hat{\psi}_{\mathrm{PS}}(t)\right) \tag{16.29}$$

i.e., $\hat{y}_{\mathrm{PS}}(t)$ equals k_{Co}, 0 or $-k_{\mathrm{Co}}$ if, respectively, $p(t)>\hat{\psi}_{\mathrm{PS}}(t)$, $p(t)=\hat{\psi}_{\mathrm{PS}}(t)$ or $p(t)<\hat{\psi}_{\mathrm{PS}}(t)$: see Figure 4b. The lowercase notation, $\hat{y}$ or $\hat{\psi}$, is used *only when the solution is unique*. In general, the solution sets for (16.12)–(16.16) and (16.27)–(16.28) are denoted by $\hat{Y}_{\mathrm{PS}}(p;k_{\mathrm{St}},k_{\mathrm{Co}})$ and $\hat{\Psi}_{\mathrm{PS}}(p;k_{\mathrm{St}},k_{\mathrm{Co}})$. More precisely, $y\in\hat{Y}_{\mathrm{PS}}$ means that y together with $s_0=\max_t\int_0^t y(\tau)\,\mathrm{d}\tau$ (which the lowest initial stock needed for the stock $s_0-\int_0^t y(\tau)\,\mathrm{d}\tau$ never to fall below 0) solves (16.12)–(16.16).

The stock-pricing programme (16.27)–(16.28) has a solution for every $k_{\mathrm{St}}>0$ and $k_{\mathrm{Co}}>0$ (by Lemma 23.1 or Part 2 of Proposition 27.2).[82] If p is continuous, i.e., $p\in\mathcal{C}[0,T]$, then there is a unique solution $\hat{\psi}_{\mathrm{PS}}(p;k_{\mathrm{St}},k_{\mathrm{Co}})$. It follows that the plant's operating profit $\Pi_{\mathrm{SR}}^{\mathrm{PS}}$ is differentiable in $(k_{\mathrm{St}},k_{\mathrm{Co}})$; equivalently, with this technology the

[79] The one-sided limits exist at every t and are equal nearly everywhere (n.e.), i.e., everywhere except for a countable set. Specification of $\psi(t)$ between $\psi(t-)$ and $\psi(t+)$ is unnecessary.

[80] The other term, $(\psi(0)-\psi(T))^{+}$, represents a possible jump of ψ at the instant separating two consecutive cycles.

[81] Matters complicate when the ratio $k_{\mathrm{St}}/k_{\mathrm{Co}}$ is comparable to the durations between the successive local peaks and troughs of p, so that the neighbouring constancy intervals of $\hat{\psi}$ start to abut; but a similar optimality rule applies to such clusters: see [21].

[82] When $k_{\mathrm{St}}>0$ but $k_{\mathrm{Co}}=0$, any constant ψ is a solution. When $k_{\mathrm{Co}}>0$ but $k_{\mathrm{St}}=0$, a solution exists if and only if $p\in\mathrm{BV}$, in which case it is unique, viz., $\hat{\psi}_{\mathrm{PS}}=p$.

programme (5.13)–(5.14) or (24.12)–(24.15) has a unique solution $\hat{r}$. In terms of $\hat{\psi}_{\mathrm{PS}}$, the unit rental values of the reservoir and the converter (in \$/kWh and \$/kW, respectively) are:

$$\frac{\partial \Pi_{\mathrm{SR}}^{\mathrm{PS}}}{\partial k_{\mathrm{St}}} = \hat{r}_{\mathrm{St}}(p, k_{\mathrm{St}}, k_{\mathrm{Co}}) = \int_0^T \hat{\kappa}_{\mathrm{St}}(\mathrm{d}t) = \mathrm{Var}_{\mathrm{c}}^{+}\left(\hat{\psi}_{\mathrm{PS}}\right) \tag{16.30}$$

$$\frac{\partial \Pi_{\mathrm{SR}}^{\mathrm{PS}}}{\partial k_{\mathrm{Co}}} = \hat{r}_{\mathrm{Co}}(p, k_{\mathrm{St}}, k_{\mathrm{Co}}) = \int_0^T (\hat{\kappa}_{\mathrm{Pu}} + \hat{\kappa}_{\mathrm{Tu}})(t)\,\mathrm{d}t = \int_0^T \left|p(t) - \hat{\psi}_{\mathrm{PS}}(t)\right| \mathrm{d}t. \tag{16.31}$$

For proofs, see [21] or [27, Sections 6 and 9].

As for the operation problem (16.12)–(16.16), it has a solution for any $p \in L^1[0,T]$ and every $(k_{\mathrm{St}}, k_{\mathrm{Co}}) \geq 0$, by Proposition 22.1 or 27.1 (Part 2). If p has no plateau (i.e., $\mathrm{meas}\{t : p(t) = \mathsf{p}\} = 0$ for every $\mathsf{p} \in \mathbb{R}$), then there is a unique solution $\hat{y}_{\mathrm{PS}}(p; k_{\mathrm{St}}, k_{\mathrm{Co}})$. It is given either by (16.29) itself (if $(k_{\mathrm{St}}, k_{\mathrm{Co}}) \gg 0$ and $p \in \mathcal{C}$), or by (16.29) with any $\psi \in \hat{\Psi}_{\mathrm{PS}}$ instead of $\hat{\psi}_{\mathrm{PS}}$ (if $(k_{\mathrm{St}}, k_{\mathrm{Co}}) \gg 0$ but $p \notin \mathcal{C}$). For proofs, see [21] or [27, Section 8].

Comments (interpretation of ψ, and assumptions on p in the pumped-storage problem):

(1) $\psi(t)$ has the interpretation of the shadow price of energy stock at time t. Heuristically, this follows from (16.25) and the marginal interpretations of κ, ν and λ, which are that: (i) κ_{St}, as the multiplier for the upper reservoir constraint, represents the reservoir capacity value, (ii) the multiplier ν_{St} has a similar interpretation for the lower reservoir constraint, and (iii) λ is the stock value at the beginning of cycle.

(2) This interpretation of ψ can be formalized as a rigorous marginal-value result by introducing a hypothetical inflow to the reservoir, $e \in L^\infty$, as a primal parameter with its own dual variable ψ. This means that (16.15) and (16.16) are perturbed by replacing y with $y - \Delta e$. Then (16.25) becomes a constraint of the dual problem, whose solution $\hat{\psi}$ equals $\nabla_e \Pi_{\mathrm{SR}}^{\mathrm{PS}}$ at $e = 0$. (This is formally similar to the hydro case (16.51), in which e is the river flow, and $\hat{\psi}$ equals $\nabla_e \Pi_{\mathrm{SR}}^{\mathrm{H}}$ at the given, positive e.)

(3) Time-continuity of the electricity tariff p, which guarantees uniqueness (and time-continuity) of the optimal price for energy stock ψ, is acceptable as an assumption for operation and valuation of storage plants because it can be verified for the general competitive equilibrium: see [28].

(4) Unlike price continuity, the no-plateau condition on the tariff p is rather questionable: it cannot hold in an equilibrium with continuous *quantity* trajectories (since it leads to the unique optimum $\hat{y}_{\mathrm{PS}}$, which is discontinuous because it takes only the three values $\pm k_{\mathrm{Co}}$ and 0, as per (16.29)).[83] Such an equilibrium is made

[83] Furthermore, a time-continuous optimal output from storage cannot be unique (unless $k_{\mathrm{St}} = 0$ or $k_{\mathrm{Co}} = 0$). To see this in detail, take any $y \in \mathcal{C}[0,T] \cap \hat{Y}_{\mathrm{PS}}(p; k_{\mathrm{St}}, k_{\mathrm{Co}})$. With $(k_{\mathrm{St}}, k_{\mathrm{Co}}) \gg 0$, if p is nonconstant on $[0,T]$ then $0 \notin \hat{Y}_{\mathrm{PS}}$: see [21]. And if p *is* a constant then y can be chosen to be nonzero (since every feasible y is then optimal). So the open set $\{t : 0 < y(t) < k_{\mathrm{Co}}\}$ is nonempty; let A be one of its component intervals. Then $p = \psi = \mathrm{const.}$ on A for each $\psi \in \hat{\Psi}_{\mathrm{PS}}$ because: (i) $y(t) = \pm k_{\mathrm{Co}}$ whenever $p(t) \neq \psi(t)$, and (ii) $0 < s < k_{\mathrm{St}}$ on A, which implies that $\psi = \mathrm{const.}$ on A. (Both (i) and the implication in (ii) are Complementary Slackness Conditions: see [21] or [27, Section 6].) Since $p_{|A}$

possible only by the presence of intervals on which an optimal y can gradually change from 0 to $\pm k_{\mathrm{Co}}$ because $p = \psi = \mathrm{const}$. But all this means is that, at a price system consistent with output continuity, the storage operation problem is not fully solved by stock pricing alone.

For a *hydro* plant with capacities $(k_{\mathrm{St}}, k_{\mathrm{Tu}})$ and an inflow $e(t) \le k_{\mathrm{Tu}}$ (for a.e. t), the operation LP is:

$$\text{(16.32)} \qquad \text{Given } (p; k_{\mathrm{St}}, k_{\mathrm{Tu}}; e) \in L^1_+[0,T] \times \mathbb{R}_+ \times \mathbb{R}_+ \times L^\infty_+[0,T] \text{ with } k_{\mathrm{Tu}} \ge e$$

$$\text{(16.33)} \qquad \text{maximize } \int_0^T p(t)\, y(t)\, \mathrm{d}t \text{ over } y \in L^\infty[0,T] \text{ and } s_0 \in \mathbb{R}$$

$$\text{(16.34)} \qquad \text{subject to: } \quad 0 \le y(t) \le k_{\mathrm{Tu}} \quad \text{for a.e. } t$$

$$\text{(16.35)} \qquad \int_0^T (y(t) - e(t))\, \mathrm{d}t = 0$$

$$\text{(16.36)} \qquad 0 \le s_0 - \int_0^t (y(\tau) - e(\tau))\, \mathrm{d}\tau \le k_{\mathrm{St}} \quad \text{for every } t.$$

As with pumped storage, there is no explicit formula for the hydro plant's operating profit $\Pi^{\mathrm{H}}_{\mathrm{SR}}(p; k_{\mathrm{St}}, k_{\mathrm{Tu}}; e)$, and both operation and rental valuation of a hydro plant are best approached through the dual problem of fixed-input valuation, which is an LP for: (i) the flow of reservoir's unit rent κ_{St}, (ii) the flow of turbine's unit rent κ_{Tu}, and (iii) the river's unit rent, i.e., the shadow price of water ψ. By including ψ but not r among the dual variables, this is a partly inclusive form of the standard dual LP. The fully inclusive form has also r_{St} and r_{Tu}, the rental values of the reservoir (in \$/kWh) and of the turbine (in \$/kW), but these are simply the totals of κ_{St} and κ_{Tu} for the cycle. The dual variable κ_{Tu} ranges over $L^1[0,T]$, and the space for κ_{St} is the space of measures $\mathcal{M}[0,T]$, as in pumped storage. The space for ψ can be $L^1[0,T]$ formally, but actually ψ is constrained to $\mathrm{BV}(0,T)$ by (16.43). The multipliers for the nonnegativity constraints in (16.34) and (16.36) are $\nu_{\mathrm{Tu}} \in L^1[0,T]$ and $\nu_{\mathrm{St}} \in \mathcal{M}[0,T]$. The multiplier for the balance constraint (16.35) is a scalar λ. So the LP of fixed-input valuation is:

$$\text{(16.37)} \qquad \text{Given } (p; k_{\mathrm{St}}, k_{\mathrm{Tu}}; e) \text{ as in (16.32)}$$

$$\text{(16.38)} \qquad \text{minimize } k_{\mathrm{St}} \int_{[0,T]} \kappa_{\mathrm{St}}(\mathrm{d}t) + k_{\mathrm{Tu}} \int_0^T \kappa_{\mathrm{Tu}}(t)\, \mathrm{d}t + \int_0^T \psi(t)\, e(t)\, \mathrm{d}t$$

$$\text{(16.39)} \qquad \text{over } \lambda \in \mathbb{R},\ \psi \in L^1[0,T] \text{ and } (\kappa_{\mathrm{St}}, \nu_{\mathrm{St}}; \kappa_{\mathrm{Tu}}, \nu_{\mathrm{Tu}}) \in \mathcal{M} \times \mathcal{M} \times L^1 \times L^1$$

$$\text{(16.40)} \qquad \text{subject to: } (\kappa_{\mathrm{St}}, \nu_{\mathrm{St}}; \kappa_{\mathrm{Tu}}, \nu_{\mathrm{Tu}}) \ge 0$$

$$\text{(16.41)} \qquad \kappa_{\mathrm{St}}[0,T] = \nu_{\mathrm{St}}[0,T]$$

$$\text{(16.42)} \qquad p(t) = \psi(t) + \kappa_{\mathrm{Tu}}(t) - \nu_{\mathrm{Tu}}(t) \quad \text{for a.e. } t$$

$$\text{(16.43)} \qquad \psi(t) = \lambda + (\kappa_{\mathrm{St}} - \nu_{\mathrm{St}})[0,t] \quad \text{for a.e. } t.$$

$= \mathrm{const.}$, y can be modified on A, without loss of optimality, to any y' such that $\int_A y'\, \mathrm{d}t = \int_A y\, \mathrm{d}t$ and $0 \le y' \le k_{\mathrm{Co}}$ on A (with $y' = y$ outside of A). A similar argument applies to the set $\{t : -k_{\mathrm{Co}} < y(t) < 0\}$.

The dual's fully inclusive form has also the remaining dependent decision variables

$$r_{\mathrm{St}} = \int_0^T \kappa_{\mathrm{St}}\,(\mathrm{d}t) \tag{16.44}$$

$$r_{\mathrm{Tu}} = \int_0^T \kappa_{\mathrm{Tu}}\,(t)\;\mathrm{d}t. \tag{16.45}$$

Comments (comparison of the partly inclusive standard, standard, and intrinsic duals of the hydro plant operation programme):

(1) The perturbation that produces (16.37)–(16.43) as the dual of (16.32)–(16.36) includes an increment $\Delta e\,(t)$ in addition to the standard perturbation (which uses cyclically varying increments $(\Delta k_{\mathrm{St}}\,(t),\ \Delta n_{\mathrm{St}}\,(t);\ \Delta k_{\mathrm{Tu}}\,(t),\ \Delta n_{\mathrm{Tu}}\,(t))$ to the constants $(k_{\mathrm{St}},\ 0;\ k_{\mathrm{Tu}},\ 0)$ in (16.36) and (16.34), as well as a scalar $\Delta\zeta$ as an increment to the 0 on the r.h.s. of (16.35)). The resource increments $\Delta e \in L^\infty$, $\Delta k_{\mathrm{St}} \in \mathcal{C}$, $-\Delta n_{\mathrm{St}} \in \mathcal{C}$, $\Delta k_{\mathrm{Tu}} \in L^\infty$, $-\Delta n_{\mathrm{Tu}} \in L^\infty$ and $\Delta\zeta \in \mathbb{R}$ are paired with the dual variables $\psi \in L^1$, $\kappa_{\mathrm{St}} \in \mathcal{M}$, $\nu_{\mathrm{St}} \in \mathcal{M}$, $\kappa_{\mathrm{Tu}} \in L^1$, $\nu_{\mathrm{Tu}} \in L^1$ and $\lambda \in \mathbb{R}$. This perturbation scheme is described in detail in [24].

(2) Though it is more transparent to have an explicit dual variable for each parameter, the nonstandard dual variable ψ (paired with e) can be eliminated by replacing it in (16.38) and (16.42) with its equivalent in terms of the standard dual variables (16.43). This reduces the valuation LP (16.37)–(16.43) to the standard dual of the hydro operation LP (16.32)–(16.36), i.e., to the dual arising from the same perturbation as above but without Δe.

(3) By giving the unit rents' distributions over time, κ_{St} and κ_{Tu}—rather than only their totals for the cycle, r_{St} and r_{Tu}—the above dual LP is the "fine" form of the valuation problem. The "coarse" form of valuation is a case of the intrinsic dual (5.13)–(5.14); it is a programme for r_{St}, r_{Tu} and ψ.

The hydro-plant valuation LP (16.37)–(16.43) can be transformed into an *unconstrained* convex programme for the water price ψ by using the constraints (16.42) and (16.43) to substitute: $(p-\psi)^+$ and $(p-\psi)^-$ for κ_{Tu} and ν_{Tu}, $(\mathrm{d}\psi)^+$ and $(\mathrm{d}\psi)^-$ for κ_{St} and ν_{St}, and any number between $\psi\,(0+)$ and $\psi\,(T-)$ for λ: see [24] for details. In these terms, the fixed-input valuation problem (for a hydro plant) becomes:

$$\text{Given } (p; k_{\mathrm{St}}, k_{\mathrm{Tu}}; e) \in L^1_+\,[0,T] \times \mathbb{R}_+ \times \mathbb{R}_+ \times L^\infty_+\,[0,T] \text{ with } k_{\mathrm{Tu}} \geq e \tag{16.46}$$

$$\text{minimize } k_{\mathrm{St}}\,\mathrm{Var}^+_{\mathrm{c}}\,(\psi) + k_{\mathrm{Tu}} \int_0^T (p\,(t) - \psi\,(t))^+\,\mathrm{d}t + \int_0^T \psi\,(t)\,e\,(t)\,\mathrm{d}t \tag{16.47}$$

$$\text{over } \psi \in \mathrm{BV}\,(0,T)\,. \tag{16.48}$$

Recall that $\mathrm{Var}^+_{\mathrm{c}}\,(\psi)$, defined by (16.26), is the total of all rises of ψ over the cycle.

If $k_{\mathrm{Tu}} > e\,(t) > 0$ for every t, then the sum of the two integrals in (16.47) has a minimum at (and only at) $\psi = p$. Therefore, the programme's main feature is the trade-off between minimizing the variation (which on its own would require setting ψ at a constant value) and minimizing the sum of integrals (which on its own would require setting ψ equal to p). This trade-off is what determines the extent to which the local peaks of p should be "shaved off" and the troughs "filled in" to obtain the optimum shadow price $\hat{\psi}_{\rm H}$, at least in the case that p is piecewise strictly monotone and $k_{\rm Tu} > e > 0$ at all times. The solution is determined by constancy intervals of $\hat{\psi}_{\rm H}$. If $k_{\rm St}/\operatorname{Sup}(e)$ and $k_{\rm St}/(k_{\rm Tu} - \operatorname{Inf}(e))$, which are upper bounds on the times needed to fill up and to empty the reservoir, are sufficiently short, then the constancy intervals do not abut. Around a trough of p there is an interval $(\underline{t}, \overline{t})$ characterized by $\int_{\underline{t}}^{\overline{t}} e(t)\,\mathrm{d}t = k_{\rm St}$, on which $p(t) < \hat{\psi}_{\rm H}$ throughout. Around a local peak of p there is an interval $(\underline{t}, \overline{t})$ characterized by $\int_{\underline{t}}^{\overline{t}} (k_{\rm Tu} - e(t))\,\mathrm{d}t = k_{\rm St}$ on which $p(t) > \hat{\psi}_{\rm H}$ throughout. The optimal output has the "bang-coast-bang" form

$$\hat{y}_{\rm H}(t) = \begin{cases} k_{\rm Tu} & \text{if } p(t) > \hat{\psi}_{\rm H}(t) \\ e(t) & \text{if } p(t) = \hat{\psi}_{\rm H}(t) \\ 0 & \text{if } p(t) < \hat{\psi}_{\rm H}(t) \end{cases} . \tag{16.49}$$

The lowercase notation, $\hat{y}$ or $\hat{\psi}$, is used *only when the solution is unique*. In general, the solution sets for (16.32)–(16.36) and (16.46)–(16.48) are denoted by $\hat{Y}_{\rm H}(p; k_{\rm St}, k_{\rm Tu}; e)$ and $\hat{\Psi}_{\rm H}(p; k_{\rm St}, k_{\rm Tu}; e)$.

The shadow-pricing programme (16.46)–(16.48) has a solution by Lemma 23.1, if

$$k_{\rm St} > 0 \text{ and } k_{\rm Tu} > \operatorname{EssSup}(e) \geq \operatorname{EssInf}(e) > 0. \tag{16.50}$$

If additionally p is continuous, i.e., $p \in \mathcal{C}_+[0,T]$, then there is a unique solution

$$\hat{\psi}_{\rm H}(p; k_{\rm St}, k_{\rm Tu}; e) = \nabla_e \Pi_{\rm SR}^{\rm H}(p; k_{\rm St}, k_{\rm Tu}; e). \tag{16.51}$$

This is the TOU price of water (unit value of the river flow). It follows that the plant's operating profit $\Pi_{\rm SR}^{\rm H}$ is also differentiable in $(k_{\rm St}, k_{\rm Tu})$. In terms of $\hat{\psi}_{\rm H}$, the unit rental values of the reservoir and the turbine (in \$/kWh and \$/kW, respectively) are:

$$\hat{r}_{\rm St}(p; k_{\rm St}, k_{\rm Tu}; e) = \frac{\partial \Pi_{\rm SR}^{\rm H}}{\partial k_{\rm St}} = \operatorname{Var}_{\rm c}^{+}\left(\hat{\psi}_{\rm H}\right) \tag{16.52}$$

$$\hat{r}_{\rm Tu}(p; k_{\rm St}, k_{\rm Tu}; e) = \frac{\partial \Pi_{\rm SR}^{\rm H}}{\partial k_{\rm Tu}} = \int_0^T \left(p(t) - \hat{\psi}_{\rm H}(t)\right)^+ \mathrm{d}t. \tag{16.53}$$

For proofs, see [24].

As for the operation problem (16.32)–(16.36), it has a solution for any $p \in L_+^1[0,T]$ and every $(k_{\rm St}, k_{\rm Tu}) \geq 0$ and $e \leq k_{\rm Tu}$, by Proposition 22.1. If p has no plateau (i.e., $\operatorname{meas}\{t : p(t) = \mathsf{p}\} = 0$ for every $\mathsf{p} \in \mathbb{R}$), then there is a unique solution $\hat{y}_{\rm H}(p; k_{\rm St}, k_{\rm Tu}; e)$. It is given either by (16.49) itself (if (16.50) holds and $p \in \mathcal{C}$), or by (16.29) with any $\psi \in \hat{\Psi}_{\rm H}$ instead of $\hat{\psi}_{\rm H}$ (if (16.50) holds but $p \notin \mathcal{C}$). For proofs, see [24].

Comments (on assumptions on p and properties of water value ψ in the hydro problem):

(1) As in the case of thermal generation with pumped storage, time-continuity of the electricity tariff p, which guarantees uniqueness and continuity of the optimal water price ψ, can be verified for the general competitive equilibrium with hydro-thermal generation. The much less important condition that p have no plateau is, again, questionable: it cannot hold in an equilibrium with continuous *quantity*

trajectories (since it leads to the unique optimum $\hat{y}_{\mathrm{H}}$, which is discontinuous under (16.50) because it takes only the values k_{Tu}, $e(t)$, and 0, as per (16.49)).

(2) When $e \nleq k_{\mathrm{Tu}}$ (i.e., when the policy of pure "coasting", $y = e$ with no spillage, is infeasible), the hydro operation and valuation LPs must be modified in the way indicated in [24]. This complicates the solution, and an optimal water price ψ need *not* then be unique or continuous over time (despite the continuity of the electricity price p).

Comments (on choice of space for dual variables):

(1) For "automatic" proofs of the dual LPs' solubility, which are based on Slater's Condition, the dual-variable spaces must be the norm-duals of the corresponding primal perturbation spaces (L^∞ and $\mathcal{C}$). This means using $L^{\infty *}$, instead of L^1, as the space for each of the dual variables paired to those primal perturbations that range over L^∞ (viz., for κ_θ and ν_θ in (16.6)–(16.9), for κ_{Tu} and κ_{Pu} in (16.17)–(16.22), and for ψ, κ_{Tu} and ν_{Tu} in (16.37)–(16.43))—just as $\mathcal{M} = \mathcal{C}^*$ serves as the space for the dual variables paired to perturbations that range over $\mathcal{C}$ (viz., for κ_{St} and ν_{St}). This is because, like $\mathcal{C}_+$, the nonnegative cone L^∞_+ has a nonempty norm-interior, and so positivity of the capacities k_θ, $(k_{\mathrm{St}}, k_{\mathrm{Co}})$ or $(k_{\mathrm{St}}, k_{\mathrm{Tu}})$, together with (16.50) for the hydro plant, imply that Slater's Condition, as generalized in [44, (8.12)] to infinite-dimensional inequality constraints, holds with the supremum norm topology on the primal parameter spaces L^∞ and $\mathcal{C}$. This ensures the existence of a dual optimum in the norm-dual spaces (i.e., κ_θ and ν_θ in $L^{\infty *}$, κ_{Tu} and κ_{Pu} in $L^{\infty *}$, κ_{St} and ν_{St} in $\mathcal{M}$, and ψ, κ_{Tu} and ν_{Tu} in $L^{\infty *}$). Density representation of the dual variables other than κ_{St} and ν_{St} comes only from the problems' structures *and* the assumption that p is a density: since $p \in L^1$, every optimal κ_θ and ν_θ (for a thermal plant) is actually in L^1 by (16.11), as is every optimal κ_{Tu} and κ_{Pu} (for a storage plant), and every optimal κ_{Tu} and ν_{Tu} (for a hydro plant). And every feasible ψ is in $\mathrm{BV} \subset L^1$ by (16.43). This is what justifies the use of L^1 (rather than $L^{\infty *}$) in the above formulations of the dual LPs (when $p \in L^1$).

(2) In the more general case of a $p \in L^{\infty *}$, the generating capacity's optimal rent flow, κ_θ or κ_{Tu}, are in $L^{\infty *}$ (although the corresponding ν_θ and κ_{Pu} or ν_{Tu} are in L^1 because $p \geq 0$). Also, when $p \in L^{\infty *}$, the degenerate case of zero storage capacity (with a positive conversion capacity) provides an example of a duality gap (Appendix A).

17. Peak-load pricing of electricity with pumped storage or hydro generation

Our introductory application of the short-run approach to electricity pricing, in Section 2, is made simple by cross-price independence of short-run supply and the assumed cross-price independence of demand. In such a case, the short-run general equilibrium (SRGE) can be found separately for each time instant (by intersecting the demand and supply curves). It is equally simple to calculate the unit operating profit, and use it as an imputed capacity value to work out the long-run general equilibrium (LRGE).

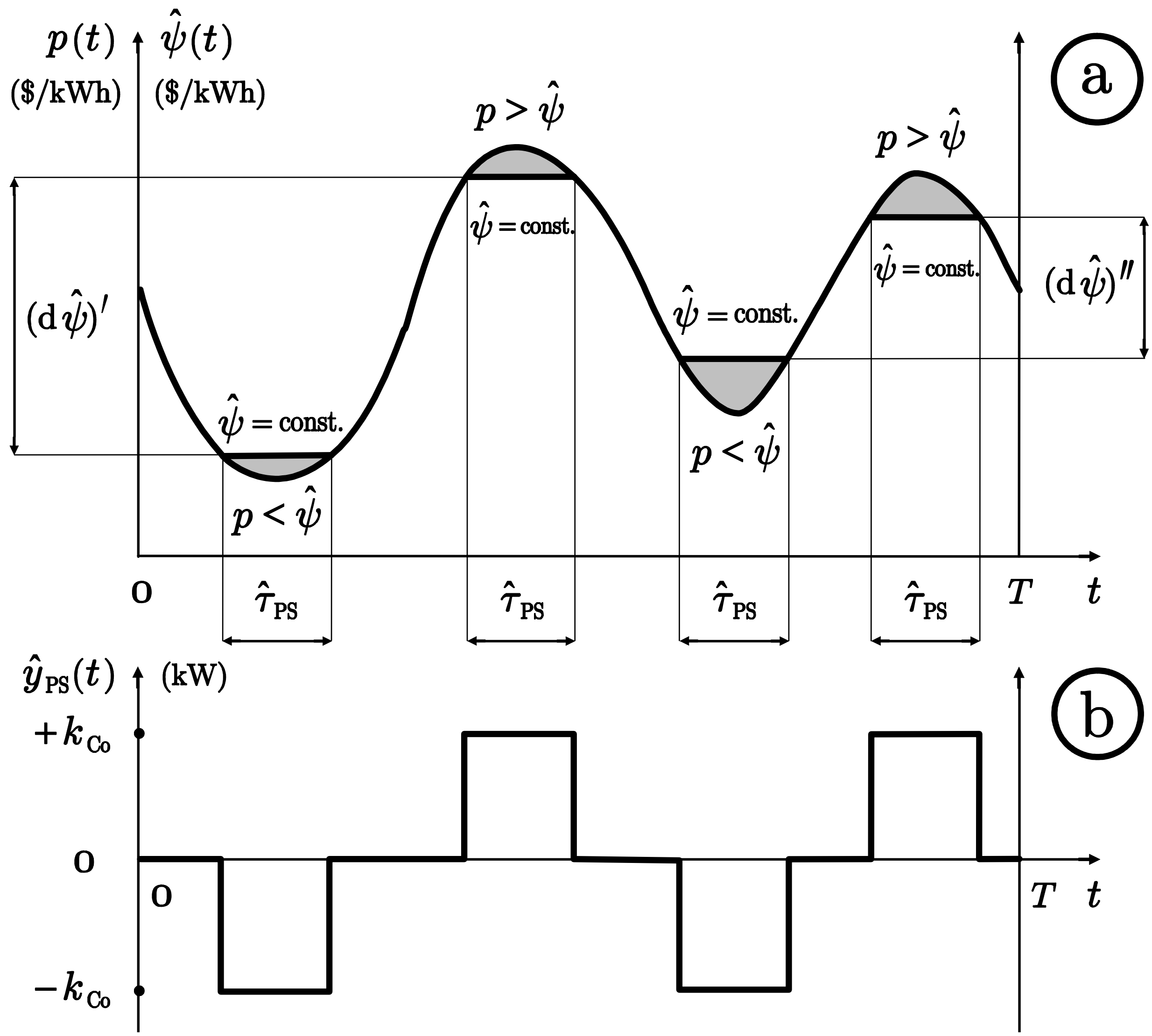

FIGURE 4. Trajectories of: (a) shadow price of stock $\hat{\psi}$, and (b) output of pumped-storage plant (optimum storage policy) $\hat{y}_{\mathrm{PS}}$ in Section 16, and in Theorem 17.1. Unit rent for storage capacity is $\mathrm{Var}_{\mathrm{c}}^{+}\left(\hat{\psi}\right) = \left(\mathrm{d}\hat{\psi}\right)' + \left(\mathrm{d}\hat{\psi}\right)''$, the sum of rises of $\hat{\psi}$. Unit rent for conversion capacity is $\int_0^T \left|p\left(t\right) - \hat{\psi}\left(t\right)\right| \mathrm{d}t$, the sum of grey areas. By definition, $\hat{\tau}_{\mathrm{PS}} = k_{\mathrm{St}}/k_{\mathrm{Co}}$.

That analysis is now extended to apply to cross-price dependent demand and to include storage or hydro plants, whose profit-maximizing output is also cross-price dependent. Though the resulting general equilibrium problem cannot be solved by explicit formulae, the short-run approach does make it tractable: first, short-run supply can be determined by solving the plant operation LPs; then an iterative procedure (such as Walrasian tatonnement) can be used to find the short-run equilibrium; and finally plant valuations, obtained from dual LP solutions, can be used to find the long-run equilibrium by another iteration (as is indicated in Figure 3). A system of equilibrium conditions required for this approach is obtained by placing the operation and valuation results for the ESI's plants into the SRP programme-based LRGE system, (13.11)–(13.15) with (13.18)–(13.19). We do this first for an electricity supply technology that combines thermal generation with pumped storage.

Except for the storage capacity, all the ESI's inputs are taken to have *fixed* prices: $r^{\mathrm{F}}_{\mathrm{Th}} = \left(r^{\mathrm{F}}_1, \ldots, r^{\mathrm{F}}_{\Theta}\right)$ for the thermal generating capacities, $w = (w_1, \ldots, w_{\Theta})$ for the corresponding fuels, and $r^{\mathrm{F}}_{\mathrm{Co}}$ for the storage plant's converter. There is a location where an energy reservoir of capacity k_{St} can be constructed at a cost $G(k_{\mathrm{St}})$. Usually, the marginal cost is increasing, i.e., the construction cost is a strictly convex and increasing function, $G\colon \left[0, \overline{k}_{\mathrm{St}}\right] \to \mathbb{R}_+$ with $G(0) = 0$.[84] (This is especially so with the PWES and CAES techniques, which utilize special geological features.) In the terminology of Section 13, the reservoir is the single equilibrium-priced capital input; all the others have fixed prices. Formally, $\Phi^{\mathrm{E}}_{\mathrm{PS}} = \{\mathrm{St}\}$, $\Phi^{\mathrm{F}}_{\mathrm{PS}} = \{\mathrm{Co}\}$, and $\Phi^{\mathrm{F}}_{\theta} = \Phi_{\theta} = \{\theta\}$ for each $\theta \in \Theta$ (the set of thermal plant types).

All input demand for electricity is taken to come from a single Industrial User, who produces a final good from inputs of electricity and the numeraire, z and n. His production function, $(z, n) \mapsto F(z, n)$, is assumed to be strictly concave and increasing, and Mackey continuous, i.e., $\mathrm{m}(L^{\infty} \times \mathbb{R}, L^1 \times \mathbb{R})$-continuous on $L^{\infty}_{+}[0, T] \times \mathbb{R}_+$. One example is the additively separable form for $F(\cdot, n)$, i.e., the integral functional $F(z, n) = \int_0^T f(t, z(t), n)\,\mathrm{d}t$, where f meets the conditions of [7, p. 535].[85]

A complete commodity bundle consists, then, of electricity (differentiated over time), the ESI's inputs (viz., the thermal capacities, the fuels, and the storage and conversion capacities), the produced final good and the numeraire. These quantities and their prices are always listed in this order, but those which are irrelevant in a particular context are omitted (as in Section 13). So a complete price system is $(p; r_{\mathrm{Th}}; w; r_{\mathrm{PS}}; \varrho, 1)$ with $r_{\mathrm{PS}} = (r_{\mathrm{St}}, r_{\mathrm{Co}})$, but a consumer price system is just $(p; \varrho, 1) \in L^1[0, T] \times \mathbb{R}^2$—since a consumption bundle consists of electricity, the produced final good and the numeraire, denoted by $(x; \varphi, m) \in L^{\infty}[0, T] \times \mathbb{R}^2$. The utility function, U_h for household h, is also assumed to be Mackey continuous, i.e., $\mathrm{m}(L^{\infty} \times \mathbb{R}^2, L^1 \times \mathbb{R}^2)$-continuous on the consumption set $L^{\infty}_{+}[0, T] \times \mathbb{R}^2_{+}$. Each household's initial endowment is a quantity of the

[84] A typical *non*-convex G is one that is concave on an "initial" interval $\left[0, \widetilde{k}\right]$, and convex on $\left(\widetilde{k}, \overline{k}\right]$. A limiting case of this arises from a nonzero setup cost $G(0+) > 0$, with G convex on $\left(0, \overline{k}\right]$. Supply (of storage capacity) is then discontinuous at the price equal to the minimum average cost, which is attained at some $\underline{k}$ greater than the point of inflection $\widetilde{k}$, i.e., at the price $\underline{r} := \min_k (G(k)/k) =: G(\underline{k})/\underline{k}$. The profit-maximizing supply is 0 at $r < \underline{r}$, but it exceeds $\underline{k}$ at $r > \underline{r}$. At $r = \underline{r}$, it takes the two values $\{0, \underline{k}\}$, but none of the intermediate values. The total supply curve for this form of marginal and average costs is discussed in, e.g., [17, 4-4: Figure 4-5].

[85] That is, the function $t \mapsto f(t, \mathsf{z}, n)$ is integrable on $[0, T]$ for each $(\mathsf{z}, n) \in \mathbb{R}^2_+$, and the function $(\mathsf{z}, n) \mapsto f(t, \mathsf{z}, n)$ is concave, increasing and continuous on $\mathbb{R}^2_+$, with $f(t, 0, 0) = 0$ for every $t \in [0, T]$. For a short proof of the Mackey continuity of F, see [25].

numeraire $m_h^{\mathrm{En}} > 0$. The household's share in the User Industry's profit is $\varsigma_{h\,\mathrm{IU}}$, and its share of profit from supplying the storage capacity is $\varsigma_{h\,\mathrm{St}}$.

By feeding the programming results summarized in Section 16 into the framework of Section 13, we next characterize long-run equilibrium by optimality of the ESI's investments in addition to the SRGE system, which is either (17.4)–(17.9) for pumped storage or (17.14)–(17.19) for hydro-thermal generation. For simplicity, we assume that all the equilibrium capacities are positive, i.e., that each type of plant is built (in general, some plant types might not be built because of their costs).

Theorem 17.1 (Characterization of long-run equilibrium with pumped storage)**.** *Assume that the ESI's technology consists of thermal generation techniques (Θ) and a pumped storage technique. Then a price system made up of:*

- *a time-continuous electricity tariff $p^\star \in \mathcal{C}[0,T]$*
- *a rental price for storage capacity $r_{\mathrm{St}}^\star$*
- *a price $\varrho^\star > 0$ for the produced final good*
- *the given prices for fuel and the generating capacities (viz., r_θ^{F} for thermal capacity of type θ and w_θ for its fuel, and r_{Co} for the converter capacity)*

and an allocation made up of:

- *an output $y_\theta^\star \in L_+^\infty[0,T]$ from the thermal plant of type θ with*
  - *a capacity $k_\theta^\star > 0$*
  - *a fuel input $v_\theta^\star$ (for each θ)*
- *an output $y_{\mathrm{PS}}^\star \in L^\infty[0,T]$ from a pumped-storage plant with*
  - *a storage capacity $k_{\mathrm{St}}^\star > 0$*
  - *a conversion capacity $k_{\mathrm{Co}}^\star > 0$*
- *a consumption bundle $(x_h^\star, \varphi_h^\star, m_h^\star) \in L_+^\infty[0,T] \times \mathbb{R}_+ \times \mathbb{R}_+$ for each household h*
- *an input-output bundle of the User Industry $(-z^\star, F(z^\star, n^\star), -n^\star) \in L_-^\infty[0,T] \times \mathbb{R}_+ \times \mathbb{R}_-$*

form a long-run competitive equilibrium if and only if:

(1) (a) *(Equality of ESI's capital-input prices to profit-imputed marginal values) For each $\theta = 1, \ldots, \Theta$*

$$r_\theta^{\mathrm{F}} = \int_0^T (p^\star(t) - w_\theta)^+ \, \mathrm{d}t \tag{17.1}$$

$$r_{\mathrm{St}}^\star = \mathrm{Var}_{\mathrm{c}}^+(\psi^\star) \tag{17.2}$$

$$r_{\mathrm{Co}} = \int_0^T |p^\star(t) - \psi^\star(t)| \, \mathrm{d}t \tag{17.3}$$

where $\psi^\star := \hat{\psi}_{\mathrm{PS}}(p^\star, k_{\mathrm{St}}^\star, k_{\mathrm{Co}}^\star)$ is the optimal price of energy stock, i.e., the unique solution to the programme (16.27)–(16.28) with $(p^\star; k_{\mathrm{St}}^\star, k_{\mathrm{Co}}^\star)$ as data.[86]

[86] Since $p^\star \in \mathcal{C}[0,T]$, the optimal ψ is indeed unique [27, Lemma 8].

(b) *(Operating profit maximization by ESI) For each* θ

$$y_\theta^\star(t) \in S\left(p^\star(t), k_\theta^\star, w_\theta\right) := \begin{cases} \{0\} & \text{if } p^\star(t) < w_\theta \\ [0, k_\theta^\star] & \text{if } p^\star(t) = w_\theta \\ \{k_\theta^\star\} & \text{if } p^\star(t) > w_\theta \end{cases} \quad \text{for a.e. } t \tag{17.4}$$

$$v_\theta^\star = \int_0^T y_\theta^\star(t)\, \mathrm{d}t. \tag{17.5}$$

And, with $(p^\star; k_{\mathrm{St}}^\star, k_{\mathrm{Co}}^\star)$ *as the data,*

$$y_{\mathrm{PS}}^\star \text{ solves the linear programme (16.12) to (16.16)} \tag{17.6}$$

(which implies that $y_{\mathrm{PS}}^\star(t) = k_{\mathrm{Co}}$ *when* $p^\star(t) > \psi^\star(t)$ *and* $y_{\mathrm{PS}}^\star(t) = -k_{\mathrm{Co}}$ *when* $p^\star(t) < \psi^\star(t)$*).*

(2) *(Profit maximization by User Industry)*[87]

$$(p^\star, 1) \in \varrho^\star \widehat{\partial} F(z^\star, n^\star). \tag{17.7}$$

(3) *(Consumer utility maximization) For each* h, $(x_h^\star, \varphi_h^\star, m_h^\star)$ *maximizes* U_h *on the budget set*

$$\left\{(x, \varphi, m) \geq 0 : \int_0^T p^\star(t)\, x(t)\, \mathrm{d}t + \varrho^\star \varphi + m \leq \hat{M}_h\left(p^\star, r_{\mathrm{St}}^\star, \varrho^\star\right)\right\}$$

where

$$\hat{M}_h(p, r_{\mathrm{St}}, \varrho) = m_h^{\mathrm{En}} + \varsigma_{h\,\mathrm{St}} \sup_{k_{\mathrm{St}}}\left(r_{\mathrm{St}} k_{\mathrm{St}} - G(k_{\mathrm{St}})\right) + \varsigma_{h\,\mathrm{IU}} \sup_{z,n}\left(\varrho F(z,n) - \int_0^T p(t)\, z(t)\, \mathrm{d}t - n\right). \tag{17.8}$$

(4) *(Market clearance)*

$$y_{\mathrm{PS}}^\star + \sum_\theta y_\theta^\star = z^\star + \sum_h x_h^\star \quad \text{and} \quad F(z^\star, n^\star) = \sum_h \varphi_h^\star. \tag{17.9}$$

(5) *(MC pricing of storage capacity)*

$$r_{\mathrm{St}}^\star \in \partial G\left(k_{\mathrm{St}}^\star\right). \tag{17.10}$$

Proof. Given the results of Section 16, this is a formality—except for verifying the absence of a duality gap. Note first that Conditions 2 to 5 of the theorem are simply specializations, to the ESI case, of the corresponding parts of the definition of a long-run equilibrium (Section 13). What has to be shown is the equivalence of the theorem's Condition 1 (optimal operation and valuation of the ESI's plants) to the definition's Condition 1 (LRP maximization). As a general principle, this has been established in

[87] Since F is taken to be $-\infty$ outside of $L_+^\infty \times \mathbb{R}_+$, $\widehat{\partial} F$ contains a term arising from this nonnegativity constraint. To spell this out, assume that F, as a function on its effective domain $L_+^\infty \times \mathbb{R}_+$, has a Mackey continuous, concave and Gateaux differentiable extension F^{Ex} defined on all of $L^\infty \times \mathbb{R}$. Then (17.7) means that $(z^\star, n^\star) \geq 0$ and $(1/\varrho^\star)\, p^\star = \nabla_z F^{\mathrm{Ex}}(z^\star, n^\star) + \mu$ and $1/\varrho^\star = \left(\partial F^{\mathrm{Ex}}/\partial n\right)(z^\star, n^\star) + \nu$ for some $\mu \in L_+^1$ vanishing a.e. on the set $\{t : z^\star(t) > 0\}$, with $\nu = 0$ if $n^\star > 0$. (If $p^\star$ were in $L^{\infty *}$ but not in L^1 then μ would be an element of $L_+^{\infty *}$ concentrated on $\{t : z^\star(t) \leq \epsilon\}$ for each $\epsilon > 0$.)

Section 4 and restated in Section 6 (by taking account of Section 5). Its substance is that, in the long run, competitive profit maximization is equivalent—as a system of conditions on both quantities and prices—to the conjunction of: (i) maximization of the operating profit (short-run profit), which includes minimization of the operating cost, (ii) minimization of the fixed-input value by shadow pricing (which is identified as the dual programme), and (iii) equality of the maximum SRP to the minimum FIV (absence of a duality gap). For each of the ESI's plants, the SRP and FIV programmes are spelt out in Section 16, and it remains only to show that their values are equal. (In formal terms, (13.4)–(13.5) is (3.5) at equilibrium prices, which, as is noted before the Comment in Section 6, is equivalent to the conjunction of (4.2)–(4.3), (6.2) and (6.3). And, for the ESI's technology, (4.2)–(4.3) and (6.2) can be put as (17.4)–(17.6) and (17.1)–(17.3). It remains only to prove (6.3) for each of the ESI's plants.)

To this end, note first that the thermal operation LP (16.1)–(16.3) and its dual (16.6)–(16.9) always have the same value: it is $k_\theta \int_0^T (p(t) - w_\theta)^+ \, \mathrm{d}t$ for each (p, k_θ, w_θ), by (16.4) and by (16.5) or (16.10). For pumped storage, however, the equality of values of the operation LP (16.12)–(16.16) and its dual—in the form of either the standard dual LP (16.17)–(16.22) or the equivalent CP (16.27)–(16.28)—relies on the properties of its data in the general equilibrium, $(p^\star; k^\star_{\mathrm{St}}, k^\star_{\mathrm{Co}})$. It can be proved in two ways because it follows from *either* of two assumptions: that $(k^\star_{\mathrm{St}}, k^\star_{\mathrm{Co}}) \gg 0$ and that $p^\star \in L^1[0,T]$. Strict positivity of the fixed-input bundle $(k_{\mathrm{St}}, k_{\mathrm{Co}})$ is a case of the generalized Slater's Condition for infinite-dimensional inequality constraints, formulated in [44, (8.12)]. *A fortiori*, it is a case of Slater's Condition for generalized perturbed CPs, formulated in [44, Theorem 18 (a)]. So it guarantees the continuity of $\Pi^{\mathrm{PS}}_{\mathrm{SR}}(p, \cdot)$ on a neighbourhood of $(k_{\mathrm{St}}, k_{\mathrm{Co}})$, for every $p \in L^{\infty *}$: see Part 1 of Lemma 23.1 for details. (The same argument applies more generally to c.f.c. techniques: see Proposition 27.2.) The other, alternative proof derives upper semicontinuity of $\Pi^{\mathrm{PS}}_{\mathrm{SR}}(p, \cdot)$ from the assumption that $p \in L^1$. This is a case of a price system in the predual of the commodity space: here, L^1 is the Banach predual of $L^\infty[0,T]$. The maximand $\langle p \,|\, \cdot \rangle$ is therefore continuous for the weak* topology $\mathrm{w}(L^\infty, L^1)$, and one can show that the maximum value, $\Pi^{\mathrm{PS}}_{\mathrm{SR}}(p, \cdot)$, is u.s.c. by exploiting the weak*-compactness of the short-run production set

$$\{y \in L^\infty[0,T] : (y; -k_{\mathrm{St}}, -k_{\mathrm{Co}}) \in \mathbb{Y}_{\mathrm{PS}}\} \subseteq \{y \in L^\infty : |y| \le k_{\mathrm{Co}}\}$$

where $\mathbb{Y}_{\mathrm{PS}}$ is given by (15.4); formally, Lemma 21.3 applies.[88] (A stronger result can be obtained by applying the dual-value continuity criterion of [44, Theorem 18' (e)]: this shows that the convex function $\overline{\Pi}^{\mathrm{PS}}_{\mathrm{SR}}(\cdot, k_{\mathrm{St}}, k_{\mathrm{Co}})$ is norm-continuous on L^1, which implies that the concave function $\Pi^{\mathrm{PS}}_{\mathrm{SR}}(p, \cdot)$ is upper semicontinuous at $(k_{\mathrm{St}}, k_{\mathrm{Co}})$ for each $p \in L^1$, by Lemma 20.1.) Finally, Lemma 20.1 shows that the equality $\overline{\Pi}^{\mathrm{PS}}_{\mathrm{SR}} = \Pi^{\mathrm{PS}}_{\mathrm{SR}}$ at $(p; k_{\mathrm{St}}, k_{\mathrm{Co}})$ follows from upper semicontinuity, and *a fortiori* from continuity, of $\Pi^{\mathrm{PS}}_{\mathrm{SR}}(p, \cdot)$ at $(k_{\mathrm{St}}, k_{\mathrm{Co}})$. Since $p^\star \in L^1$ and $(k^\star_{\mathrm{St}}, k^\star_{\mathrm{Co}}) \gg 0$, either method applies to this data point. □

[88]For this technology, the Proof of Lemma 21.3 simplifies to a direct application of Berge's Maximum Theorem [6, VI.3: Theorem 2]. This is because K is the finite-dimensional space $\mathbb{R}^2$, and because the set $\bigcup_{k \in B} \mathbb{Y}_{\mathrm{SR}}(k)$ is itself bounded when B is (i.e., the operation vmax is not needed). The same applies more generally to c.f.c. techniques: see Proposition 27.2.

We next present a similar result with hydroelectric generation (H) instead of pumped storage (PS). The thermal technology remains the same, and its inputs have fixed prices, $r^{\mathrm{F}}_{\mathrm{Th}} = \left(r^{\mathrm{F}}_1, \ldots, r^{\mathrm{F}}_{\Theta}\right)$ and $w = (w_1, \ldots, w_{\Theta})$. The hydro turbine also has a fixed price, $r^{\mathrm{F}}_{\mathrm{Tu}}$. There is a river with a single location where a dam can be constructed to create a water reservoir of a capacity k_{St}, at a cost $G(k_{\mathrm{St}})$. The river has a fixed, periodic flow, $e(t)$ at time $t \in [0, T]$, which (on the assumption of a constant head) means a given energy inflow.[89] Its price, $\psi(t)$ at time t, is determined in the long-run equilibrium.[90] The river's total rent is $\int_0^T \psi e \, \mathrm{d}t$, and household h's share of the rent is $\varsigma_{h\,\mathrm{Ri}}$. Its share of profit from supplying the storage capacity is $\varsigma_{h\,\mathrm{St}}$. As before, there is a single Industrial User of electricity (whose production function is F), and the household's share in his profit is $\varsigma_{h\,\mathrm{IU}}$.

Theorem 17.2 (Characterization of long-run equilibrium with hydro-thermal generation). *Assume that the ESI's technology consists of thermal generation techniques (Θ) and a hydroelectric technique. Then a price system made up of:*

- *a time-continuous electricity tariff $p^\star \in \mathcal{C}[0, T]$*
- *a rental price for the hydro reservoir capacity $r^\star_{\mathrm{St}}$*
- *a price $\varrho^\star$ for the produced final good*
- *the given prices for fuel and the generating capacities (viz., r^{F}_θ for thermal capacity of type θ and w_θ for its fuel, and $r^{\mathrm{F}}_{\mathrm{Tu}}$ for the turbine capacity)*

and an allocation made up of:

- *an output $y^\star_\theta \in L^\infty_+[0, T]$ from the thermal plant of type θ with*
  - *a capacity $k^\star_\theta > 0$*
  - *a fuel input $v^\star_\theta$ (for each θ)*
- *an output $y^\star_{\mathrm{H}} \in L^\infty[0, T]$ from a hydro plant with*
  - *reservoir and turbine capacities $k^\star_{\mathrm{St}} > 0$ and $k^\star_{\mathrm{Tu}} > 0$*
  - *the given river flow $e \in L^\infty_+[0, T]$, which is assumed to meet Condition (16.50)*[91]
- *a consumption bundle $(x^\star_h, \varphi^\star_h, m^\star_h) \in L^\infty_+[0, T] \times \mathbb{R}_+ \times \mathbb{R}_+$ for each household h*
- *an input-output bundle of the User Industry $(-z^\star, F(z^\star, n^\star), -n^\star) \in L^\infty_-[0, T] \times \mathbb{R}_+ \times \mathbb{R}_-$*

form a long-run competitive equilibrium if and only if:

(1) (a) *(Equality of ESI's capital-input prices to their profit-imputed marginal values) For each $\theta = 1, \ldots, \Theta$*

$$r^{\mathrm{F}}_\theta = \int_0^T \left(p^\star(t) - w_\theta\right)^+ \mathrm{d}t \tag{17.11}$$

$$r^\star_{\mathrm{St}} = \mathrm{Var}^+_{\mathrm{c}}(\psi^\star) \tag{17.12}$$

[89] More generally, it might be possible to improve the watershed to obtain a river flow e at a cost $G_{\mathrm{Ri}}(e)$, a convex function of e. The case of a fixed, unimprovable river flow $\overline{e}$ can be obtained by setting $G_{\mathrm{Ri}}(e)$ equal to 0 for $e = \overline{e}$ and $+\infty$ otherwise.

[90] A complete price system is $(p; r_{\mathrm{Th}}; w; r_{\mathrm{H}}, \psi; \varrho, 1)$ with $r_{\mathrm{H}} = (r_{\mathrm{St}}, r_{\mathrm{Tu}})$.

[91] The assumption can be dropped, but this complicates the problem and, as a result, an optimal water price function need not be unique or continuous: see [24].

$$r_{\mathrm{Tu}}^{\mathrm{F}}=\int_0^T\left(p^{\star}(t)-\psi^{\star}(t)\right)^{+}\mathrm{d}t \tag{17.13}$$

where $\psi^{\star}:=\hat{\psi}_{\mathrm{H}}\left(p^{\star};k_{\mathrm{St}}^{\star},k_{\mathrm{Tu}}^{\star};e\right)$ *is the optimal price of water, i.e., the unique solution to the programme (16.46)–(16.48) with* $\left(p^{\star};k_{\mathrm{St}}^{\star},k_{\mathrm{Tu}}^{\star};e\right)$ *as the data.*

(b) *(Operating profit maximization by ESI) For each* θ

$$y_{\theta}^{\star}(t)\in S\left(p^{\star}(t),k_{\theta}^{\star},w_{\theta}\right):=\begin{cases}\{0\} & \text{if } p^{\star}(t)<w_{\theta}\\ [0,k_{\theta}^{\star}] & \text{if } p^{\star}(t)=w_{\theta}\\ \{k_{\theta}^{\star}\} & \text{if } p^{\star}(t)>w_{\theta}\end{cases}\quad\text{for a.e. } t \tag{17.14}$$

$$v_{\theta}^{\star}=\int_0^T y_{\theta}^{\star}(t)\,\mathrm{d}t \tag{17.15}$$

and, with $\left(p^{\star};k_{\mathrm{St}}^{\star},k_{\mathrm{Tu}}^{\star};e\right)$ *as the data,*

$$y_{\mathrm{H}}^{\star}\text{ solves the linear programme (16.32) to (16.36)} \tag{17.16}$$

(which implies that $y_{\mathrm{H}}^{\star}(t)=k_{\mathrm{Tu}}$ *when* $p^{\star}(t)>\psi^{\star}(t)$ *and* $y_{\mathrm{H}}^{\star}(t)=0$ *when* $p^{\star}(t)<\psi^{\star}(t)$*).*

(2) *(Profit maximization by User Industry)*

$$\left(p^{\star},1\right)\in\varrho^{\star}\widehat{\partial}F\left(z^{\star},n^{\star}\right). \tag{17.17}$$

(3) *(Consumer utility maximization) For each* h, $\left(x_h^{\star},\varphi_h^{\star},m_h^{\star}\right)$ *maximizes* U_h *on the budget set*

$$\left\{(x,\varphi,m)\geq 0:\int_0^T p^{\star}(t)\,x(t)\,\mathrm{d}t+\varrho^{\star}\varphi+m\leq\hat{M}_h\left(p^{\star},r_{\mathrm{St}}^{\star},\psi^{\star},\varrho^{\star}\right)\right\}$$

where

$$\begin{aligned}\hat{M}_h\left(p,r_{\mathrm{St}},\varrho,\psi\right)&=m_h^{\mathrm{En}}+\varsigma_{h\,\mathrm{St}}\left(\sup_{k_{\mathrm{St}}}\left(r_{\mathrm{St}}k_{\mathrm{St}}-G\left(k_{\mathrm{St}}\right)\right)\right)\\&+\varsigma_{h\,\mathrm{IU}}\left(\sup_{z,n}\left(\varrho F(z,n)-\int_0^T p(t)\,z(t)\,\mathrm{d}t-n\right)\right)+\varsigma_{h\,\mathrm{Ri}}\int_0^T\psi(t)\,e(t)\,\mathrm{d}t.\end{aligned} \tag{17.18}$$

(4) *(Market clearance)*

$$y_{\mathrm{H}}^{\star}+\sum_{\theta}y_{\theta}^{\star}=z^{\star}+\sum_h x_h^{\star}\quad\text{and}\quad F\left(z^{\star},n^{\star}\right)=\sum_h\varphi_h^{\star}. \tag{17.19}$$

(5) *(MC pricing of reservoir capacity)*

$$r_{\mathrm{St}}^{\star}\in\partial G\left(k_{\mathrm{St}}^{\star}\right). \tag{17.20}$$

Proof. This is proved like Theorem 17.1 (taking into account the last Comment in Section 13). □

Remark 17.3 (Value of site for reservoir)**.** *The rental value of the hydro or storage site is* $r_{\mathrm{St}}^{*}k_{\mathrm{St}}^{*}-G\left(k_{\mathrm{St}}^{*}\right)$ *per cycle (the reservoir's value less its construction cost).*

Comments (multiple sites): A similar analysis applies when there is a number of storage sites (or hydro sites) with different development costs, G_l for location l. Reservoir capacity is then a good differentiated by its location, and so is the river flow in the case of hydro. Therefore, some of the long-run equilibrium prices and quantities may depend on l:

(1) Consider first the case of pumped storage. Since $\left(\partial \Pi_{\mathrm{SR}}^{\mathrm{PS}} / \partial k_{\mathrm{Co}}\right)\left(k_{\mathrm{St},l}^{\star}, k_{\mathrm{Co},l}^{\star}\right)$ equals $r_{\mathrm{Co}}^{\mathrm{F}}$, which is independent of l, and since the derivative is homogeneous of degree 0 in $(k_{\mathrm{St}}, k_{\mathrm{Co}})$, the equilibrium capacity ratio $k_{\mathrm{St},l}^{\star} : k_{\mathrm{Co},l}^{\star}$ is independent of l. Therefore, the equilibrium price of storage capacity $r_{\mathrm{St}}^{\star}$ is also the same for each l (since it equals $\partial \Pi_{\mathrm{SR}}^{\mathrm{PS}} / \partial k_{\mathrm{St}}$, which is also homogeneous of degree 0). And this is so because the production technique has just one input whose supply cost depends on the location. However, the plant's size depends on l, since $k_{\mathrm{St},l}^{\star}$ meets the condition $r_{\mathrm{St}}^{\star} \in \partial G_l\left(k_{\mathrm{St},l}^{\star}\right)$. The site's rent, $r_{\mathrm{St}}^{*} k_{\mathrm{St},l}^{\star} - G_l\left(k_{\mathrm{St},l}^{\star}\right)$, also depends on l.

(2) In hydro generation, both the reservoir construction cost function G_l and the fixed river flow e_l depend on the location l. So, in hydro, the equilibrium capacity ratio $k_{\mathrm{St},l}^{\star} / k_{\mathrm{Tu},l}^{\star}$, the price of reservoir capacity $r_{\mathrm{St},l}^{\star}$ and the shadow price of water $\psi_l^{\star}$ do all depend on l. (So do the reservoir's size $k_{\mathrm{St},l}^{\star}$ and the site's rent $r_{\mathrm{St}}^{\star} k_{\mathrm{St},l}^{\star} - G_l\left(k_{\mathrm{St},l}^{\star}\right)$.)

Comment (optimum of thermal output in terms of SRMC): Competitive profit maximization by the thermal plants can be reformulated as SRMC pricing by the thermal generating system, i.e., by using the system's instantaneous SRMC curve. With a finite number of plant types, Θ, this curve is actually a "right-angled" broken line:[92] it consists of (i) Θ "horizontal" segments

$$[k_1 + \ldots + k_{\theta-1}, k_1 + \ldots + k_\theta] \times \{w_\theta\} \quad \text{for } \theta = 1, \ldots, \Theta$$

(with $k_0 := 0$) and (ii) $\Theta + 1$ "vertical" segments

$$\{k_1 + \ldots + k_\theta\} \times [w_\theta, w_{\theta+1}] \quad \text{for } \theta = 0, 1, \ldots, \Theta$$

(with $w_{\Theta+1} := +\infty$, and with $w_0 := -\infty$ unless free disposal is included). Formally, Condition (17.4) or (17.14) for each θ is equivalent to:

$$p^{\star}(t) \in \partial_{\mathsf{y}} c_{\mathrm{SR}}\left(y^{\star}(t), k_{\mathrm{Th}}^{\star}, w\right) \quad \text{for a.e. } t$$

where $k_{\mathrm{Th}}^{\star} = (k_1^{\star}, \ldots, k_{\Theta}^{\star})$ and c_{SR} is the instantaneous short-run cost per unit time. With 1_A denoting the 0-1 indicator of a set A (1 on A and 0 outside), the instantaneous SRC can be given as

$$c_{\mathrm{SR}}(\mathsf{y}, k_{\mathrm{Th}}, w) := \int_0^{\mathsf{y}} \sum_{\theta=1}^{\Theta} w_\theta 1_{[k_1+\ldots+k_{\theta-1}, k_1+\ldots+k_\theta]}(\mathsf{q})\, \mathrm{d}\mathsf{q} \tag{17.21}$$

$$= w_1 \mathsf{y} + \sum_{\theta=1}^{\Theta-1} (w_{\theta+1} - w_\theta)\left(\mathsf{y} - (k_1 + \ldots + k_\theta)\right)^{+}$$

[92] In a model with a "continuum" of plant types, the SRMC curve is a general "complete nondecreasing curve", in the terminology of [42, 24.3]. But even the continuum model does *not* make the SRC curve differentiable: it still has a kink at the peak output, and typically it has offpeak kinks, too—see [24].

if $0 \le \mathsf{y} \le \sum_{\theta=1}^{\Theta} k_\theta$ (otherwise $c_{\mathrm{SR}} = +\infty$). This is an increasing and convex (though piecewise linear) function of $\mathsf{y} \in \left[0, \sum_{\theta=1}^{\Theta} k_\theta\right]$, with $c_{\mathrm{SR}}(0) = 0$. The SRMC curve is the graph of the subdifferential correspondence $\mathsf{y} \mapsto \partial c_{\mathrm{SR}}(\mathsf{y})$, in the instantaneous quantity-price plane. When $k_{\mathrm{Th}} \gg 0$ (i.e., $k_\theta > 0$ for each θ),

$$\partial_{\mathsf{y}} c_{\mathrm{SR}}(\mathsf{y}, k_{\mathrm{Th}}, w) = \begin{cases} (-\infty, w_1] & \text{if } \mathsf{y} = 0 \\ \{w_\theta\} & \text{if } \mathsf{y} \in (k_1 + \ldots + k_{\theta-1}, k_1 + \ldots + k_\theta) \\ [w_\theta, w_{\theta+1}] & \text{if } \mathsf{y} = k_1 + \ldots + k_\theta \text{ and } 1 \le \theta \le \Theta - 1 \\ [w_\Theta, +\infty) & \text{if } \mathsf{y} = \sum_{\theta=1}^{\Theta} k_\theta \\ \emptyset & \text{if } \mathsf{y} > \sum_{\theta=1}^{\Theta} k_\theta \text{ or } \mathsf{y} < 0 \end{cases} . \tag{17.22}$$

(For the case of $\Theta = 1$, the SRMC and SRC curves have been used in Section 2 and are shown in Figures 1a and c; the supply and subdifferential correspondences, $\mathsf{p} \mapsto S(\mathsf{p})$ and $\mathsf{y} \mapsto \partial c_{\mathrm{SR}}(\mathsf{y})$, are inverse to each other.)

18. Derivation of the dual programmes (proofs for Section 5)

The dual programmes are next derived formally by using the framework of [44].

Proposition 18.1 (Dual to SRP programme). *The dual of the short-run profit maximization programme (3.6)–(3.7), with k as the primal parameter ranging over the space K paired with R as the range for the dual variable r, is the fixed-input shadow-pricing programme (5.6), or equivalently (5.13)–(5.14) when $\mathbb{Y}$ is a cone. The dual parameter is (p, w).*

Proof. Given (p, k, w), the parametric primal constrained maximand is $\langle p \mid y \rangle - \langle w \mid v \rangle$ minus $\delta(y, -k, -v \mid \mathbb{Y})$, where y and v are the primal decision variables, and k is the primal parameter (paired with the dual decision variable r). Let d' and d'' denote the dual perturbations (paired with y and $-v$). By [44, (4.17)] with the primal problem reoriented to maximization, the (perturbed) dual constrained minimand is, as a function of r and (d', d''),

$$\begin{aligned} &\sup_{y,v;\Delta k} \{\langle d', d'' \mid y, -v\rangle - \langle r \mid \Delta k\rangle + \langle p \mid y\rangle - \langle w \mid v\rangle - \delta(y, -k - \Delta k, -v \mid \mathbb{Y})\} \\ &\qquad = \langle r \mid k\rangle + \sup_{y,v,\Delta k} \{\langle p + d', r, w + d'' \mid y, -k - \Delta k, -v\rangle : (y, -k - \Delta k, -v) \in \mathbb{Y}\} \\ &\qquad = \langle r \mid k\rangle + \sup_{y,v,k} \{\langle p + d', r, w + d'' \mid y, -k, -v\rangle : (y, -k, -v) \in \mathbb{Y}\} \\ &\qquad = \langle r \mid k\rangle + \Pi_{\mathrm{LR}}(p + d', r, w + d'') . \end{aligned}$$

So, by setting $d' = 0$ and $d'' = 0$, the dual programme is (5.6); and when $\mathbb{Y}$ is a cone, the dual is to minimize $\langle r \mid k\rangle + \delta(p, r, w \mid \mathbb{Y}^\circ)$ over r (since $\Pi_{\mathrm{LR}} = \delta^{\#}(\cdot \mid \mathbb{Y}) = \delta(\cdot \mid \mathbb{Y}^\circ)$, i.e., the support function of a cone is the indicator function of the polar cone). Finally, d' and d'' perturb the dual like increments to p and w, which therefore are the dual parameters (and so d' and d'' may be renamed to Δp and Δw). $\square$

The other duals are derived in the same way; the dual of the SRC programme is spelt out.

Proposition 18.2 (Dual to SRC programme)**.** *The dual of the short-run cost minimization programme (3.10)–(3.11), with* (y,k) *as the primal parameter ranging over the space* $Y \times K$ *paired with* $P \times R$ *as the range for the dual variable* (p,r)*, is the output-and-fixed-input pricing programme (5.4), or equivalently (5.9)–(5.10) when* $\mathbb{Y}$ *is a cone. The dual parameter is* w*.*

19. Shephard-Hotelling Lemmas and their dual counterparts (expansion and proofs for Sections 4, 6, 7, 8 and 11)

Programme solutions can always be characterized as marginal values of Type Two, i.e., the primal solution set is equal to the subdifferential of the *primal* optimal value as a (convex) function of the *dual* parameter. Likewise, the dual solution set is equal to the superdifferential of the *dual* optimal value as a (concave) function of the *primal* parameter (when the primal is to maximize, and the dual is to minimize). This derivative property is next stated for the profit or cost optimization programmes and their duals. All six results are obtained by applying either the Inversion Rule (Theorem C.2) and the FOC (C.12) or (C.24) or the Derivative Property of the Conjugate (C.33), which combines the Inversion Rule and the FOC. The same techniques apply to the reduced programmes: see the end of this section.

Notation: As a superscript, the symbol # indicates the Fenchel-Legendre *convex conjugate* (of a convex function), defined by (C.1) in Appendix C. As a subscript, # indicates the *concave conjugate* (of a concave function), defined by (C.20). In either position, # means the *total conjugate*, i.e., the conjugate w.r.t. all of the function's arguments (except for those indicated as fixed). Partial conjugates w.r.t. one variable (say, the first or the second variable) are denoted by $\#_1$, $\#_2$, , etc.; these are defined by (C.6). The partial conjugate w.r.t. the first and second variables *together* is denoted by $\#_{1,2}$ (for a bivariate function, this means the same as #).

Lemma 19.1 (Hotelling's Lemma for the short run)**.** *Assume that* $\mathbb{Y}$ *is closed. Then* $(y,-v) \in \partial_{p,w}\Pi_{\mathrm{SR}}(p,k,w)$ *if and only if* (y,v) *solves the short-run profit maximization programme (3.6)–(3.7).*

Proof. By definition, $\Pi_{\mathrm{SR}}(\cdot,k,\cdot)$ is $\delta^{\#}(\cdot \mid \mathbb{Y}_{\mathrm{SR}}(k))$, i.e., it is the support function of the section of $\mathbb{Y}$ through $-k$. This is a closed convex subset of $Y \times V$; so if it is also nonempty, then

$$\partial_{p,w}\Pi_{\mathrm{SR}}(p,k,w) = \{(y,-v) \in \mathbb{Y}_{\mathrm{SR}}(k) : \langle p \,|\, y\rangle - \langle w \,|\, v\rangle = \Pi_{\mathrm{SR}}(p,k,w)\}$$

by (C.36). Even in the degenerate case of $\mathbb{Y}_{\mathrm{SR}}(k) = \emptyset$, the subdifferential and the solution set are equal: each is $Y \times V$ (since every vector is then a subgradient of $\Pi(\cdot,k,\cdot) = -\infty$, and since every point solves, albeit improperly,[93] the then infeasible programme (3.6)–(3.7)). □

Lemma 19.2 (Dual of SR Hotelling's Lemma)**.** $r \in \widehat{\partial}_k \overline{\Pi}_{\mathrm{SR}}(p,k,w)$ *if and only if* r *solves the fixed-input pricing programme (5.6).*

[93] See the Comment on proper and improper solutions in Appendix C, after (C.36)–(C.38).

Proof. By the definition of $\overline{\Pi}_{\mathrm{SR}}$ as the optimal value of (5.6), and by (3.14),

$$\overline{\Pi}_{\mathrm{SR}} = \left(-\Pi_{\mathrm{LR}}\right)_{\#_2} \text{ and } -\Pi_{\mathrm{LR}} = \Pi_{\mathrm{SR}\#_2} \tag{19.1}$$

(in other words, $\overline{\Pi}_{\mathrm{SR}}(p,\cdot,w) = \left(-\Pi_{\mathrm{LR}}(p,\cdot,w)\right)_{\#}$ and $-\Pi_{\mathrm{LR}}(p,\cdot,w) = \Pi_{\mathrm{SR}}(p,\cdot,w)_{\#}$). From the second equality of (19.1), $\left(-\Pi_{\mathrm{LR}}\right)_{\#_2\#_2} = -\Pi_{\mathrm{LR}}$ (i.e., $\left(-\Pi_{\mathrm{LR}}(p,\cdot,w)\right)_{\#\#} = -\Pi_{\mathrm{LR}}(p,\cdot,w)$) by using (C.22). This and the first equality of (19.1) mean that the Inversion Rule (C.32) can be applied (to $\Pi_{\mathrm{LR}}(p,\cdot,w)$ in place of Π) to give

$$r \in \widehat{\partial}_k \overline{\Pi}_{\mathrm{SR}}(p,k,w) \Leftrightarrow -k \in \partial_r \Pi_{\mathrm{LR}}(p,r,w)$$
$$\Leftrightarrow r \text{ minimizes } \Pi_{\mathrm{LR}}(p,\cdot,w) + \langle \cdot \mid k \rangle$$

by the FOC (C.12). Alternatively, apply the Derivative Property (C.33) to conflate the two steps. □

Alternative Proof of Lemma 19.2 (under c.r.t.s.) If $\mathbb{Y}$ is a cone, this can be proved like Lemma 19.1: $\overline{\Pi}_{\mathrm{SR}}(p,\cdot,w)$ is then the inf-support function of the polar cone's section

$$\mathbb{Y}^{\circ}_{p,w} := \{r : (p,r,w) \in \mathbb{Y}^{\circ}\} \tag{19.2}$$

and so (C.38) applies. □

Lemma 19.3 (Shephard's Lemma for the short run). *Assume that $\mathbb{Y}$ is closed. Then $(k,v) \in \widehat{\partial}_{r,w} C_{\mathrm{LR}}(y,r,w)$ if and only if (k,v) solves the long-run cost minimization programme (3.8)–(3.9).*

Proof. Like Lemma 19.1, this is a case of differentiating a support function: $C_{\mathrm{LR}}(y,\cdot,\cdot)$ is the inf-support function of the section of $\mathbb{Y}$ through y, so (C.38) applies. □

Lemma 19.4 (Dual of SR Shephard's Lemma). *$p \in \partial_y \underline{C}_{\mathrm{LR}}(y,r,w)$ if and only if p solves the long-run output-pricing programme (5.5).*

Proof. Like Lemma 19.2, this follows from the definitional relationships between Π_{LR}, the value function being differentiated and the value dual to it—i.e., from

$$\underline{C}_{\mathrm{LR}} = \Pi_{\mathrm{LR}}^{\#_1} \text{ and } \Pi_{\mathrm{LR}} = C_{\mathrm{LR}}^{\#_1} \tag{19.3}$$

(in other words, from $\underline{C}_{\mathrm{LR}}(\cdot,r,w) = \Pi_{\mathrm{LR}}(\cdot,r,w)^{\#}$ and $\Pi_{\mathrm{LR}}(\cdot,r,w) = C_{\mathrm{LR}}(\cdot,r,w)^{\#}$) by applying the Inversion Rule (C.31) and the FOC (C.12). Alternatively, apply the Derivative Property (C.33) to conflate the two steps. □

Alternative Proof of Lemma 19.4 (under c.r.t.s.) If $\mathbb{Y}$ is a cone then $\underline{C}_{\mathrm{LR}}(\cdot,r,w)$ is the support function of $\{p : (p,r,w) \in \mathbb{Y}^{\circ}\}$, and so (C.36) applies. □

Lemma 19.5 (Shephard's Lemma for the long run). *Assume that $\mathbb{Y}$ is closed. Then $v \in \widehat{\partial}_w C_{\mathrm{SR}}(y,k,w)$ if and only if v solves the short-run cost minimization programme (3.10)–(3.11).*

Proof. Like Lemmas 19.1 and 19.3, this is a case of differentiating a support function: here, $C_{\mathrm{SR}}(y,k,\cdot)$ is the inf-support function of the section of $\mathbb{Y}$ through $(y,-k)$, so (C.38) applies. □

Lemma 19.6 (Dual of LR Shephard's Lemma). *$(p,-r) \in \partial_{y,k} \underline{C}_{\mathrm{SR}}(y,k,w)$ if and only if (p,r) solves the output-and-fixed-input pricing programme (5.4).*

Proof. Like Lemmas 19.2 and 19.4, this follows from the definitional relationships between Π_{LR}, the value function being differentiated and the value dual to it—i.e., from

$$\underline{C}_{\mathrm{SR}}(y,k,w)=\Pi_{\mathrm{LR}}^{\#_{1,2}}(y,-k,w) \text{ and } \Pi_{\mathrm{LR}}(p,r,w)=C_{\mathrm{SR}}^{\#_{1,2}}(p,-r,w) \tag{19.4}$$

by applying the Inversion Rule (C.31) and the FOC (C.12). Alternatively, apply the Derivative Property (C.33) to conflate the two steps. □

Alternative Proof of Lemma 19.6 (under c.r.t.s.) If $\mathbb{Y}$ is a cone then $\underline{C}_{\mathrm{SR}}(\cdot,\cdot,w)$ is the support function of $\{(p,-r):(p,r,w)\in\mathbb{Y}^{\circ}\}$, and so (C.36) applies. □

The two marginal values of Type Two are actually of Type One if (and only if) there is no duality gap. This is next applied (thrice) to complement the preceding six lemmas.

Remark 19.7. *$(y,-v)\in\partial_{p,w}\overline{\Pi}_{\mathrm{SR}}(p,k,w)$ if and only if $(y,-v)\in\partial_{p,w}\Pi_{\mathrm{SR}}(p,k,w)$ and $\overline{\Pi}_{\mathrm{SR}}(p,k,w)=\Pi_{\mathrm{SR}}(p,k,w)$.*

Remark 19.8. *$r\in\widehat{\partial}_k\Pi_{\mathrm{SR}}(p,k,w)$ if and only if $r\in\widehat{\partial}_k\overline{\Pi}_{\mathrm{SR}}(p,k,w)$ and $\overline{\Pi}_{\mathrm{SR}}(p,k,w)=\Pi_{\mathrm{SR}}(p,k,w)$.*

Remark 19.9. *$(k,v)\in\widehat{\partial}_{r,w}\underline{C}_{\mathrm{LR}}(y,r,w)$ if and only if $(k,v)\in\widehat{\partial}_{r,w}C_{\mathrm{LR}}(y,r,w)$ and $\underline{C}_{\mathrm{LR}}(y,r,w)=C_{\mathrm{LR}}(y,r,w)$.*

Remark 19.10. *$p\in\partial_y C_{\mathrm{LR}}(y,r,w)$ if and only if $p\in\partial_y\underline{C}_{\mathrm{LR}}(y,r,w)$ and $\underline{C}_{\mathrm{LR}}(y,r,w)=C_{\mathrm{LR}}(y,r,w)$.*

Remark 19.11. *$v\in\widehat{\partial}_w\underline{C}_{\mathrm{SR}}(y,k,w)$ if and only if $v\in\widehat{\partial}_w C_{\mathrm{SR}}(y,k,w)$ and $\underline{C}_{\mathrm{SR}}(y,k,w)=C_{\mathrm{SR}}(y,k,w)$.*

Remark 19.12. *$(p,-r)\in\partial_{y,k}C_{\mathrm{SR}}(y,k,w)$ if and only if $(p,-r)\in\partial_{y,k}\underline{C}_{\mathrm{SR}}(y,k,w)$ and $\underline{C}_{\mathrm{SR}}(y,k,w)=C_{\mathrm{SR}}(y,k,w)$.*

Since the primal and dual values are assumed to be equal only *at* a particular data point (and not on a whole neighbourhood of it), Remarks 19.7–19.12 do require a proof. This can be based on (C.18), which is a criterion of subdifferentiability in terms of the function's second conjugate. It applies because the dual value (in this context, and under c.r.t.s., the imputed value of the given quantities) is the second conjugate of the primal value (profit or cost) as a function of the primal parameters (the quantity data). Likewise, the primal value is the second conjugate of the dual value as a function of the dual parameters (the price data). For example, $\overline{\Pi}_{\mathrm{SR}}$ is the second concave conjugate of Π_{SR} as a function of k, with (p,w) fixed. Dually, Π_{SR} is the second convex conjugate of $\overline{\Pi}_{\mathrm{SR}}$ as a function (p,w), with k fixed. (Similarly, $\underline{C}_{\mathrm{LR}}$ and $\underline{C}_{\mathrm{SR}}$ are the second convex conjugates of C_{LR} and C_{SR} as functions of, respectively, y and (y,k), with (r,w) or w fixed. Dually, C_{LR} and C_{SR} are the second concave conjugates of $\underline{C}_{\mathrm{LR}}$ and $\underline{C}_{\mathrm{SR}}$ as functions of, respectively, (r,w) and w, with y or (y,k) fixed.) These bi-conjugacy relationships are next recorded for use in proving Remarks 19.7–19.12.

Lemma 19.13. *If $\mathbb{Y}$ is closed then $\Pi_{\mathrm{SR}}=\overline{\Pi}_{\mathrm{SR}}^{\#_{1,3}\#_{1,3}}$ (i.e., $\Pi_{\mathrm{SR}}(\cdot,k,\cdot)=\overline{\Pi}_{\mathrm{SR}}(\cdot,k,\cdot)^{\#\#}$ on $Y\times W$ for each $k\in K$).*

Proof. Since $\Pi_{\mathrm{SR}}(\cdot,k,\cdot)$ is by definition the conjugate of $\delta(\cdot\mid\mathbb{Y}_{\mathrm{SR}}(k))$, it suffices to show that this is, in turn, the conjugate of $\overline{\Pi}_{\mathrm{SR}}(\cdot,k,\cdot)$. Since $\mathbb{Y}$ is closed (and convex), the definitional relationship $\Pi_{\mathrm{LR}} := \delta^{\#}(\cdot \mid \mathbb{Y})$ can be inverted to give

$$\begin{aligned}(19.5)\qquad \delta\left(y,-v \mid \mathbb{Y}_{\mathrm{SR}}(k)\right) &:= \delta\left(y,-k,-v \mid \mathbb{Y}\right) = \Pi^{\#}_{\mathrm{LR}}(y,-k,-v) \\ &:= \sup_{p,r,w}\left(\langle p \,|\, y\rangle - \langle r \,|\, k\rangle - \langle w \,|\, v\rangle - \Pi_{\mathrm{LR}}(p,r,w)\right) \\ &= \sup_{p,w}\Big(\langle p \,|\, y\rangle - \langle w \,|\, v\rangle - \overline{\overline{\Pi}}_{\mathrm{SR}}(p,k,w)\Big)\end{aligned}$$

since $\overline{\overline{\Pi}}_{\mathrm{SR}}$ is the optimal value of (5.6). □

Lemma 19.14. *$\overline{\overline{\Pi}}_{\mathrm{SR}} = \Pi_{\mathrm{SR}\#_2\#_2}$ (i.e., $\overline{\overline{\Pi}}_{\mathrm{SR}}(p,\cdot,w) = \Pi_{\mathrm{SR}}(p,\cdot,w)_{\#\#}$ on K for each $(p,w) \in P \times W$).*

Proof. Combine the definitional relationships (19.1) between Π_{LR} and each of the two functions ($\overline{\overline{\Pi}}_{\mathrm{SR}}$ and Π_{SR}). □

Lemma 19.15. *If $\mathbb{Y}$ is closed then $C_{\mathrm{LR}} = \underline{C}_{\mathrm{LR}}(y,\cdot,\cdot)_{\#_{2,3}\#_{2,3}}$ (i.e., $C_{\mathrm{LR}}(y,\cdot,\cdot) = \underline{C}_{\mathrm{LR}}(y,\cdot,\cdot)_{\#\#}$ on $R \times W$ for each $y \in Y$).*

Proof. Like Lemma 19.13, this is proved by iterating conjugacy (using the second equality of (19.5) in the process). □

Lemma 19.16. *$\underline{C}_{\mathrm{LR}} = C^{\#_1\#_1}_{\mathrm{LR}}$ (i.e., $\underline{C}_{\mathrm{LR}}(\cdot,r,w) = C_{\mathrm{LR}}(\cdot,r,w)^{\#\#}$ on Y for each $(r,w) \in R \times W$).*

Proof. Like Lemma 19.14, this follows from the definitional relationships (19.3) between Π_{LR} and each of the two functions (here, $\underline{C}_{\mathrm{LR}}$ and C_{LR}). □

Lemma 19.17. *If $\mathbb{Y}$ is closed then $C_{\mathrm{SR}} = \underline{C}_{\mathrm{SR}\#_3\#_3}$ (i.e., $C_{\mathrm{SR}}(y,k,\cdot) = \underline{C}_{\mathrm{SR}}(y,k,\cdot)_{\#\#}$ on W for each $(y,k) \in Y \times K$).*

Proof. Like Lemmas 19.13 and 19.15, this can be proved by iterating conjugacy (using the second equality of (19.5) in the process). □

Lemma 19.18. *$\underline{C}_{\mathrm{SR}} = C^{\#_{1,2}\#_{1,2}}_{\mathrm{SR}}$ (i.e., $\underline{C}_{\mathrm{SR}}(\cdot,\cdot,w) = C_{\mathrm{SR}}(\cdot,\cdot,w)^{\#\#}$ on $Y \times K$ for each $w \in W$).*

Proof. Like Lemmas 19.14 and 19.16, this follows from the definitional relationships (19.4) between Π_{LR} and each of the two functions (here, $\underline{C}_{\mathrm{SR}}$ and C_{SR}). □

Remarks 19.7–19.12 can now be deduced (all in the same way).

Proof of Remark 19.7. Fix any k, and abbreviate $\Pi_{\mathrm{SR}}(\cdot,k,\cdot)$ to Π. Then $\Pi = \overline{\overline{\Pi}}^{\#\#}$ (on $P \times W$) by Lemma 19.13. So, for each (p,w),

$$(y,-v) \in \partial\overline{\overline{\Pi}}(p,w) \Leftrightarrow \Big((y,-v) \in \partial\Pi(p,w) \text{ and } \Pi(p,w) = \overline{\overline{\Pi}}(p,w)\Big)$$

by (C.18). □

Proof of Remark 19.8. Fix any (p,w), and abbreviate $\Pi_{\mathrm{SR}}(p,\cdot,w)$ to Π. Then $\overline{\overline{\Pi}} = \Pi_{\#\#}$ (on K) by Lemma 19.14. So, for each k,

$$r \in \widehat{\partial}\Pi(k) \Leftrightarrow \left(r \in \widehat{\partial}\overline{\overline{\Pi}}(k) \text{ and } \overline{\overline{\Pi}}(k) = \Pi(k)\right)$$

by (C.26). □

Proof of Remark 19.9. Fix any y, and abbreviate $C_{\mathrm{LR}}(y,\cdot,\cdot)$ to C. Then $C = \underline{C}_{\#\#}$ (on $R \times W$) by Lemma 19.15. So, for each (r,w),

$$(k,v) \in \widehat{\partial}\underline{C}(r,w) \Leftrightarrow \Big((k,v) \in \widehat{\partial}C(r,w) \text{ and } C(r,w) = \underline{C}(r,w)\Big)$$

by (C.26) □

Proof of Remark 19.10. Fix any (r,w), and abbreviate $C_{\mathrm{LR}}(\cdot,r,w)$ to C. Then $\underline{C} = C^{\#\#}$ (on Y) by Lemma 19.16. So, for each y,

$$p \in \partial C(y) \Leftrightarrow (p \in \partial\underline{C}(y) \text{ and } \underline{C}(y) = C(y))$$

by (C.18). □

Proof of Remark 19.11. Fix any (y,k), and abbreviate $C_{\mathrm{SR}}(y,k,\cdot)$ to C. Then $C = \underline{C}_{\#\#}$ (on W) by Lemma 19.17. So, for each w,

$$v \in \widehat{\partial}\underline{C}(w) \Leftrightarrow (v \in \partial\underline{C}(y,k) \text{ and } \underline{C}(v) = C(v))$$

by (C.26). □

Proof of Remark 19.12. Fix any w, and abbreviate $C_{\mathrm{SR}}(\cdot,\cdot,w)$ to C. Then $\underline{C} = C^{\#\#}$ (on $Y \times K$) by Lemma 19.18. So, for each (y,k),

$$(p,-r) \in \partial C(y,k) \Leftrightarrow ((p,-r) \in \partial\underline{C}(y,k) \text{ and } \underline{C}(y,k) = C(y,k))$$

by (C.18). □

When there is no duality gap, programme solutions are therefore equal to marginal values of Type One: the dual solution is then equal to the primal value's derivative w.r.t. the primal parameter, and, similarly, the primal solution is the dual value's derivative w.r.t. the dual parameter. A pair of solutions with *equal* values is therefore the same as a pair of sub/super-gradients, w.r.t. primal and dual parameters, of just one of the two value functions (either primal or dual). Here, this means that Π and C can replace $\overline{\Pi}$ and $\underline{C}$ in Lemmas 19.2, 19.4 and 19.6—which can then be combined with Lemmas 19.1, 19.3 and 19.5 (respectively) to form subdifferential systems purely in terms of either Π or C (i.e., without $\overline{\Pi}$ or $\underline{C}$). Similarly, $\overline{\Pi}$ and $\underline{C}$ can replace Π and C in Lemmas 19.1, 19.3 and 19.5—which can then be combined with Lemmas 19.2, 19.4 and 19.6 to form subdifferential systems purely in terms of either $\overline{\Pi}$ or $\underline{C}$ (i.e., without Π or C). This is next stated formally.

Corollary 19.19 (Equivalence of subdifferential and solution systems)**.** *Assume that* $\mathbb{Y}$ *is closed. Then:*

1. *The following are equivalent to one another:*
   - (a) *the SRP subdifferential system (8.4)–(8.5)*
   - (b) *the SRP optimization system (6.1)–(6.3)*
   - (c) *the FIV subdifferential system (8.6)–(8.7).*
2. *The following are equivalent to one another:*
   - (a) *the LRC subdifferential system (8.8)–(8.9)*
   - (b) *the LRC optimization system (6.4)–(6.6)*
   - (c) *the OV subdifferential system (8.10)–(8.11).*
3. *The following are equivalent to one another:*

(a) *the SRC subdifferential system (8.12)–(8.13)*
(b) *the SRC optimization system (6.7)–(6.9)*
(c) *the OFIV subdifferential system (8.14)–(8.15).*

Therefore, each of these systems fully characterizes a long-run producer optimum, i.e., is equivalent to (3.3).

Proof. For Part 1, to prove that 1b is equivalent to 1a: (i) use Lemma 19.1, and (ii) combine Lemma 19.2 with Remark 19.8. To prove that 1b is equivalent to 1c: (i) use Lemma 19.2, and (ii) combine Lemma 19.1 with Remark 19.7.

For Part 2, to prove that 2b is equivalent to 2a: (i) use Lemma 19.3, and (ii) combine Lemma 19.4 with Remark 19.10. To prove that 2b is equivalent to 2c: (i) use Lemma 19.4, and (ii) combine Lemma 19.3 with Remark 19.9.

For Part 3, to prove that 3b is equivalent to 3a: (i) use Lemma 19.5, and (ii) combine Lemma 19.6 with Remark 19.12. To prove that 3b is equivalent to 3c: (i) use Lemma 19.6, and (ii) combine Lemma 19.5 with Remark 19.11.

Finally, each of the three optimization systems (1b, 2b, 3b) is equivalent to (3.3), as has been noted in Section 6. □

The same derivative properties of cost and profit functions, and the FOCs, serve to transform split optimization systems into their saddle differential equivalents.

Corollary 19.20 (Equivalence of saddle differential and solution systems)**.** *Assume that* $\mathbb{Y}$ *is closed. Then:*

(1) *The following are equivalent to one another:*
  (a) *the SRC-P saddle differential system (7.1)–(7.3)*
  (b) *the split SRP optimization system (4.2)–(4.5)*
  (c) *the SRP optimization system (6.1)–(6.3).*
(2) *The following are equivalent to one another:*
  (a) *the L-SRC saddle differential system (11.8)–(11.10)*
  (b) *the split LRC optimization system (11.11)–(11.14)*
  (c) *the LRC optimization system (6.4)–(6.6).*
(3) *The following are equivalent to one another:*
  (a) *the FIV saddle differential system (8.1)–(8.3)*
  (b) *the split SRC optimization system (6.8)–(6.11)*
  (c) *the SRC optimization system (6.7)–(6.9).*
(4) *The following are equivalent to one another:*
  (a) *the OV saddle differential system (11.15)–(11.17)*
  (b) *the reverse-split SRC optimization system (11.18)–(11.21)*
  (c) *the SRC optimization system (6.7)–(6.9).*

Therefore, each of these systems fully characterizes a long-run producer optimum, i.e., is equivalent to (3.3).

Proof. First note that, in all four parts, it is obvious that the optimization system is equivalent to the split optimization system (1b to 1c, 2b to 2c, 3b to 3c, and 3b to 3c): this is two-stage solving.

Next, for Part 1, to prove that 1b is equivalent to 1a: (i) use Lemma 19.5, (ii) combine Lemma 19.2 with Remark 19.8, and (iii) apply the FOC (C.12) to C_{SR} (as a function of y).

For Part 2, to prove that 2b is equivalent to 2a: (i) use Lemma 19.5, (ii) combine Lemma 19.4 with Remark 19.10, and (iii) apply the FOC (C.12) to C_{SR} (as a function of k).

For Part 3, to prove that 3b is equivalent to 3a: (i) use Lemma 19.2, (ii) combine Lemma 19.5 with Remark 19.11, and (iii) apply the FOC (C.12) to $\overline{\Pi}_{\mathrm{SR}}$ (as a function of p).

For Part 4, to prove that 4b is equivalent to 4a: (i) use Lemma 19.4, (ii) combine Lemma 19.5 with Remark 19.11, and (iii) apply the FOC (C.24) to $\underline{C}_{\mathrm{LR}}$ (as a function of r).

Finally, as in the Proof of Corollary 19.19, each of the three optimization systems (1c, 2c and 3c, repeated as 4c) is equivalent to (3.3). □

Together, Corollaries 19.19 and 19.20 establish the equivalence of all the various systems of Sections 4, 6–8 and 11. This includes the subdifferential and saddle-differential systems, whose equivalence is thus shown indirectly, through optimization systems ((“direct” proofs by the relevant rules of convex calculus have been given in Section 9).

For the reduced short-run programmes—viz., the reduced SRP programme for y in (3.13) and (4.2) and the short-run output reduced shadow-pricing programme for p in (5.8) and (6.10), whose solution sets are denoted by $\hat{Y}(p,k,w)$ and $\check{P}(y,k,w)$—there are the following “reduced” versions of Shephard-Hotelling Lemmas: a version of the short-run Hotelling's Lemma that is limited to output quantities, and a version of the dual to the short-run Shephard's Lemma that is limited to output prices.

Lemma 19.21 (SR Hotelling's Lemma for outputs only)**.** *The following conditions are equivalent to each other:*

1. $y \in \hat{Y}(p,k,w)$*, i.e.,* y *yields the supremum in (3.13), which is* $\Pi_{\mathrm{SR}}(p,k,w)$*.*
2. $y \in \partial_p \Pi_{\mathrm{SR}}(p,k,w)$ *and* $C_{\mathrm{SR}}(y,k,w) = \Pi_{\mathrm{SR}}^{\#_1}(y,k,w)$*, i.e.,* $C_{\mathrm{SR}} = C_{\mathrm{SR}}^{\#_1\#_1}$ *at* (y,k,w)*.*

The last equality holds if $\underline{C}_{\mathrm{SR}} = C_{\mathrm{SR}}$ *at* (y,k,w)*. Also, if* $\overline{\Pi}_{\mathrm{SR}} = \Pi_{\mathrm{SR}}$ *at* (p,k,w) *and* $y \in \partial_p \Pi_{\mathrm{SR}}(p,k,w)$ *then* $y \in \partial_p \overline{\Pi}_{\mathrm{SR}}(p,k,w)$*.*

Proof. Apply the Inversion Rule (Theorem C.2) and the FOC (C.12) to C_{SR} and its conjugate Π_{SR} as functions of y and p (with k and w fixed); alternatively, apply the Derivative Property (C.33) to conflate the two steps. This shows that Condition 1 and 2 are equivalent.

Fix any w and recall that $\underline{C}_{\mathrm{SR}} = C_{\mathrm{SR}}^{\#_{1,2}\#_{1,2}}$ by Lemma 19.18. So $\underline{C}_{\mathrm{SR}} = C_{\mathrm{SR}}$ at (y,k,w) if and only if $C_{\mathrm{SR}}^{\#_{1,2}\#_{1,2}} = C_{\mathrm{SR}}$ at (y,k,w), and then *a fortiori* $C_{\mathrm{SR}}^{\#_1\#_1} = C_{\mathrm{SR}}$ at (y,k,w) by Remark C.1.[94]

[94]For an alternative proof of this, note that: (i) by (5.8), $\underline{C}_{\mathrm{SR}} = \overline{\Pi}_{\mathrm{SR}}^{\#_1} \le \Pi_{\mathrm{SR}}^{\#_1} = C_{\mathrm{SR}}^{\#_1\#_1}$ by (3.13), with the inequality holding because $\overline{\Pi}_{\mathrm{SR}} \ge \Pi_{\mathrm{SR}}$, and (ii) $C_{\mathrm{SR}}^{\#_1\#_1} \le C_{\mathrm{SR}}$ by (C.4) without the middle term. So $\underline{C}_{\mathrm{SR}} \le C_{\mathrm{SR}}^{\#_1\#_1} \le C_{\mathrm{SR}}$ everywhere.

Finally, recall that $\overline{\Pi}_{\mathrm{SR}} \geq \Pi_{\mathrm{SR}}$ everywhere (on $P \times K \times W$). So if $\overline{\Pi}_{\mathrm{SR}} = \Pi_{\mathrm{SR}}$ at (p, k, w) and $y \in \partial_p \Pi_{\mathrm{SR}}(p, k, w)$ then also $y \in \partial_p \overline{\Pi}_{\mathrm{SR}}(p, k, w)$ by the subgradient inequality (C.11). □

Lemma 19.22 (Dual of SR Shephard's Lemma for outputs only). *The following conditions are equivalent to each other:*

1. $p \in \check{P}(y, k, w)$, *i.e., p yields the supremum in (5.8), which is* $\underline{C}_{\mathrm{SR}}(y, k, w)$.
2. $p \in \partial_y \underline{C}_{\mathrm{SR}}(y, k, w)$ *and* $\overline{\Pi}_{\mathrm{SR}}(p, k, w) = \underline{C}_{\mathrm{SR}}^{\#_1}(p, k, w)$, *i.e.,* $\overline{\Pi}_{\mathrm{SR}} = \overline{\Pi}_{\mathrm{SR}}^{\#_1\#_1}$ *at* (p, k, w).

The last equality holds if $\Pi_{\mathrm{SR}} = \overline{\Pi}_{\mathrm{SR}}$ *at* (p, k, w). *Also, if* $C_{\mathrm{SR}} = \underline{C}_{\mathrm{SR}}$ *at* (y, k, w) *and* $p \in \partial_y \underline{C}_{\mathrm{SR}}(y, k, w)$ *then* $p \in \partial_y C_{\mathrm{SR}}(y, k, w)$.

Proof. Being a "mirror image" of Lemma 19.21, this is proved by the same arguments, with $\underline{C}_{\mathrm{SR}}(y)$, $\overline{\Pi}_{\mathrm{SR}}(p)$ and $\Pi_{\mathrm{SR}}(p)$ in place of $\Pi_{\mathrm{SR}}(p)$, $C_{\mathrm{SR}}(y)$ and $\underline{C}_{\mathrm{SR}}(y)$, respectively. To spell this out, apply the Inversion Rule (Theorem C.2) and the FOC (C.12) to $\overline{\Pi}_{\mathrm{SR}}$ and its conjugate $\underline{C}_{\mathrm{SR}}$ as functions of p and y (with k and w fixed); alternatively, apply the Derivative Property (C.33) to conflate the two steps. This shows that Condition 1 and 2 are equivalent.

Fix any k and recall that $\Pi_{\mathrm{SR}} = \overline{\Pi}_{\mathrm{SR}}^{\#_{1,3}\#_{1,3}}$ by Lemma 19.13. So $\Pi_{\mathrm{SR}} = \overline{\Pi}_{\mathrm{SR}}$ at (p, w) if and only if $\overline{\Pi}_{\mathrm{SR}}^{\#_{1,3}\#_{1,3}} = \overline{\Pi}_{\mathrm{SR}}$ at (p, w), and then *a fortiori* $\overline{\Pi}_{\mathrm{SR}}^{\#_1\#_1}(p, w) = \overline{\Pi}_{\mathrm{SR}}(p, w)$ by Remark C.1.

Finally, recall that $C_{\mathrm{SR}} \geq \underline{C}_{\mathrm{SR}}$ everywhere (on $Y \times K \times W$). So if $C_{\mathrm{SR}} = \underline{C}_{\mathrm{SR}}$ at (y, k, w) and $p \in \partial_y \underline{C}_{\mathrm{SR}}(y, k, w)$ then also $p \in \partial_y C_{\mathrm{SR}}(y, k, w)$ by the subgradient inequality (C.11). □

Corollary 19.23. *Assume both that* $\underline{C}_{\mathrm{SR}} = C_{\mathrm{SR}}$ *at* (y, k, w) *and that* $\overline{\Pi}_{\mathrm{SR}} = \Pi_{\mathrm{SR}}$ *at* (p, k, w). *Then the following conditions are equivalent to one another:*

1. $y \in \hat{Y}(p, k, w)$, *i.e., y yields the supremum in (3.13), which is* $\Pi_{\mathrm{SR}}(p, k, w)$.
2. $p \in \partial_y C_{\mathrm{SR}}(y, k, w)$.
3. $y \in \partial_p \Pi_{\mathrm{SR}}(p, k, w)$.
4. $p \in \check{P}(y, k, w)$, *i.e., p yields the supremum in (5.8), which is* $\underline{C}_{\mathrm{SR}}(y, k, w)$.
5. $y \in \partial_p \overline{\Pi}_{\mathrm{SR}}(p, k, w)$.
6. $p \in \partial_y \underline{C}_{\mathrm{SR}}(y, k, w)$.

Proof. Lemmas 19.21 and 19.22 state that Conditions 1, 3, 4 and 6 are equivalent; to add Conditions 2 and 5, recall from the Proofs that these are the FOCs for the optima in Conditions 1 and 4. □

20. Preclusion of duality gaps by semicontinuity of optimal values

Once a pair of solutions (to a primal-dual programme pair) is found, a direct comparison of their values shows whether there is a duality gap. But there is also a method of checking for a gap at the outset—before solving the programmes. Namely, absence of a duality gap is equivalent to Type One semicontinuity of either optimal value, primal or dual (i.e., to the semicontinuity of primal value w.r.t. primal parameters, or of dual value w.r.t. dual parameters). This well-known result is next stated for the SRP, LRC and SRC problems. It is later complemented by sufficient criteria for value semicontinuity or continuity (Sections 21 and 23), which apply to profit or cost as functions of the quantities—Π_{SR} of k, C_{LR} of y, and C_{SR} of (y,k). Together, these results can serve to preclude duality gaps. (By contrast, semicontinuity of profit or cost in prices—Π_{SR} in (p,w), C_{LR} in (r,w), and C_{SR} in w—is an automatic Type Two property that does not rule out a duality gap: the primal value is always semicontinuous w.r.t. the dual parameter.)

Lemma 20.1 (Continuity conditions for equality of SRP to dual value)**.** *Assume that $\mathbb{Y}$ is closed. Then, for each $(p,w) \in P \times W$, the following conditions are equivalent to one another:*

1. $\overline{\Pi}_{\mathrm{SR}}(p,k,w) = \Pi_{\mathrm{SR}}(p,k,w)$.
2. *The concave function $\Pi_{\mathrm{SR}}(p,\cdot,w)$ is upper semicontinuous at k, and the primal (3.6)–(3.7) and the dual (5.6) are not both infeasible.*[95]
3. *The convex function $\overline{\Pi}_{\mathrm{SR}}(\cdot,k,\cdot)$ is lower semicontinuous at (p,w), and the primal (3.6)–(3.7) and the dual (5.6) are not both infeasible.*

Proof. To prove that Conditions 1 and 2 are equivalent, let Π and $\overline{\Pi}$ mean $\Pi_{\mathrm{SR}}(p,\cdot,w)$ and $\overline{\Pi}_{\mathrm{SR}}(p,\cdot,w)$, which are functions on K. Recall that

$$\Pi \le \operatorname{usc}\Pi \le \Pi_{\#\#} = \overline{\Pi} \tag{20.1}$$

by (C.21) and Lemma 19.14. The second inequality in (20.1) is strict if and *only* if its sides are oppositely infinite. So $\Pi(k) = \overline{\Pi}(k)$ if and only if: (i) $\Pi(k) = \operatorname{usc}\Pi(k)$, and (ii) $\Pi(k) > -\infty$ or $\overline{\Pi}(k) < +\infty$ (i.e., it is not the case that both $\Pi(k) = -\infty$ and $\overline{\Pi}(k) = +\infty$).

The equivalence of Conditions 1 and 3 is proved similarly: now let Π and $\overline{\Pi}$ mean $\Pi_{\mathrm{SR}}(\cdot,k,\cdot)$ and $\overline{\Pi}_{\mathrm{SR}}(\cdot,k,\cdot)$, which are functions on $P \times W$. Then

$$\overline{\Pi} \ge \operatorname{lsc}\overline{\Pi} \ge \overline{\Pi}^{\#\#} = \Pi \tag{20.2}$$

by (C.4) and Lemma 19.13. So $\overline{\Pi}(p,w) = \Pi(p,w)$ if and only if: (i) $\overline{\Pi}(p,w) = \operatorname{lsc}\overline{\Pi}(p,w)$, and (ii) $\overline{\Pi}(p,w) < +\infty$ or $\Pi(p,w) > -\infty$. □

Lemma 20.2 (Continuity conditions for equality of LRC to dual value)**.** *Assume that $\mathbb{Y}$ is closed. Then, for each $(r,w) \in R \times W$, the following conditions are equivalent to one another:*

1. $\underline{C}_{\mathrm{LR}}(y,r,w) = C_{\mathrm{LR}}(y,r,w)$.
2. *The convex function $C_{\mathrm{LR}}(\cdot,r,w)$ is lower semicontinuous at y, and the primal (3.8)–(3.9) and the dual (5.5) are not both infeasible.*
3. *The concave function $\underline{C}_{\mathrm{LR}}(y,\cdot,\cdot)$ is upper semicontinuous at (r,w), and the primal (3.8)–(3.9) and the dual (5.5) are not both infeasible.*

Proof. This can be proved like Lemma 20.1: to prove that Conditions 1 and 2 are equivalent, let C and $\underline{C}$ mean $C_{\mathrm{LR}}(\cdot,r,w)$ and $\underline{C}_{\mathrm{LR}}(\cdot,r,w)$, which are functions on Y. Recall

[95] The primal (3.6)–(3.7) or the dual (5.6) is feasible if and only if $\mathbb{Y}_{\mathrm{SR}}(k) \ne \emptyset$ or $\Pi_{\mathrm{LR}}(p,\cdot,w) \ne +\infty$, respectively. When $\mathbb{Y}$ is a cone (i.e., under c.r.t.s.), this means that $\mathbb{Y}_{\mathrm{SR}}(k) \ne \emptyset$ or $\mathbb{Y}^{\circ}_{p,w} \ne \emptyset$; the two sections are defined by (21.1) and (19.2).

that

$$C \geq \operatorname{lsc} C \geq C^{\#\#} = \underline{C} \tag{20.3}$$

by (C.4) and Lemma 19.16. The second inequality in (20.3) is strict if and *only* if its sides are oppositely infinite. So $C(y) = \underline{C}(y)$ if and only if: (i) $C(y) = \operatorname{lsc} C(y)$, and (ii) $C(y) < +\infty$ or $\underline{C}(y) > -\infty$.

The equivalence of Conditions 1 and 3 is proved similarly: now let C and $\underline{C}$ mean $C_{\mathrm{LR}}(y, \cdot, \cdot)$ and $\underline{C}_{\mathrm{LR}}(y, \cdot, \cdot)$, which are functions on $R \times W$. Then

$$\underline{C} \leq \operatorname{usc} \underline{C} \leq \underline{C}_{\#\#} = C \tag{20.4}$$

by (C.21) and Lemma 19.15. So $\underline{C}(r, w) = C(r, w)$ if and only if: (i) $\underline{C}(r, w) = \operatorname{usc} \underline{C}(r, w)$, and (ii) $\underline{C}(r, w) > -\infty$ or $C(r, w) < +\infty$. □

Lemma 20.3 (Continuity conditions for equality of SRC to dual value). *Assume that $\mathbb{Y}$ is closed. Then, for each $w \in W$, the following conditions are equivalent to one another:*

1. $\underline{C}_{\mathrm{SR}}(y, k, w) = C_{\mathrm{SR}}(y, k, w)$.
2. *The convex function $C_{\mathrm{SR}}(\cdot, \cdot, w)$ is lower semicontinuous at (y, k), and the primal (3.10)–(3.11) and the dual (5.4) are not both infeasible.*
3. *The concave function $\underline{C}_{\mathrm{SR}}(y, k, \cdot)$ is upper semicontinuous at w, and the primal (3.10)–(3.11) and the dual (5.4) are not both infeasible.*

Proof. This can be proved like Lemmas 20.1 and 20.2: to prove that Conditions 1 and 2 are equivalent, let C and $\underline{C}$ mean $C_{\mathrm{SR}}(\cdot, \cdot, w)$ and $\underline{C}_{\mathrm{SR}}(\cdot, \cdot, w)$, which are functions on $Y \times K$. Recall that

$$C \geq \operatorname{lsc} C \geq C^{\#\#} = \underline{C} \tag{20.5}$$

by (C.4) and Lemma 19.18. The second inequality in (20.5) is strict if and *only* if its sides are oppositely infinite. So $C(y) = \underline{C}(y)$ if and only if: (i) $C(y, k) = \operatorname{lsc} C(y, k)$, and (ii) $C(y, k) < +\infty$ or $\underline{C}(y, k) > -\infty$.

The equivalence of Conditions 1 and 3 is proved similarly: now let C and $\underline{C}$ mean $C_{\mathrm{SR}}(y, k, \cdot)$ and $\underline{C}_{\mathrm{SR}}(y, k, \cdot)$, which are functions on W. Then

$$\underline{C} \leq \operatorname{usc} \underline{C} \leq \underline{C}_{\#\#} = C \tag{20.6}$$

by (C.21) and Lemma 19.17. So $\underline{C}(w) = C(w)$ if and only if: (i) $\underline{C}(w) = \operatorname{usc} \underline{C}(w)$, and (ii) $\underline{C}(w) > -\infty$ or $C(w) < +\infty$. □

Comment: Profit and cost are always semicontinuous in prices (as are the dual values in quantities), i.e., for every p, y, r, k, w and v:

1. (i) $\Pi_{\mathrm{SR}}(\cdot, k, \cdot)$ is l.s.c. convex on $P \times W$, (ii) $C_{\mathrm{LR}}(y, \cdot, \cdot)$ is u.s.c. concave on $R \times W$, and (iii) $C_{\mathrm{SR}}(y, k, \cdot)$ is u.s.c. concave on W.
2. (i) $\overline{\Pi}_{\mathrm{SR}}(p, \cdot, w)$ is u.s.c. concave on K, (ii) $\underline{C}_{\mathrm{LR}}(\cdot, r, w)$ is l.s.c. convex on Y, and (iii) $\underline{C}_{\mathrm{SR}}(\cdot, \cdot, w)$ is l.s.c. convex on $Y \times K$.

These Type Two results (which are part of Lemmas 19.13–19.18) follow directly from the definitions: e.g., $\Pi_{\mathrm{SR}}(\cdot, k, \cdot)$ is the pointwise supremum of a family of continuous (and linear) functions on $P \times W$ (and likewise $\overline{\Pi}_{\mathrm{SR}}(p, \cdot, w)$ is the pointwise infimum of such functions on K). This also shows that Π_{SR} is proper convex in (p, w), and that $\overline{\Pi}_{\mathrm{SR}}$ is proper concave in k, unless the one or the other is an infinite constant. (What is more, $\Pi_{\mathrm{SR}}(\cdot, k, \cdot)$ and $\overline{\Pi}_{\mathrm{SR}}(p, \cdot, w)$ are second conjugates of $\overline{\Pi}_{\mathrm{SR}}(\cdot, k, \cdot)$ and $\Pi_{\mathrm{SR}}(p, \cdot, w)$, respectively, by Lemmas 19.13 and 19.14.)

21. Semicontinuity of cost or profit in the quantity variables (complement to Sections 4, 5 and 20)

As Section 20 shows, to preclude duality gaps between the SRP or SRC programme and its dual—which are used in the short-run approach (Sections 12 and 13)—one needs to ensure that $\Pi_{\mathrm{SR}}(p, k, w)$ is u.s.c. in k, or that $C_{\mathrm{SR}}(y, k, w)$ is l.s.c. in (y, k). A setting that, by itself, guarantees this Type One semicontinuity of the optimal value is *finite* linear programming: see, e.g., [44, Example 1', p. 24] for a proof based on polyhedral convexity. So finite LPs can have no duality gaps, except when both the primal and the dual are actually infeasible (i.e., when their values are oppositely infinite); hence, any paired solutions have equal values.[96] Here, this applies when all the spaces (Y, etc.) are finite-dimensional *and* the constraint set is polyhedral (i.e., when $\mathbb{Y}$ is the intersection of a finite number of finite-dimensional closed half-spaces). But this does not extend to infinite LPs (see Appendix A), which therefore require other methods of ensuring semicontinuity (and thus ruling out a gap and ensuring that the marginal values are of Type One).

One way to obtain such results for general convex programmes in infinite-dimensional spaces is to apply Berge's Theorem under a suitable compactness condition on the constraint set. Here, such a condition is met when the relevant subsets of the production set $\mathbb{Y}$ are bounded and the commodity space is the dual of a completely normed vector lattice, i.e., a dual Banach lattice (with a norm $\|\cdot\|$ and a vector order $\leq$). Therefore, Y, K, and V are henceforth taken to be duals of Banach lattices: $(Y, \|\cdot\|, \leq)$ is the dual of some $(Y', \|\cdot\|', \leq)$, etc., the nonnegative cones in Y and Y' are denoted by Y_+ and Y'_+, and P is either Y' or Y^*—except for Sections 24, 25 and 28, in which P is any space paired with Y (which need not be a dual Banach space or a lattice).

Notation and definitions: Every nonnegative linear functional on a Banach lattice is norm-continuous (Birkhoff's Theorem): see, e.g., [2, 12.5], [8, XV: Theorem 18] or [39, 1.3.7]. In other words, the norm-dual and the order-dual of a Banach lattice are equal—so both can be called the *Banach dual*. The Banach dual of Y, denoted by $(Y^*, \|\cdot\|^*, \leq)$, contains the Banach predual of Y, i.e., $Y^* = Y'^{**} \supseteq Y'$, but Y^* can be larger than Y'. Either can serve as the price space P, and the general equilibrium price system may belong to Y' or to Y^*, depending on the assumptions: see [7] and [26].

The weak and the Mackey topologies on Y for its pairing with P (which is either Y' or Y^*) are denoted by $\mathrm{w}(Y, P)$ and $\mathrm{m}(Y, P)$. These are the weakest and the strongest of those locally convex topologies on Y which yield P as the continuous dual space. Since every *convex* $\mathrm{m}(Y, P)$-closed set is $\mathrm{w}(Y, P)$-closed,[97] a *convex* function on Y (with values in $\mathbb{R} \cup \{\pm\infty\}$) is $\mathrm{m}(Y, P)$-lower semicontinuous if and *only* if it is $\mathrm{w}(Y, P)$-l.s.c. So these properties can be simply called *P-closedness*

[96] This can also be proved by using the simplex algorithm: see, e.g., [11, 5.1 and 9.1].

[97] This is a corollary to the Hahn-Banach Separation Theorem: see, e.g., [18, 12A: Corollary 1].

(of a convex subset of Y) and *P-lower semicontinuity* (of a convex function on Y).

When $P = Y'$, the notation $\mathrm{w}(Y, Y') = \mathrm{w}(P^*, P)$ and $\mathrm{m}(Y, Y') = \mathrm{m}(P^*, P)$ is abbreviated to w^* and m^*, and these are called the *weak** and the *Mackey topologies*. For comparison, $\mathrm{w}(Y, Y^*)$ is simply called the *weak topology*, whilst $\mathrm{m}(Y, Y^*)$ is identical to the norm topology of Y.

The *bounded weak* topology* on Y, denoted by bw^*, is a locally convex topology stronger than w^*, but weaker than m^*. It can be defined as the topology of uniform convergence on norm-compact subsets of Y', or by stipulating that a subset of Y is bw^*-closed if and only if its intersection with any closed ball in Y is w^*-closed (or, equivalently, w^*-compact): see, e.g., [18, 18D: Corollary (b)].

From here on, conditions on production set $\mathbb{Y}$ are selected from those listed below. To capture any long-run constraint on producible outputs, we use the projection of $\mathbb{Y}$ on Y, which is[98]

$$\mathrm{proj}_Y(\mathbb{Y}) := \{y \in Y : \exists (k, v) \ (y, -k, -v) \in \mathbb{Y}\}.$$

Some of the conditions use sections of $\mathbb{Y}$, viz., the short-run production set

$$\mathbb{Y}_{\mathrm{SR}}(k) := \{(y, -v) : (y, -k, -v) \in \mathbb{Y}\} \tag{21.1}$$

(which is the section through $-k$), and the long-run and short-run input requirement sets

$$\mathbb{I}_{\mathrm{LR}}(y) := \{(k, v) : (y, -k, -v) \in \mathbb{Y}\}$$
$$\mathbb{I}_{\mathrm{SR}}(y, k) := \{v : (y, -k, -v) \in \mathbb{Y}\}.$$

When Z is one of these sets, denote by

$$\mathrm{vmax}\, Z \quad \text{and} \quad \mathrm{vmin}\, Z$$

the sets of all the maximal and the minimal points for the vector order $\leq$ restricted to Z. Such points represent the efficient output or input bundles. These are next assumed to form norm-bounded sets—which, for brevity, are called *bounded* (but need not be order-bounded, unless the space is L^∞ with the supremum norm $\|\cdot\|_\infty$).

Production Set Assumptions (PSAs). Five assumptions are maintained from here on (though not all are always needed):

(1) $\mathbb{Y}$ is a cone (with a vertex at 0).
(2) $\mathbb{Y}$ is convex.
(3) $\mathbb{Y}$ is weakly* closed, i.e., $\mathrm{w}(Y \times K \times V, Y' \times K' \times V')$-closed.
(4) $\mathbb{Y}$ includes free disposal of inputs and of producible outputs; i.e., if $k \leq \widetilde{k}$, $v \leq \widetilde{v}$, $y \geq \widetilde{y} \in \mathrm{proj}_Y(\mathbb{Y})$ and $(y, -k, -v) \in \mathbb{Y}$, then $\left(\widetilde{y}, -\widetilde{k}, -\widetilde{v}\right) \in \mathbb{Y}$.
(5) $\mathbb{I}_{\mathrm{LR}}(y) \subseteq K_+ \times V_+$ (i.e., $\mathbb{Y} \subseteq Y \times K_+ \times V_+$).

The following assumptions are made *selectively* where needed:

(6) For every $(k, v) \in \mathbb{I}_{\mathrm{LR}}(y)$, there exists a $\left(\widetilde{k}, \widetilde{v}\right) \in \mathrm{vmin}\, \mathbb{I}_{\mathrm{LR}}(y)$ with $\left(\widetilde{k}, \widetilde{v}\right) \leq (k, v)$.

[98] $\mathrm{proj}_Y(\mathbb{Y})$ need not be comprehensive downwards (i.e., it need not contain $Y_- = -Y_+$).

(7) For every bounded set $B \subset Y$, the set $\bigcup_{y \in B} \operatorname{vmin} \mathbb{I}_{\mathrm{LR}}(y)$ is also bounded.[99]
(8) For every $(y, -v) \in \mathbb{Y}_{\mathrm{SR}}(k)$, there exists a $(\widetilde{y}, -\widetilde{v}) \in \operatorname{vmax} \mathbb{Y}_{\mathrm{SR}}(k)$ with $(\widetilde{y}, -\widetilde{v}) \geq (y, -v)$.
(9) For every bounded set $B \subset K$, the set $\bigcup_{k \in B} \operatorname{vmax} \mathbb{Y}_{\mathrm{SR}}(k)$ is also bounded.
(10) For every $v \in \mathbb{I}_{\mathrm{SR}}(y, k)$, there exists a $\widetilde{v} \in \operatorname{vmin} \mathbb{I}_{\mathrm{SR}}(y, k)$ with $\widetilde{v} \leq v$.
(11) For every bounded set $B \subset Y \times K$, the set $\bigcup_{(y,k) \in B} \operatorname{vmin} \mathbb{I}_{\mathrm{SR}}(y, k)$ is also bounded. (This follows from either of PSAs 7 and 9.)

Our PSAs are similar to the conditions of [13, p. 134] and [14, p. 580] for the finite-dimensional case (see also the end of this section for further comments). When the commodity space is $\mathbb{R}^n$ or, more generally, a dual Banach lattice (e.g., L^∞ or L^ϱ with $\varrho > 1$), the assumptions that efficient points exist (PSAs 6, 8 and 10) can be derived from simpler conditions by using the following lemma.

Lemma 21.1 (Existence of maximal points)**.** *Let $(L, \|\cdot\|, \leq)$ be the dual of a Banach lattice $\left(L', \|\cdot\|', \leq\right)$. If B is a norm-bounded and $\mathrm{w}(L, L')$-closed nonempty subset of L, then the restriction, to B, of the lattice order $\leq$ has a maximal element.*

Proof. Given any chain H in B (i.e., a subset of B that is totally ordered by $\leq$), define a linear functional y_H on L' by[100]

$$\langle p \,|\, y_H \rangle := \sup_{y \in H} \langle p \,|\, y \rangle \quad \text{for } p \in L'_+$$

where the supremum is finite because $\sup_{y \in H} \|y\| \leq \sup_{y \in B} \|y\| < +\infty$. Then $y_H \in L$ (in other words, y_H is the supremum of H in the lattice L). This can be shown in two ways: (i) $y_H - y \geq 0$ for any $y \in H$, and every nonnegative linear functional on L' belongs to L, and (ii) $\|y_H\| \leq \sup_{y \in H} \|y\|$. Next, to show that $y_H \in B$, note that

$$\langle p \,|\, y_H \rangle := \sup_{y \in H} \left\langle p^+ \,|\, y \right\rangle - \sup_{y \in H} \left\langle p^- \,|\, y \right\rangle = \lim_{y \nearrow, \, y \in H} \langle p \,|\, y \rangle$$

for each $p \in L'$. This exhibits y_H as the $\mathrm{w}(L, L')$-limit of a net in B (the identity map on H can serve as such a net). So $y_H \in B$ (since B is weakly* closed). Thus the assumption of Zorn's Lemma is verified for $\leq$ as a partial order on B (and so a maximal point exists). $\square$

Corollary 21.2 (Existence of efficient points)**.** *Assume PSA 3. Then:*

(1) *PSA 8 holds if the set*

$$\mathbb{Y}_{\mathrm{SR}}(k) \cap ((y, -v) + (Y_+ \times V_+)) \tag{21.2}$$

is bounded, for each y, k and v.

(2) *Similarly, PSAs 6 and 10 follow from PSA 5.*

Proof. For Part 1, apply Lemma 21.1 to the bounded set (21.2), which is $\mathrm{w}(Y \times V, Y' \times V')$-closed by PSA 3.

[99] If L^∞ is the space Y in PSA 7, or K in PSA 9, or $Y \times K$ in PSA 11, then it obviously suffices to make this assumption for each singleton set (instead of B).

[100] This construction is used for proving related but different results in, e.g., [2, 14.11] and [33, X.4: Theorem 6].

For Part 2, apply Lemma 21.1 to the negatives of the sets

$$\mathbb{I}_{\mathrm{LR}}(y) \cap ((k,v) - (K_+ \times V_+)) \quad \text{and} \quad \mathbb{I}_{\mathrm{SR}}(y,k) \cap (v - V_+) \tag{21.3}$$

which are bounded (and even order-bounded) by PSA 5, and are weakly* closed by PSA 3. □

To exploit weak* compactness of the efficient boundary of the short-run production set (PSA 8), the maximand $\langle p, w \mid \cdot, \cdot \rangle$ must be weakly* continuous (i.e., p and w must be in Y' and in V'). As is shown next, this guarantees that Π_{SR} is u.s.c. in k (and obviously the condition on p and w is restrictive only when the spaces are infinite-dimensional and nonreflexive, i.e., when $Y' \neq Y^*$ or $V' \neq V^*$).

Lemma 21.3 (Semicontinuity of SRP in fixed inputs)**.** *Under the PSAs 8 and 9 in addition to PSAs 2 and 3, the concave function* $\Pi_{\mathrm{SR}}(p, \cdot, w) \colon K \to \mathbb{R} \cup \{-\infty\}$ *is* K'*-upper semicontinuous (on* K*), for each* $(p,w) \in Y'_+ \times V'_+$*.*[101]

Proof. That $\Pi(k) < +\infty$ for every k follows from PSAs 8 and 9 for $B = \{k\}$, with Π meaning $\Pi_{\mathrm{SR}}(p, \cdot, w)$. Next, since Π is concave, it suffices to prove that it is u.s.c. for the bounded weak* topology, i.e., that Π is weakly* u.s.c. on any norm-bounded set $B \subset K = K'^*$. (This is because every bw*-closed *convex* set is w*-closed, by the Krein-Smulian Theorem: see, e.g., [18, 18E: Corollary 2].) And a bound on k implies a bound on the *efficient* combinations of y and v (which are the only ones that matter because $(p,w) \geq 0$). In precise terms, the set

$$A := \mathrm{cl}_{\mathrm{w}(Y \times V, Y' \times V')} \bigcup_{k \in B} \mathrm{vmax}\, \mathbb{Y}_{\mathrm{SR}}(k)$$

is $\mathrm{w}(Y \times V, Y' \times V')$-compact by PSA 9 and the Banach-Alaoglu Theorem; and for every $k \in B$

$$\Pi_{\mathrm{SR}}(k) = \sup_{y,\,v} \{ \langle p \mid y \rangle - \langle w \mid v \rangle : (y, -v) \in \mathbb{Y}_{\mathrm{SR}}(k) \cap A \} \tag{21.4}$$

by PSA 8. Since $(p,w) \in Y' \times V'$, the maximand in (21.4) is $\mathrm{w}(Y \times V, Y' \times V')$-u.s.c. (and actually continuous) in (y,v) jointly. In addition, since $\mathbb{Y}$ is $\mathrm{w}(Y \times K \times V, Y' \times K' \times V')$-closed (and A is compact), the constraint correspondence $k \mapsto \mathbb{Y}_{\mathrm{SR}}(k) \cap A$ is compact-valued and upper hemicontinuous ($\mathrm{w}(K, K')$-to-$\mathrm{w}(Y \times V, Y' \times V')$): see, e.g., [34, 7.1.16]. So Π_{SR} is $\mathrm{w}(K, K')$-u.s.c. on B by the relevant part of Berge's Maximum Theorem [6, VI.3: Theorem 2].[102] □

Similar Type One semicontinuity results are next given for the other functions: C_{LR} is l.s.c. in y, and C_{SR} is l.s.c. in (y,k) jointly.

Lemma 21.4 (Semicontinuity of LRC in outputs)**.** *Under the PSAs 6 and 7 in addition to PSAs 2, 3 and 5, the convex function* $C_{\mathrm{LR}}(\cdot, r, w) \colon Y \to \mathbb{R}_+ \cup \{+\infty\}$ *is* Y'*-lower semicontinuous (on* Y*), for each* $(r,w) \in K'_+ \times V'_+$*.*

[101] Also, under PSA 4, if $0 \in \mathbb{Y}$ then $\Pi_{\mathrm{SR}}(p, \cdot, w) \geq 0$ on K_+ (and it is $-\infty$ outside of K_+).

[102] Another way to apply Berge's Theorem [6, VI.3: Theorem 2] is to take $\langle p \mid y \rangle - \langle w \mid v \rangle - \delta(y, -k, -v \mid \mathbb{Y})$ as the maximand (u.s.c. in (y,k,v)) and A as the constraint set (compact and independent of $k \in B$).

Proof. This is proved like Lemma 21.3: since the function $C := C_{\mathrm{LR}}(\cdot, r, w)$ is convex (on Y), it suffices to prove that C is l.s.c. for the bounded weak* topology, i.e., that C is weakly* l.s.c. on any norm-bounded set $B \subset Y = Y'^{*}$. And a bound on y implies a bound on the *efficient* combinations of k and v (which are the only ones that matter with $(r, w) \geq 0$). In precise terms, the set

$$A := \mathrm{cl}_{\mathrm{w}(K \times V, K' \times V')} \bigcup_{y \in B} \mathrm{vmin}\, \mathbb{I}_{\mathrm{LR}}(y)$$

is $\mathrm{w}(K \times V, K' \times V')$-compact by PSA 7 and the Banach-Alaoglu Theorem; and for every $y \in B$

$$C_{\mathrm{LR}}(y) := \inf_{k,v} \{\langle r \mid k \rangle + \langle w \mid v \rangle : (k, v) \in \mathbb{I}_{\mathrm{LR}}(k) \cap A\} \tag{21.5}$$

by PSA 6. Since $(r, w) \in K' \times V'$, the minimand in (21.5) is $\mathrm{w}(K \times V, K' \times V')$-l.s.c. (and actually continuous) in (k, v). In addition, since $\mathbb{Y}$ is $\mathrm{w}(Y \times K \times V, Y' \times K' \times V')$-closed (and A is compact), the constraint correspondence $y \mapsto \mathbb{I}_{\mathrm{LR}}(y) \cap A$ is compact-valued and upper hemicontinuous ($\mathrm{w}(Y, Y')$-to-$\mathrm{w}(K \times V, K' \times V')$): see, e.g., [34, 7.1.16]. So C is $\mathrm{w}(Y, Y')$-l.s.c. on B by the relevant part of Berge's Maximum Theorem [6, VI.3: Theorem 2], reoriented to minimization. Finally, $C \geq 0$ by PSA 5. □

Lemma 21.5 (Semicontinuity of SRC in fixed quantities)**.** *Under the PSAs 10 and 11 in addition to PSAs 2, 3 and 5, the convex function $C_{\mathrm{SR}}(\cdot, \cdot, w) : Y \times K \to \mathbb{R}_{+} \cup \{+\infty\}$ is $(Y' \times K')$-lower semicontinuous (on $Y \times K$), for each $w \in V'_{+}$.*

Proof. This is proved like Lemmas 21.3 and 21.4: since the function $C := C_{\mathrm{SR}}(\cdot, \cdot, w)$ is convex, it suffices to show that it is l.s.c. for the bounded weak* topology, i.e., that C is weakly* l.s.c. on any norm-bounded set $B \subset Y \times K = (Y' \times K')^{*}$. And bounds on k and on y imply a bound on the *efficient* v's (which are the only ones that matter with $w \geq 0$). In precise terms, the set

$$A := \mathrm{cl}_{\mathrm{w}(V, V')} \bigcup_{(y,k) \in B} \mathrm{vmin}\, \mathbb{I}_{\mathrm{SR}}(y, k)$$

is $\mathrm{w}(V, V')$-compact by PSA 11 and the Banach-Alaoglu Theorem; and for every $(y, k) \in B$

$$C_{\mathrm{SR}}(y, k) := \inf_{v} \{\langle w \mid v \rangle : v \in \mathbb{I}_{\mathrm{SR}}(y, k) \cap A\} \tag{21.6}$$

by PSA 10. Since $w \in V'$, the minimand in (21.6) is $\mathrm{w}(V, V')$-l.s.c. (and actually continuous) in v. In addition, since $\mathbb{Y}$ is $\mathrm{w}(Y \times K \times V, Y' \times K' \times V')$-closed (and A is compact), the constraint correspondence $(y, k) \mapsto \mathbb{I}_{\mathrm{SR}}(y, k) \cap A$ is compact-valued and upper hemicontinuous ($\mathrm{w}(Y \times K, Y' \times K')$-to-$\mathrm{w}(V, V')$): see, e.g., [34, 7.1.16]. So C is $\mathrm{w}(Y \times K, Y' \times K')$-l.s.c. on B by the relevant part of Berge's Maximum Theorem [6, VI.3: Theorem 2], reoriented to minimization. Finally, $C \geq 0$ by PSA 5. □

So profit and cost are semicontinuous functions on the commodity space paired with its *pre*-dual, on condition that the given price system lies in that predual space. On the same condition, these functions are *a fortiori* semicontinuous for the pairing of the commodity space with its dual (instead of the predual) as the price space. Since either space (dual or predual) can serve as the range for the decision variable of the dual programme, the "predual"-semicontinuity does add to the results that would follow from just the "dual"-semicontinuity of the optimal value. In ruling out a duality gap (and thus ensuring that the marginal values are of Type One), the stronger property of "predual"-semicontinuity means that the primal and the dual programmes have the same value when the dual decision variable ranges only over the predual space (hence, *a fortiori*, also when it ranges over the dual space).

In symbols, if (i) $w \in V'_+$, and (ii) $p \in Y'_+$ or (iii) $r \in K'_+$ then, respectively: (i) $\Pi_{\mathrm{SR}}(p, \cdot, w)$ is K'-u.s.c. (and *a fortiori* also K^*-u.s.c.) on K, (ii) $C_{\mathrm{SR}}(\cdot, \cdot, w)$ is $(Y' \times K')$-l.s.c. (and *a fortiori* also $(Y^* \times K^*)$-l.s.c.) on $Y \times K$, and (iii) $C_{\mathrm{LR}}(\cdot, r, w)$ is Y'-l.s.c. (and *a fortiori* also Y^*-l.s.c.) on Y. In our notation for the dual values ($\overline{\Pi}$, $\underline{C}$) and the marginal values ($\widehat{\partial}_k \overline{\Pi}$, $\widehat{\partial}_k \Pi$, $\partial_{y,k} \underline{C}$, $\partial_{y,k} C$), the cases of either the predual or the dual as the price space are distinguished by the superscripts $'$ and $*$—although the $*$ is suppressed from ∂^* in [21], [24], [26] and [28]. (Both $'$ and $*$ are also suppressed when the predual equals the dual, e.g., when the space is finite-dimensional.) Since $Y' \subseteq Y^*$, $V' \subseteq V^*$ and $K' \subseteq K^*$,

$$\overline{\Pi}'_{\mathrm{SR}}(p,k,w) \geq \overline{\Pi}^*_{\mathrm{SR}}(p,k,w) = \overline{\Pi}_{\mathrm{SR}}(p,k,w) \geq \Pi_{\mathrm{SR}}(p,k,w)$$
$$\underline{C}'_{\mathrm{LR}}(y,r,w) \leq \underline{C}^*_{\mathrm{LR}}(y,r,w) = \underline{C}_{\mathrm{LR}}(y,r,w) \leq C_{\mathrm{LR}}(y,r,w)$$
$$\underline{C}'_{\mathrm{SR}}(y,k,w) \leq \underline{C}^*_{\mathrm{SR}}(y,k,w) = \underline{C}_{\mathrm{SR}}(y,k,w) \leq C_{\mathrm{SR}}(y,k,w)$$

for every $p \in Y^*$, $w \in V^*$ and $r \in K^*$—with equalities when $p \in Y'$, $w \in V'$ and $r \in K'$ (by Lemmas 21.3–21.5 and Propositions 20.1–20.3). Then (Section 19) the solution set for the "starred" dual (viz., (5.6), (5.5) or (5.4) with $P = Y^*$, $W = V^*$, $R = K^*$) is equal to $\widehat{\partial}^*_k \Pi_{\mathrm{SR}}(p,k,w)$, $\partial^*_y C_{\mathrm{LR}}(y,r,w)$ or $\partial^*_{y,k} C_{\mathrm{SR}}(y,k,w)$. Likewise, the solution set for the "primed" dual (viz., (5.6), (5.5) or (5.4) with $P = Y'$, $W = V'$, $R = K'$) is then equal to $\widehat{\partial}'_k \Pi_{\mathrm{SR}}$, $\partial'_y C_{\mathrm{LR}}$ or $\partial'_{y,k} C_{\mathrm{SR}}$—which always equals $K' \cap \widehat{\partial}^*_k \Pi$, $Y' \cap \partial^*_y C_{\mathrm{LR}}$ or $(Y' \times K') \cap \partial^*_{y,k} C_{\mathrm{SR}}$, by definition. It follows that any solutions to the "primed" dual are exactly those solutions to the "starred" dual which do belong to the smaller, "primed" space for the dual variable.[103]

Comment: Y'-semicontinuity of $C := C_{\mathrm{SR}}(\cdot, k, w)$ is also useful in subdifferentiating its conjugate, $\Pi := \Pi_{\mathrm{SR}}(\cdot, k, w)$, as a function on Y^* but *at* a $p \in Y'$ (with k and w fixed). Namely,

$$\left. \begin{array}{c} C \text{ is } Y'\text{-l.s.c. proper convex on } Y \\ (\text{and } \Pi = C^{\#} \text{ on } Y^*) \end{array} \right\} \Rightarrow \partial \Pi(p) = \partial \left(\Pi_{|Y'}\right)(p) \text{ for } p \in Y' \tag{21.7}$$

i.e., at any $p \in Y'$ the subdifferential of Π: $Y^* \to \mathbb{R} \cup \{+\infty\}$ can be evaluated after restricting Π to the subspace Y' (which makes the task easier). This can be proved by applying the Inversion Rule (C.31) to the cases of either Y' or Y^* as P, and comparing the results. (In other words, it follows from the "reduced" version of Hotelling's Lemma (Lemma 19.21) applied to the cases of either Y' or Y^* as P.)

Comments (on the Proofs of Lemmas 21.3–21.5):

[103]This is of course true whenever the "primed" and "starred" dual values are equal, whether or not the common dual value equals the primal value, e.g., whenever $\overline{\Pi}'_{\mathrm{SR}} = \overline{\Pi}^*_{\mathrm{SR}}$, whether or not this equals Π_{SR}.

(1) These proofs exemplify the advantage of using the bounded weak* topology to exploit convexity: for a convex function C on a dual Banach space Y, the question of weak* lower semicontinuity is reduced to bounded sets $B \subset Y$—even though a bounded set is never a weak* neighbourhood (unless Y is finite-dimensional). By itself, an application of Berge's Theorem [6, VI.3: Theorem 2] can prove only that C is weakly* l.s.c. on every ball B.[104] The Krein-Smulian Theorem upgrades this result to weak* l.s. continuity on Y (and not just on each B). To obtain a continuity result on Y without this step, one would have to put the norm topology on Y to make B a neighbourhood—and then the conclusion would be weaker, viz., only that C is norm-l.s.c. on Y (i.e., that it is Y^*-l.s.c. but not that it is Y'-l.s.c.).
(2) The bounded weak* topology can be equally useful in other contexts: e.g., in [21] and [24] we use the Krein-Smulian Theorem to show that the production sets for storage technologies are weakly* closed (in an L^∞ space). In [25], we devise another "localization" technique, and we combine the two for a simple proof that the additively separable consumer utility is Mackey continuous on L^∞_+.
(3) Since duality of programmes is symmetric, absence of a duality gap could also be proved by showing that the dual value is semicontinuous in the dual parameter (instead of showing that the primal value is semicontinuous in the primal parameter), i.e., by verifying Condition 3 (instead of 2) of Lemma 20.1, etc.

Comments (on the Production Set Assumptions):
(1) PSA 9 formalizes the notion that fixed inputs impose capacity constraints.[105]
(2) Unlike the fixed inputs, the variable inputs alone need not impose any bound on $\|y\|$: see (15.1).
(3) Unlike the inputs k and v, which are always nonnegative by PSA 5, the "output" can be a signed bundle $y = y^+ - y^-$, where $y^\pm$ are the nonnegative and nonpositive parts. This is convenient when, e.g., y represents a single good differentiated over time, and the dated commodities cannot be classified as net inputs or net outputs *a priori*. For example, the output from storage y is always signed, i.e., $y^+ \neq 0 \neq y^-$ unless $y = 0$: see (15.4).[106]

22. Solubility of cost and profit programmes

In addition to semicontinuity of Π_{SR}, C_{LR} and C_{SR} (which rules out duality gaps), the PSAs also guarantee solubility of the primal programmes of SRP, LRC and SRC optimization, when p, r and w are in the preduals (Y', K' and V'). This is because the relevant subsets of the constraint set ($\mathbb{Y}$) are then weakly* compact (so Weierstrass's Extreme Value Theorem applies). This is next recorded formally.

[104] For Berge's Theorem to apply, the (efficient) range of the decision variable must be contained in a weak* compact that is independent of the parameter (y) as it ranges over a set $B \subset Y$—and so B must be bounded. (The result stated in [44, Example 4' after (5.13)] also applies, but it is a special case of Berge's.)

[105] PSAs 7, 9 and 11 make it possible (in Lemmas 21.3–21.5) to prove semicontinuity of profit and costs at *every* $(p, r, w) \geq 0$ (and not only at strictly positive prices as is done in [13] or [14], for finite-dimensional spaces).

[106] When, by contrast, a signed y could arise *only* from free disposal, the good is essentially a net output.

Proposition 22.1 (Solubility of SRP programme)**.** *Under the Production Set Assumptions 3, 8 and 9,*[107] *if* $(p,w) \in Y'_+ \times V'_+$ *and the short-run profit maximization programme (3.6)–(3.7) is feasible, then it has a solution.*

Proof. It is similar to the Proof of Lemma 21.3, but simpler. A fixed k imposes a bound on the efficient combinations of y and v: in precise terms, the set

$$E(k) := \mathrm{cl}_{\mathrm{w}(Y\times V, Y'\times V')} \operatorname{vmax} \mathbb{Y}_{\mathrm{SR}}(k)$$

is $\mathrm{w}(Y\times V, Y'\times V')$-compact by PSA 9 and the Banach-Alaoglu Theorem, and it is contained in $\mathbb{Y}_{\mathrm{SR}}(k)$ by PSA 3. Since $(p,w) \geq 0$,

$$\Pi_{\mathrm{SR}}(k) = \sup_{y,v} \{\langle p \mid y \rangle - \langle w \mid v \rangle : (y,-v) \in E(k)\} \tag{22.1}$$

by PSA 8 (as part of which, $E(k) \neq \emptyset$ if $\mathbb{Y}_{\mathrm{SR}}(k) \neq \emptyset$). Since $(p,w) \in Y' \times V'$, the maximand in (22.1) is $\mathrm{w}(Y\times V, Y'\times V')$-u.s.c. (and actually continuous) in (y,v) jointly. So, by Weierstrass's Theorem, it attains its supremum on $E(k)$, which equals its supremum on $\mathbb{Y}_{\mathrm{SR}}(k)$ by (22.1). □

Proposition 22.2 (Solubility of LRC programme)**.** *Under the Production Set Assumptions 3, 6 and 7,*[108] *if* $(r,w) \in K'_+ \times V'_+$ *and the long-run cost minimization programme (3.8)–(3.9) is feasible, then it has a solution.*

Proof. Again, it is similar to the Proof of Lemma 21.4, but simpler. A fixed y imposes a bound on the efficient combinations of k and v: in precise terms, the set

$$E(y) := \mathrm{cl}_{\mathrm{w}(K\times V, K'\times V')} \operatorname{vmin} \mathbb{I}_{\mathrm{LR}}(y)$$

is $\mathrm{w}(K\times V, K'\times V')$-compact by PSA 7 and the Banach-Alaoglu Theorem, and it is contained in $\mathbb{I}_{\mathrm{LR}}(y)$ by PSA 3. Since $(r,w) \geq 0$,

$$C_{\mathrm{LR}}(y) = \inf_{k,v} \{\langle r \mid k \rangle + \langle w \mid v \rangle : (k,v) \in E(y)\} \tag{22.2}$$

by PSA 6 (as part of which, $E(y) \neq \emptyset$ if $\mathbb{I}_{\mathrm{LR}}(y) \neq \emptyset$). Since $(r,w) \in K' \times V'$, the minimand in (22.2) is $\mathrm{w}(K\times V, K'\times V')$-l.s.c. (and actually continuous) in (k,v) jointly. So, by Weierstrass's Theorem, it attains its infimum on $E(y)$, which equals its infimum on $\mathbb{I}_{\mathrm{LR}}(y)$ by (22.2). □

Proposition 22.3 (Solubility of SRC programme)**.** *Under the Production Set Assumptions 3, 10 and 11,*[109] *if* $w \in V'_+$ *and the short-run cost minimization programme (3.8)–(3.9) is feasible, then it has a solution.*

Proof. Again, it is similar to the Proof of Lemma 21.5, but simpler. A fixed (y,k) imposes a bound on the efficient and v's: in precise terms precise terms, the set

$$E(y,k) := \mathrm{cl}_{\mathrm{w}(V,V')} \operatorname{vmin} \mathbb{I}_{\mathrm{SR}}(y,k)$$

[107]Here, it suffices to assume PSA 9 for $B = \{k\}$, i.e., that $\operatorname{vmax} \mathbb{Y}_{\mathrm{SR}}(k)$ is bounded for each $k \in K$.

[108]Here, it suffices to assume PSA 7 for $B = \{y\}$, i.e., that $\operatorname{vmin} \mathbb{I}_{\mathrm{LR}}(y)$ is bounded for each $y \in Y$.

[109]Here, it suffices to assume PSA 7 for $B = \{(y,k)\}$, i.e., that $\operatorname{vmin} \mathbb{I}_{\mathrm{SR}}(y,k)$ is bounded for each $y \in Y$ and $k \in K$.

is $\mathrm{w}(V, V')$-compact by PSA 11 and the Banach-Alaoglu Theorem, and it is contained in $\mathbb{I}_{\mathrm{SR}}(y, k)$ by PSA 3. Since $w \geq 0$,

$$C_{\mathrm{SR}}(y, k) := \inf_{v} \{\langle w \mid v \rangle : v \in \mathbb{I}_{\mathrm{SR}}(y, k)\} \tag{22.3}$$

by PSA 10 (as part of which, $E(y, k) \neq \emptyset$ if $\mathbb{I}_{\mathrm{SR}}(y, k) \neq \emptyset$). Since $w \in K'$, the minimand in (22.3) is $\mathrm{w}(V, V')$-l.s.c. (and actually continuous) in v. So, by Weierstrass's Theorem, it attains its infimum on $E(y, k)$, which equals its infimum on $\mathbb{I}_{\mathrm{SR}}(y, k)$ by (22.3). □

23. Continuity of profit or cost in the quantity variables and solubility of the shadow-pricing programmes

Slater's Condition is sufficient for Type One continuity, and not just semicontinuity, of the optimal value (albeit only locally, on a neighbourhood of a particular parameter point, rather than globally as in Lemmas 21.3–21.5). Type One continuity of the primal value guarantees not only that there is no duality gap but also that a dual solution exists (and can be obtained as a cluster point of any sequence of approximate optima): see, e.g., [44, Theorem 17]. As is spelt out next, this applies to the value function $\Pi_{\mathrm{SR}}(p, \cdot, w)$ when its domain, K, carries the norm topology. (A weaker topology would not do because the effective domain of Π_{SR} is typically K_+, and to have a nonempty interior it must carry the norm topology as well as have a nonempty core a.k.a. algebraic interior.)

A similar result is given for $C_{\mathrm{LR}}(\cdot, r, w)$—but not for $C_{\mathrm{SR}}(\cdot, \cdot, w)$ because, without modifications, it would be vacuous in the cases of most interest: see a Comment at the end of this section.

Lemma 23.1 (Solubility of dual to SRP programme)**.** *Assume PSAs 8 and 9. If a $k \in K$ has a norm-neighbourhood N for which there exists a (y, v) such that $\left(y, -\widetilde{k}, -v\right) \in \mathbb{Y}$ for every $\widetilde{k} \in N$ then, for each $(p, w) \in Y^*_+ \times V^*_+$:*

1. *The concave function $\Pi_{\mathrm{SR}}(p, \cdot, w) : K \to \mathbb{R} \cup \{-\infty\}$ is finite and norm-continuous at k (and hence K^*-u.s.c. at k).*[110]
2. *So $\widehat{\partial}^*_k \Pi_{\mathrm{SR}}(p, k, w) \neq \emptyset$ (when K is paired with K^* as R). Equivalently, the fixed-input shadow-pricing programme (5.6) has a solution in the norm-dual K^*, and its value equals $\Pi_{\mathrm{SR}}(p, k, w)$.*

Proof. This is because Slater's Condition, as formulated in [44, Theorem 18 (a)] for generalized perturbed CPs, is met (when K is topologized by the norm). Spelt out, this argument means here that the concave function $\Pi_{\mathrm{SR}}(p, \cdot, w)$ is locally bounded from below (by the constant $\langle p \mid y \rangle - \langle w \mid v \rangle$, on N), so it is continuous: see, e.g., [18, 14A: Theorem], [44, Theorem 8] or [48, 5.20]. Therefore, it has a supergradient in K^* (by a version of the Hahn-Banach Theorem): see, e.g., [18, 14B], [44, Theorem 11 (a)] or [48, 5.35]. And this means, by Remark 19.8 and Lemma 19.2 that: (i) the dual (5.6) has a solution (in K^*), and (ii) $\overline{\Pi}_{\mathrm{SR}} = \Pi_{\mathrm{SR}}$ at (p, k, w). □

[110] As in Lemma 21.3, $\Pi_{\mathrm{SR}} < +\infty$ everywhere by PSAs 8 and 9 with $B = \{k\}$.

Lemma 23.2 (Solubility of dual to LRC programme)**.** *Assume PSAs 6 and 7. If a $y \in Y$ has a norm-neighbourhood N for which there exists a (k, v) such that $(\widetilde{y}, -k, -v) \in \mathbb{Y}$ for every $\widetilde{y} \in N$ then, for each $(r, w) \in K_+^* \times V_+^*$:*

1. *The convex function $C_{\mathrm{LR}}(\cdot, r, w) : Y \to \mathbb{R}_+ \cup \{+\infty\}$ is finite and norm-continuous at y (and hence Y^*-l.s.c. at y).*[111]
2. *So $\partial_y^* C_{\mathrm{LR}}(y, r, w) \neq \emptyset$ (when Y is paired with Y^* as P). Equivalently, the output shadow-pricing programme (5.5) has a solution in the norm-dual Y^*, and its value equals $C_{\mathrm{LR}}(y, r, w)$.*

Proof. This is because Slater's Condition, as formulated in [44, Theorem 18 (a)] for generalized perturbed CPs, is met (when Y is topologized by the norm). Spelt out, this means here that the convex function $C_{\mathrm{LR}}(\cdot, r, w)$ is locally bounded from above (by the constant $\langle r \mid k \rangle + \langle w \mid v \rangle$, on N), so it is continuous: see, e.g., [18, 14A: Theorem], [44, Theorem 8] or [48, 5.20]. Therefore, it has a subgradient in Y^* (by a version of the Hahn-Banach Theorem): see, e.g., [18, 14B], [44, Theorem 11 (a)] or [48, 5.35]. And this means, by Remark 19.10 and Lemma 19.4 that: (i) the dual (5.5) has a solution (in Y^*), and (ii) $\underline{C}_{\mathrm{SR}} = C_{\mathrm{SR}}$ at (y, r, w). □

Comments:

1. With $C_{\mathrm{SR}}(\cdot, \cdot, w)$ as the value function, Slater's Condition usually fails at efficient points of $Y \times K$, e.g., when k imposes an active capacity constraint on y: if $\sup_t y(t) = k$, it is impossible to maintain the constraint $y \leq k$ under small but otherwise arbitrary variations of (y, k). In conjunction with additional arguments, however, Slater's Condition may still be of use because it may hold for a modified problem (in which the effective domain C_{SR} is artificially extended): see [24].
2. That $\widehat{\partial}^* \Pi_{\#\#}(k) \neq \emptyset$, where Π means $\Pi_{\mathrm{SR}}(p, \cdot, w)$—i.e., that the dual (5.6) is soluble—can also be shown by minimizing $\langle r \mid k \rangle - \Pi_{\#}(r)$ over r: the function's sublevel sets are $\mathrm{w}(K^*, K)$-compact if Π is norm-continuous at k (i.e., if the primal value is continuous at the given primal parameter point): see, e.g., [36, 6.3.9], [42, 14.2.2 with 10.1] or [44, Theorem 10 (b)]. So a minimum point exists by Weierstrass's Theorem, and it belongs to $\widehat{\partial}^* \Pi_{\#\#}(k)$ by the Derivative Property (C.33) reoriented for concave conjugacy. The Hahn-Banach Theorem is still needed to show that there is no duality gap, i.e., that the minimum in question, $\Pi_{\#\#}(k)$, actually equals $\Pi(k)$—or, equivalently, that $\widehat{\partial}^* \Pi(k) = \widehat{\partial}^* \Pi_{\#\#}(k) \neq \emptyset$. This is a roundabout argument, but it provides a check as well as stating another result (viz., the duality between the continuity and inf-compactness properties).

24. Long-run producer optimum with conditionally fixed technical coefficients

Such technologies have already been encountered in the context of electricity: both thermal generation and pumped storage, though not hydro, are examples (Section 15). More generally, a production technique has *conditionally fixed coefficients* (c.f.c.) if the conditional input demands are price-independent, i.e., if the cost-minimizing input quantities are functions not of the input prices (r, w), but of the output bundle y alone.

[111] As in Lemma 21.4, $C_{\mathrm{LR}} \geq 0$ everywhere by PSA 5.

Denoted by $\check{k}_\phi(y)$ and $\check{v}_\xi(y)$, these are called the *input requirements* for a fixed input $\phi \in \Phi$ and a variable input $\xi \in \Xi$ (since the input requirement set is an orthant with $\left(\check{k}(y), \check{v}(y)\right)$ as its vertex). There may also be a constraint that applies to any producible output bundle in the long as in the short run (e.g., $\int y \, \mathrm{d}t = 0$ when y is the net flow from storage, as in (15.4)). In these terms, the long-run production set for a c.f.c. technique is the convex cone

$$\mathbb{Y} = \left\{(y, -k, -v) : \check{k}(y) \le k, \ \check{v}(y) \le v, \ y \in Y_0\right\} \tag{24.1}$$

where each of the (real-valued) functions $\check{k}_\phi$ and $\check{v}_\xi$ is: (i) sublinear, i.e., convex and positively linearly homogeneous (p.l.h.) on Y, and (ii) nonnegative on Y_0, which is a convex cone in the output space Y. Usually

$$Y_0 = \{y : \langle a_j \,|\, y\rangle = 0, \ b_l(y) \ge 0 \quad \text{for } j \in J, \ l \in L\} \tag{24.2}$$

where each a_j is a linear functional, and each b_l is a superlinear (p.l.h. concave) function on Y. The polar of Y_0 is then

$$Y_0^\circ = \left\{\sum_j \alpha_j a_j - \sum_l \beta_l \widehat{\partial} b_l(0) : \alpha \in \mathbb{R}^J, \ \beta \in \mathbb{R}^L, \ \beta \ge 0\right\}.$$

The finite sets J and L may both be empty (in which case $Y_0 = Y$ and $Y_0^\circ = \{0\}$). Note, also, that unless the output is a scalar (i.e., unless $Y = \mathbb{R}$), this need not be an ordinary fixed-coefficients technology: see also the Comment at the end of this section.

A direct route to characterizing a long-run producer optimum in terms of the functions $\check{k}$ and $\check{v}$ is to note that, for $r \ge 0$ and $w \ge 0$,

$$C_{\mathrm{LR}}(y, r, w) = r \cdot \check{k}(y) + w \cdot \check{v}(y) + \delta(y \mid Y_0) \tag{24.3}$$

and to use either the LRC optimization system (6.4)–(6.6) or its differential equivalent (8.8)–(8.9) or, easiest of all, the conjunction of (6.4) and (8.9). In the c.f.c. case, it is no problem to split the joint programme (3.8)–(3.9) for k and v in (6.4): the optimal k's and v's can be found separately from each other (as functions of y); equivalently, $\partial_{r,w} C_{\mathrm{LR}} = \partial_r C_{\mathrm{LR}} \times \partial_w C_{\mathrm{LR}}$. When $r \gg 0$ and $w \gg 0$ (and $y \in Y_0$), the unique optima are $\check{k}(y)$ and $\check{v}(y)$—so our use of this notation in the context of a c.f.c. technique is essentially consistent with the earlier meaning of $\check{v}(y, k, w)$ and $\check{k}(y, r, w)$ for a general technology (Sections 4 and 11 after (11.14)). However, when r and w are nonnegative but not strictly positive, the solution set for (6.4) is not just the single point $\left(\check{k}(y), \check{v}(y)\right)$: it is the Cartesian product of the sub-orthants

$$\check{K}(y, r) := \left\{k : k \ge \check{k}(y), \ r \cdot \left(k - \check{k}(y)\right) = 0\right\}$$
$$\check{V}(y, r) := \{v : v \ge \check{v}(y), \ w \cdot (v - \check{v}(y)) = 0\}.$$

So a bundle $(y, -k, -v)$ is a long-run producer optimum at prices (p, r, w) if and only if

$$k \ge \check{k}(y), \ y \in Y_0 \text{ and } v \ge \check{v}(y) \tag{24.4}$$

$$r \ge 0 \text{ and } w \ge 0 \tag{24.5}$$

$$r \cdot \left(k - \check{k}(y)\right) = 0 \text{ and } w \cdot (v - \check{v}(y)) = 0 \tag{24.6}$$

$$p \in r \partial \check{k}(y) + w \partial \check{v}(y) + \mathrm{N}(y \mid Y_0) \tag{24.7}$$

where $\mathrm{N}(y \mid Y_0)$ is the *outward normal cone* to Y_0 at y, i.e.,[112]

$$\mathrm{N}(y \mid Y_0) := \partial\delta(y \mid Y_0) = \{\lambda \in Y_0^\circ : \langle \lambda \mid y \rangle = 0\} \tag{24.8}$$
$$= \left\{ \sum_j \alpha_j a_j - \sum_{l:\, \beta_l \neq 0} \beta_l \widehat{\partial} b_l(y) : \alpha \in \mathbb{R}^J,\ \beta \in \mathbb{R}^L,\ \beta \geq 0,\ \beta \cdot b(y) = 0 \right\}$$
$$\text{and} \quad r\partial\check{k}(y) + w\partial\check{v}(y) := \sum_{\phi:\, r_\phi \neq 0} r_\phi \partial\check{k}_\phi(y) + \sum_{\xi:\, w_\xi \neq 0} w_\xi \partial\check{v}_\xi(y).$$

Comment: The qualifications $\beta_l \neq 0$, $r_\phi \neq 0$ and $w_\xi \neq 0$ in (24.8), and later in Lemma 25.1, may seem superfluous—and so they are when Y and P, the output quantity and price spaces, are finite-dimensional (because, if $\check{k}$ is a finite convex function on Y, then $\partial\check{k}(y) \neq \emptyset$, and the term $0\partial\check{k}(y) = \{0\}$ has no effect on any sum that contains it). But when P is infinite-dimensional, the P-part of the algebraic subdifferential of a finite convex function $\check{k}$ on Y can be empty (i.e., $\partial\check{k}(y) := P \cap \partial^{\mathrm{a}}\check{k}(y)$ can be the empty set $\emptyset$ even though $\partial^{\mathrm{a}}\check{k}(y) \neq \emptyset$).[113] Without the restriction to nonzero coefficients, the term $0\partial\check{k}(y) = 0\emptyset = \emptyset$ would then make the whole sum empty, instead of having no effect.

The sum (24.7) decomposes a marginal cost $p \in \partial_y C_{\mathrm{LR}}$ into the sum of operating charges and capital charges (plus a term arising from Y_0 if $Y_0 \neq Y$).

The system (24.4)–(24.7) can be recognized as the Kuhn-Tucker Optimality Conditions for any of the programme pairs—either SRP or LRC or SRC optimization together with the dual. For the case of SRP, this is proved formally in Proposition 25.3 (Section 25). The roles of the variables (p, y; r, k; w, v)—as primal/dual decisions or parameters—differ from case to case, of course (Sections 3 and 5).

Although this system is easiest to derive by using the LRC programme and function (24.3) to find (k, v) and p in terms of (y, r, w), the short-run profit approach requires inverting this map partially to find (y, v) and r (given p, k and w). Since this means solving the SRP programme with its dual, it is of interest to spell out both programmes in terms of $\check{k}$, $\check{v}$ and Y_0 (even though the primal is obvious, and the dual might be left implicit because, whatever it is, a characterization of optimality for the programme pair is already known, from (24.4)–(24.7)).

Since the short-run cost is known—it is $w \cdot \check{v}(y)$ whenever the SRC programme is feasible and $w \geq 0$—we focus on the *reduced* SRP programme, introduced in (3.13) and (4.2). Since the fixed capacities k are thought of as a plant, it is called the (reduced) *profit-maximizing plant operation* programme. It can be formulated as the following CP (an ordinary CP with an "abstract" constraint, the set Y_0):

$$\text{Given } p, k \text{ and } w \geq 0 \tag{24.9}$$

$$\text{maximize } \langle p \mid y \rangle - w \cdot \check{v}(y) \text{ over } y \tag{24.10}$$

---

[112] At 0, the normal cone equals the polar cone Y_0°. When Y_0 is a vector subspace of Y, as in (15.5), the normal cone is the same at every y: it is the annihilator space (a.k.a. orthogonal complement) $Y_0^\perp$.

[113] For example, the function $\check{k}(y) := \mathrm{EssSup}(y) := \operatorname{ess\,sup}_t y(t)$, for $y \in Y := L^\infty[0,T]$, has no subgradient in $P := L^1[0,T]$ at any y with $\mathrm{meas}\{t : y(t) = \mathrm{EssSup}(y)\} = 0$. This is because $\gamma \in L^1 \cap \partial^{\mathrm{a}}\mathrm{EssSup}(y)$ if and only if $\gamma \geq 0$, $\int_0^T \gamma(t)\,\mathrm{d}t = 1$ and $\gamma = 0$ on $\{t : y(t) < \mathrm{EssSup}(y)\}$: see, e.g., [32, 4.5.1: Example 3].

$$\text{subject to:}\quad \check{k}(y) \le k \text{ and } y \in Y_0. \tag{24.11}$$

The dual programme (5.13)–(5.14) consists in *plant valuation*; this is the standard dual of (24.9)–(24.11), and so its variables (r) are the Lagrange multipliers for the primal inequality constraints. It can be formulated as the following CP:

$$\text{Given } p, k \text{ and } w \ge 0 \tag{24.12}$$

$$\text{minimize } r \cdot k \text{ over } r \tag{24.13}$$

$$\text{subject to:}\quad r \ge 0 \tag{24.14}$$

$$p \in r\partial\check{k}(0) + w\partial\check{v}(0) + Y_0^{\circ}. \tag{24.15}$$

Formally, this is because the condition $(p, r, w) \in \mathbb{Y}^{\circ}$ of (5.14) can be expanded into (24.15) when $\mathbb{Y}$ is given by (24.1): see Lemma 25.1.

Comment (on the Kuhn-Tucker and the FFE Conditions with a c.f.c. technique):

(1) The Kuhn-Tucker Conditions on y and r to solve the *reduced* operation programme (24.9)–(24.11) and its dual (24.12)–(24.15) are also (24.4)–(24.7), but with $v = \check{v}(y)$, which makes parts of (24.4) and (24.6) redundant (and $w \ge 0$ is now an assumption needed for the reduction).

(2) For a c.f.c. technique, the FFE characterization of a solution pair (as a pair of feasible points giving equal values to the primal maximand and the dual minimand) is:[114]

$$k \ge \check{k}(y),\ y \in Y_0 \text{ and } v \ge \check{v}(y) \tag{24.16}$$

$$r \ge 0 \text{ and } w \ge 0 \tag{24.17}$$

$$p \in r\partial\check{k}(0) + w\partial\check{v}(0) + Y_0^{\circ} \tag{24.18}$$

$$\langle p \,|\, y \rangle = r \cdot k + w \cdot v. \tag{24.19}$$

This system's equivalence to the Kuhn-Tucker Conditions (24.4)–(24.7) can be seen from a variant of Euler's Theorem on p.l.h. functions (C.41): applied to each $\check{k}_{\phi}$ and $\check{v}_{\xi}$ (in place of C), it shows that the LRMC pricing condition (24.7) can be equivalently recast as the conjunction of price consistency (24.18) and the LRC recovery condition

$$\langle p \,|\, y \rangle = r \cdot \check{k}(y) + w \cdot \check{v}(y). \tag{24.20}$$

And, under the feasibility conditions ((24.16), (24.17) and (24.18)),

$$\langle p \,|\, y \rangle \le r \cdot \check{k}(y) + w \cdot \check{v}(y) \le r \cdot k + w \cdot v$$

so (24.20) and (24.6) together are equivalent to (24.19), i.e., to the equality of values at the two feasible points.

(3) The FFE Conditions on y and r to solve the *reduced* operation programmes (24.9)–(24.11) and its dual (24.12)–(24.15) are also (24.16)–(24.19), but with with $v = \check{v}(y)$.

[114] To see that (24.16)–(24.19) is indeed the FFE system, recall from (3.5) that primal and dual feasibilities mean that $(y, -k, -v) \in \mathbb{Y}$ and $(p, r, w) \in \mathbb{Y}^{\circ}$. In the c.f.c. case, the two feasibility conditions expand into (24.16) and (24.17)–(24.18).

(4) The Kuhn-Tucker Conditions (24.4)–(24.7) can also be derived by using, instead of the LRC function (24.3), the SRC function

$$C_{\mathrm{SR}}\left(y,k,w\right)=\begin{cases} w\cdot\check{v}\left(y\right) & \text{if } \check{k}\left(y\right)\leq k \text{ and } y\in Y_0 \\ +\infty & \text{otherwise} \end{cases} \tag{24.21}$$

to find v and $(p,-r)$ in terms of (y,k,w) from the conjunction of (6.8) and (8.13). When all the capacity constraints are active (i.e., $k_\phi=\check{k}_\phi\left(y\right)$ for each ϕ), subdifferentiation of (24.21) gives[115]

$$\partial_{y,k}C_{\mathrm{SR}}\left(y,k,w\right)=\left\{\left(w\partial\check{v}\left(y\right)+r\partial\check{k}\left(y\right)+\lambda,-r\right):r\geq 0,\ \lambda\in\mathrm{N}\left(y\mid Y_0\right)\right\}. \tag{24.22}$$

Since the function C_{SR} represents, by (24.21), the capacity constraints as well as the variable cost actually incurred, the sum representing the (multi-valued) SRMC in (24.22) contains capacity premia $\kappa_\phi\in r_\phi\partial\check{k}_\phi\left(y\right)$, where each r_ϕ is nonnegative but otherwise completely undetermined by pure short-run cost calculations.[116] This is the short-run counterpart of the LRMC's decomposition (24.7).

(5) The inputs of a c.f.c. technique are perfect complements, in the sense that no input substitution is possible after fixing the output bundle y.[117] With y fixed, the rate of input substitution is either undefined or completely indeterminate if regarded as multi-valued.[118] Remarkably, perfect complements can substitute for one another in product-value terms; i.e., the maximum revenue

$$\sup_y\left\{\langle p\,|\,y\rangle:(y,-k,-v)\in\mathbb{Y}\right\} \tag{24.23}$$

can have ordinary derivatives w.r.t. each input quantity, k_ϕ or v_ξ. Then, *a fortiori*, the (maximum) SRP function is also differentiable in k—and so the capital inputs have definite and separate marginal values, whose ratio is a well-defined rate of substitution $(\partial\Pi_{\mathrm{SR}}/\partial k_1:\partial\Pi_{\mathrm{SR}}/\partial k_2)$. This is so with, e.g., the storage technique (15.4) when the good's price is a continuous function of time.[119] Such a substitution between perfectly complementary inputs would, of course, be impossible with a homogeneous, one-dimensional output good: in such a case the output from an input bundle (k,v) could only have the familiar fixed-coefficients form $\min\left\{k_1,\ldots,v_1,\ldots\right\}$. But with a multi-dimensional, differentiated output good, perfect complementarity would imply fixed proportions between inputs *only* if the output proportions were fixed—and they are not. With output proportions allowed to vary, it is the output price system p that aggregates the output bundle y into a scalar revenue; and, given a suitable p, substitution in revenue terms is possible. With multiple outputs, the inputs can be perfect complements without, like a nut and bolt, having to be used in a fixed proportion.

---

[115] The term corresponding to any inactive constraint must be deleted.

[116] This term is an (outward) normal vector to the intersection of the sublevel sets of $\check{k}_\phi$'s in (24.21): see, e.g., [42, 23.7.1 and 23.8.1] or [32, 4.3: Propositions 1 and 2].

[117] This is the borderline case between Hicks-Allen complements and substitutes: see, e.g., [47, 1.F.d].

[118] Formally, the multi-valued rate of substitution equals $\mathbb{R}_+=[0,+\infty)$.

[119] Shown in [21] or [27], the result is summarized and used in Sections 16 and 17 here.

25. Derivation of dual programmes and of the Kuhn-Tucker Conditions for c.f.c. techniques (proofs for Section 24)

The cones polar and normal to the production cone of a technique with conditionally fixed coefficients are calculated next. The formulae can be used to specialize the dual programmes of Section 5 to such a technology. In particular, we prove that the dual of the SRP programme is indeed (24.12)–(24.15). We also show that the Kuhn-Tucker Conditions are indeed (24.4)–(24.7).

Lemma 25.1 (Polar and normal to production cone with c.f.c.)**.** *Assume that the production set $\mathbb{Y}$ has the form (24.1)–(24.2), where:*

(1) $\check{k}_\phi \colon Y \to \mathbb{R}$ *and* $\check{v}_\xi \colon Y \to \mathbb{R}$ *are sublinear (p.l.h. convex),* $a_j \colon Y \to \mathbb{R}$ *is linear, and* $b_l \colon Y \to \mathbb{R}$ *is superlinear (p.l.h. concave), with* $\check{k}_\phi(0) = 0$*,* $\check{v}_\xi(0) = 0$ *and* $b_l(0) = 0$ *(for each* $\phi \in \Phi$*,* $\xi \in \Xi$*,* $j \in J$ *and* $l \in L$*, which are finite sets).*
(2) *There exists a* $y_0 \in Y$ *such that* $\langle a_j \mid y_0 \rangle = 0$ *for each* $j \in J$ *and* $b_l(y_0) > 0$ *for each* $l \in L$.[120]
(3) *All but at most one of the functions* $\check{k}_\phi$*,* $\check{v}_\xi$ *and* b_l *(for* $\phi \in \Phi$*,* $\xi \in \Xi$ *and* $l \in L$*) are continuous for* $\mathrm{m}(Y, P)$*, the strongest locally convex topology that makes* P *the continuous dual of* Y*. All the linear functionals* a_j *belong to* P *(for* $j \in J$*).*[121]

Then, for every $(y, -k, -v) \in \mathbb{Y}$*,*

$$\mathrm{N}(y \mid Y_0) = \left\{ \sum_j \alpha_j a_j - \sum_{l \colon \beta_l \neq 0} \beta_l \widehat{\partial} b_l(y) : \alpha \in \mathbb{R}^J,\ \beta \in \mathbb{R}^L_+,\ \beta \cdot b(y) = 0 \right\} \tag{25.1}$$

$$Y_0^\circ = \left\{ \sum_j \alpha_j a_j - \sum_{l \colon \beta_l \neq 0} \beta_l \widehat{\partial} b_l(0) : \alpha \in \mathbb{R}^J,\ \beta \in \mathbb{R}^L_+ \right\} \tag{25.2}$$

$$\mathrm{N}(y, -k, -v \mid \mathbb{Y}) = \left\{ (p, r, w) \in P \times \mathbb{R}^\Phi_+ \times \mathbb{R}^\Xi_+ : \text{(24.6) and (24.7) hold} \right\} \tag{25.3}$$

$$\mathbb{Y}^\circ = \left\{ (p, r, w) \in P \times \mathbb{R}^\Phi_+ \times \mathbb{R}^\Xi_+ : p \in r \partial \check{k}(0) + w \partial \check{v}(0) + Y_0^\circ \right\}. \tag{25.4}$$

Proof. It is based on the additivity of subdifferentiation (C.15), applied to the normal cone operation $\mathrm{N} := \partial \delta$ as per (C.16), and on a representation of normal cones to sets of two special forms: (i) the kernel of a linear map $a = (a_j)_{j \in J} \colon Y \to \mathbb{R}^J$, and (ii) the superlevel set of a continuous concave function, such as $\{y : b_l(y) \geq 0\}$, abbreviated to $\{b_l \geq 0\}$.

Unless $a = 0$, this application of the additivity property (C.15) requires continuity of *all* the functions $\check{k}_\phi$, $\check{v}_\xi$ and b_l (since the one function allowed to be discontinuous has to be $\delta(\cdot \mid \ker a)$). Therefore, (25.1)–(25.4) are first proved in the purely algebraic form, i.e., for the *algebraic* subdifferential ∂^{a} and normal cone N^{a} (instead of $\partial := P \cap \partial^{\mathrm{a}}$ and $\mathrm{N} := P \cap \mathrm{N}^{\mathrm{a}}$). In other words, the *strongest locally convex topology*, $\mathcal{T}_{\mathrm{SLC}}$, is

[120] When y is a decision variable, as in the SRP programme, this is Slater's Condition on the constraints defining Y_0.

[121] For a linear functional a_j, $\mathrm{m}(Y, P)$-continuity is equivalent to $\mathrm{w}(Y, P)$-continuity (and it means simply that $a_j \in P$). But a concave function (b_l) or a convex function ($\check{k}_\phi$, $\check{v}_\xi$) can be $\mathrm{m}(Y, P)$-continuous (and hence $\mathrm{w}(Y, P)$-u.s.c. or l.s.c., respectively) without being $\mathrm{w}(Y, P)$-continuous. The weak and Mackey topologies are also used, and briefly discussed, in Section 20.

put on Y to start with. This makes every finite convex function continuous, and hence subdifferentiable by (C.19) when Y is paired with its algebraic dual Y^{a}. (See, e.g., [5, V.3.3 (d)] or deduce from [18, Exercise 2.10 (g)].) Since $\mathcal{T}_{\mathrm{SLC}} = \mathrm{m}\,(Y, Y^{\mathrm{a}})$, its use amounts to replacing P by Y^{a} to start with.

By Assumption 1, each b_l is $\mathcal{T}_{\mathrm{SLC}}$-continuous (everywhere on Y), so (C.16) applies to show that, for every $y \in Y_0 = \{a = 0\} \cap \bigcap_l \{b_l \geq 0\}$,

$$\mathrm{N}^{\mathrm{a}}\,(y \mid Y_0) = \mathrm{N}^{\mathrm{a}}\,(y \mid \ker a) + \sum_{l \in L} \mathrm{N}^{\mathrm{a}}\,(y \mid \{y' : b_l\,(y') \geq 0\}) \tag{25.5}$$

$$= \mathrm{span}\,\{a_j : j \in J\} - \sum_{l\colon b_l(y)=0} \mathrm{cone}\,\widehat{\partial}^{\mathrm{a}} b_l\,(y) \tag{25.6}$$

by: (i) the Factorization Lemma (a.k.a. Sard's Quotient Theorem) given in, e.g., [18, 21A] and [32, 0.1.4: Corollary],[122] and (ii) the formula for the normal cone to a sublevel set of a convex function (reoriented to a concave function's superlevel set), which is given in, e.g., [32, 4.3: Proposition 2], [42, 23.7.1] and [48, 7.8]. In other words, $p \in \mathrm{N}^{\mathrm{a}}\,(y \mid Y_0)$ if and only if $p = \sum_j \alpha_j a_j - \sum_l \beta_l \widehat{\partial}^{\mathrm{a}} b_l\,(y)$ for some α and $\beta \geq 0$ with $\beta \cdot b\,(y) = 0$. This proves (25.1); and (25.2) is a case of (25.1) for $y = 0$ (since $Y_0^{\circ\mathrm{a}} = \mathrm{N}^{\mathrm{a}}\,(0 \mid Y_0)$).

Note that both the decomposition (25.5) and the representation (25.6) rely on Assumption 2. First, it guarantees that all but one of the sets in question have a common interior point that also lies in the other set: $y_0 \in \{b_l > 0\} = \mathrm{cor}\,\{b_l \geq 0\} = \mathrm{int}_{Y, \mathcal{T}_{\mathrm{SLC}}}\,\{b_l \geq 0\}$, and $y_0 \in \ker a = \bigcap_j \ker\,(a_j)$.[123] So (C.16) applies to give (25.5). Second, the existence of a point y_0 satisfying the inequality strictly is what validates the formula for the normal cone, which gives the second term in (25.6).

The same arguments apply to $\mathbb{Y}$, which is the intersection of $Y_0 \times \mathbb{R}^{\Phi} \times \mathbb{R}^{\Xi}$ with the sublevel sets of $\check{k}_\phi\,(y) - k_\phi$, etc., as functions of $(y, -k, -v)$. Their subdifferentials are: $\partial^{\mathrm{a}} \check{k}_\phi \times \{(0, \ldots, 1, 0, \ldots)\} \times \{(0, \ldots, 0)\}$, etc. And the $\mathrm{card}\,\Phi + \mathrm{card}\,\Xi$ sublevel sets do have a common interior point that lies in $Y_0 \times \mathbb{R}^{\Phi} \times \mathbb{R}^{\Xi}$: e.g., $(0, -k, -v)$ with any $(k, v) \gg 0$ will do (since each $\check{k}_\phi$ or $\check{v}_\xi$ is $\mathcal{T}_{\mathrm{SLC}}$-continuous by Assumption 1). So, for every $(y, -k, -v) \in \mathbb{Y}$, one has $(p, r, w) \in \mathrm{N}^{\mathrm{a}}\,(y, -k, -v \mid \mathbb{Y})$ if and only if: $r \geq 0$, $w \geq 0$, $r \cdot \left(k - \check{k}\,(y)\right) = 0$, $w \cdot (v - \check{v}\,(y)) = 0$ and $p \in r\partial^{\mathrm{a}} \check{k}\,(y) + w\partial^{\mathrm{a}} \check{v}\,(y) + \mathrm{N}^{\mathrm{a}}\,(y \mid Y_0)$. This proves (25.3); and (25.4) is a case of (25.3) for $y = 0$, $k = 0$ and $v = 0$.

Now that (25.1)–(25.4) have been proved for $P = Y^{\mathrm{a}}$, their extension to any P follows from Assumption 3 and the fact that if vectors $p_2, \ldots, p_n$ all belong to P then: $p_1 + \ldots + p_n \in P$ if and *only* if $p_1 \in P$. By Assumption 3 and (C.19), all but at most one of the algebraic subdifferentials—$\partial^{\mathrm{a}} \check{k}_\phi\,(y)$, $\partial^{\mathrm{a}} \check{v}_\xi\,(y)$, $\widehat{\partial}^{\mathrm{a}} b_l\,(y)$—lie wholly in P. So (with the argument y suppressed in the intermediate sums)

$$P \cap \left(\sum_\phi r_\phi \partial^{\mathrm{a}} \check{k}_\phi\,(y) + \sum_\xi w_\xi \partial^{\mathrm{a}} \check{v}_\xi\,(y) + \sum_j \alpha_j a_j - \sum_l \beta_l \widehat{\partial}^{\mathrm{a}} b_l\,(y) \right)$$

[122] [42, 22.3.1] and [48, 4.19] give only Farkas's Lemma, but this contains the Factorization Lemma.

[123] For the strongest locally convex topology, the interior of a *convex* set equals the entire core, i.e., if a $Z \subseteq Y$ is convex then $\mathrm{int}_{Y, \mathcal{T}_{\mathrm{SLC}}}\, Z = \mathrm{cor}\, Z$: see, e.g., [5, V.3.3 (b)] or [18, Exercise 2.10 (g)].

$$= \sum_{\phi} P \cap r_\phi \partial^{\mathrm{a}} \check{k}_\phi + \sum_{\xi} P \cap w_\xi \partial^{\mathrm{a}} \check{v}_\xi + \sum_j \alpha_j a_j - \sum_l P \cap \beta_l \widehat{\partial}^{\mathrm{a}} b_l$$
$$= \sum_{\phi \colon r_\phi \neq 0} r_\phi \left(P \cap \partial^{\mathrm{a}} \check{k}_\phi \right) + \sum_{\xi \colon w_\xi \neq 0} w_\xi \left(P \cap \partial^{\mathrm{a}} \check{v}_\xi \right) + \sum_j \alpha_j a_j - \sum_{l \colon \beta_l \neq 0} \beta_l \left(P \cap \widehat{\partial}^{\mathrm{a}} b_l \right)$$
$$= \sum_{\phi \colon r_\phi \neq 0} r_\phi \partial \check{k}_\phi (y) + \sum_{\xi \colon w_\xi \neq 0} w_\xi \partial \check{v}_\xi (y) + \sum_j \alpha_j a_j - \sum_{l \colon \beta_l \neq 0} \beta_l \widehat{\partial} b_l (y)$$

where: the last (third) equality holds by definition ($\partial := P \cap \partial^{\mathrm{a}}$), and the penultimate (second) equality holds because $P \cap \varrho D = \varrho (P \cap D)$ for every $D \subseteq Y^{\mathrm{a}}$ and every real number $\varrho \neq 0$,[124] whilst $P \cap 0D = \{0\}$ for every $D \neq \emptyset$; this is applied to $D = \partial^{\mathrm{a}} \check{k}_\phi (y)$, $\partial^{\mathrm{a}} \check{v}_\xi (y)$, $\widehat{\partial}^{\mathrm{a}} b_l (y)$. This shows that (25.3)–(25.4) hold as stated (i.e., also when $P \neq Y^{\mathrm{a}}$ in $\partial := P \cap \partial^{\mathrm{a}}$ and $\mathrm{N} := P \cap \mathrm{N}^{\mathrm{a}}$).

With all the functions $\check{k}_\phi$ and $\check{v}_\xi$ left out (or replaced by zeros), the same arguments derive (25.1)–(25.2) for a general space P from the case of $P = Y^{\mathrm{a}}$. □

The formula for $\mathbb{Y}^\circ$ can be used to spell out all the dual programmes (when $\mathbb{Y}$ is a c.f.c. technique).

Corollary 25.2 (Dual to SRP programme with c.f.c.)**.** *On the assumptions of Lemma 25.1, the dual to the profit-maximizing operation programme (24.9)–(24.11), with k as the primal parameter, is the plant valuation programme (24.12)–(24.15), with Y_0° given by (25.2).*

Proof. Apply Proposition 18.1 and (25.4) with (25.2). □

As has already been noted (in a Comment in Section 24), the formula (25.4) for $\mathbb{Y}^\circ$ shows also that the system (24.16)–(24.19) is the FFE characterization of a solution pair (to the SRP or LRC or SRC programme together with its dual). And, by using Euler's Theorem, this FFE characterization has been proven equivalent to (24.4)–(24.7). What still remains to be shown is that the latter system, which has already been referred to as the Kuhn-Tucker Conditions, is indeed an expansion of the Kuhn-Tucker Lagrangian saddle-point condition. This is next done for the SRP programme (the LRC and SRC cases being similar). The identification of (24.4)–(24.7) as the saddle-point condition—which is known, from general theory, to be equivalent to optimality and absence of a duality gap—will also reprove its equivalence to (24.16)–(24.19) as a case of the equivalence between Kuhn-Tucker and FFE Conditions (instead of the earlier, problem-specific argument based on Euler's Theorem).

[124]For $\varrho = 0$, this fails if and only if $D \neq \emptyset$ but $P \cap D = \emptyset$ (in which case $P \cap 0D = \{0\}$ but $0 (P \cap D) = \emptyset$).

For the profit-maximizing operation programme (24.9)–(24.11), the Lagrange function (of the primal variable y and the dual variable r) is[125]

$$\mathcal{L}(y,r) := \begin{cases} \langle p \mid y \rangle - w \cdot \check{v}(y) + r \cdot \left(k - \check{k}(y)\right) & \text{if } y \in Y_0 \text{ and } r \geq 0 \\ +\infty & \text{if } y \in Y_0 \text{ and } r \ngeq 0 \\ -\infty & \text{if } y \notin Y_0 \end{cases} . \tag{25.7}$$

Proposition 25.3 (Saddle-point condition for SRP programme with c.f.c.)**.** *On the assumptions of Lemma 25.1, and given any $(p,k,w) \in P \times \mathbb{R}_+^\Phi \times \mathbb{R}_+^\Xi$, the following conditions on a pair $(y,r) \in Y \times \mathbb{R}^\Phi$, are equivalent to one another:*

(1) *y and r are solutions of equal value to the programmes of profit-maximizing operation (24.9)–(24.11) and of plant valuation (24.12)–(24.15).*
(2) *(y,r) is a saddle point (maximum-minimum point) of the Lagrange function $\mathcal{L}$ defined by (25.7), i.e., $0 \in \widehat{\partial}_y \mathcal{L}(y,r)$ and $0 \in \partial_r \mathcal{L}(y,r)$.*
(3) *(y,r), together with $v = \check{v}(y)$, meets Conditions (24.4)–(24.7).*

Proof. A (y,r) is a pair of solutions with equal values if and only if it is a saddle point of $\mathcal{L}$: see, e.g., [44, Theorem 15 (e) and (f)]. So Conditions 1 and 2 are equivalent.

Next, note that: if $0 \in \widehat{\partial}_y \mathcal{L}$ then $y \in Y_0$; and if $0 \in \partial_r \mathcal{L}$ then $r \geq 0$. So the task is to show that Conditions 2 and 3 are equivalent when $y \in Y_0$ and $r \geq 0$. The inclusion $0 \in \partial_r \mathcal{L}$ then translates into: $k \geq \check{k}(y)$ and $k_\phi = \check{k}_\phi(y)$ if $r_\phi > 0$, which are (24.4)–(24.6). And (24.7) comes from expanding the inclusion

$$0 \in \widehat{\partial}_y \mathcal{L}(y,r) = p - \partial\left(r \cdot \check{k} + w \cdot \check{v} + \delta(\cdot \mid Y_0)\right)(y).$$

It remains to be shown that this sum can be subdifferentiated term by term (and then apply (25.1) to expand $\partial\delta(y \mid Y_0) = \mathrm{N}(y \mid Y_0)$). This is done in the same way as in the Proof of Lemma 25.1. First, note that

$$\begin{aligned} \partial^{\mathrm{a}}\left(r \cdot \check{k} + w \cdot \check{v} + \delta(\cdot \mid Y_0)\right)(y) &= \partial^{\mathrm{a}}\left(r \cdot \check{k}\right)(y) + \partial^{\mathrm{a}}(w \cdot \check{v})(y) + \partial^{\mathrm{a}}\delta(y \mid Y_0) \\ &= r\partial^{\mathrm{a}}\check{k}(y) + w\partial^{\mathrm{a}}\check{v}(y) + \mathrm{N}^{\mathrm{a}}(y \mid Y_0) \end{aligned} \tag{25.8}$$

for every $y \in Y_0$ (for $y \notin Y_0$, both sides equal $\emptyset$). This is because each $\check{k}_\phi$ or $\check{v}_\xi$, being a finite convex function, is $\mathcal{T}_{\mathrm{SLC}}$-continuous (everywhere on Y). The only other function, $\delta(\cdot \mid Y_0)$, may have no point of continuity, but this does not matter: since all but one of these functions are continuous, the algebraic subdifferential ∂^{a} is an additive operator by (C.15). And it is p.l.h. by (C.17) with (C.19).

The additivity of $\partial := P \cap \partial^{\mathrm{a}}$ follows from that of ∂^{a} by using Assumption 3 on the functions involved, as in the Proof of Lemma 25.1. This is only sketched. Say, for simplicity, that $L = \emptyset$, i.e., $Y_0 = \ker a$ and so $Y_0^\circ = \mathrm{span}\{a\} \subseteq P$. All but at most one of the sets $\partial^{\mathrm{a}}\check{k}_\phi(y)$ and $\partial^{\mathrm{a}}\check{v}_\xi(y)$ also lie wholly in P, by (C.19). So the sum of their elements (one from each set) belongs to P if and *only* if each term does. (If $p_2, \ldots, p_n$ all belong to P then: $p_1 + \ldots + p_n \in P$ if and *only* if $p_1 \in P$.) This means that (25.8) holds also with $\partial := P \cap \partial^{\mathrm{a}}$ in place of ∂^{a}. ☐

[125] This is a case of the ordinary Lagrangian (with Y_0 as an "abstract" constraint, unpriced by $\mathcal{L}$): see, e.g., [44, (4.4)], where it is derived from the generalized Lagrangian defined in [44, (4.2)].

Comments:

(1) Instead of deriving the dual programme (24.12)–(24.15) straight from the primal (as in the Proof of Proposition 18.1, applied in the earlier Proof of Corollary 25.2), one can obtain the dual through the Lagrange function—since the dual to a maximization programme consists in minimizing, over the dual variables, the supremum of the Lagrange function over the primal variables: see, e.g., [44, (4.6) or Example 1': (5.1)]. Here, this means minimizing $\sup_y \mathcal{L}(y, r)$ over r. Denote[126]

$$\Pi_{\mathrm{Exc}}(y) := \langle p \mid y \rangle - w \cdot \check{v}(y) - r \cdot \check{k}(y).$$

For $r \geq 0$, $\sup_y \mathcal{L}$ equals $r \cdot k + \sup_{y \in Y_0} \Pi_{\mathrm{Exc}}$. Since Π_{Exc} is p.l.h. in y, its supremum is either 0 or $+\infty$, and it is 0 if and only if $0 \in \partial_y \Pi_{\mathrm{Exc}}(0) + \partial \delta(0 \mid Y_0)$. This inclusion translates into (24.15). The additivity of ∂ must be verified as before (in the Proof of Lemma 25.1).

(2) Three variations on the above Lagrangian $\mathcal{L}$ are possible but only one of them is useful:

(a) Although it is simpler to reduce the problem by using the obvious cost-minimizing solution $v = \check{v}(y)$ for any variable inputs, one could retain the constraint $v \leq \check{v}(y)$ and apply the Lagrangian method to the joint SRP programme for y and v. Within the intrinsic parameterization (i.e., when only p, y, r, k, w, v serve as parameters and variables), this inequality could only be treated as another unparameterized "abstract" constraint like $y \in Y_0$ (since y and v are variables and not parameters of the SRP programme). The resulting Lagrangian would be just like (25.7), only with $v \leq \check{v}(y)$ adjoined to $y \in Y_0$ as another abstract constraint; and the Kuhn-Tucker Conditions would be the same, viz., (24.4)–(24.7).

(b) The parameterization could of course be extended by rewriting this constraint as $v - \check{v}(y) \leq \zeta$, with ζ as an extrinsic primal parameter varying around 0 and paired with a Lagrange multiplier m, say. But this would only needlessly complicate the Lagrangian to: $\mathcal{L}(y, v; r, m) = \langle p \mid y \rangle - w \cdot v + m \cdot (v - \check{v}(y)) + r \cdot \left(k - \check{k}(y)\right)$ for $y \in Y_0$ and $r \geq 0$. At a saddle point, $m = w$ from the FOC that $0 = \nabla_v \mathcal{L}$—which reduces $\mathcal{L}$ back to (25.7).

(c) By contrast, it would be sensible to parameterize the constraints defining Y_0 in (24.2) to have

$$\langle a_j \mid y \rangle = \zeta'_j \text{ and } b_l(y) \geq \zeta''_l. \tag{25.9}$$

In any programme for y subject to $y \in Y_0$, this would give a marginal-value interpretation to the coefficients α_j and β_l in (25.1), since these would be the extrinsic Lagrange multipliers paired with the extrinsic parameters ζ'_j and ζ''_l. In particular, for the profit-maximizing operation programme (24.9)–(24.11), this would mean that $\partial \Pi_{\mathrm{SR}} / \partial \zeta'_j = \alpha_j$ and $\partial \Pi_{\mathrm{SR}} / \partial \zeta''_l = \beta_l$.[127] For example,

[126] $\Pi_{\mathrm{Exc}}(y)$ is the excess a.k.a. pure profit from the output y (i.e., revenue at prices p less minimum input cost at prices r and w).

[127] The partial derivatives $(\partial \Pi_{\mathrm{SR}} / \partial \zeta)$ exist if the α_j and β_l associated by (25.1) with the optimal y are unique. If not, the derivative property still holds for the superdifferential, i.e., $\widehat{\partial}_{\zeta', \zeta''} \Pi$ contains each (α, β) that satisfies (25.1) for some optimal y (and hence for every optimal y).

in the case of the storage technology (15.4), the constraint $\int y(t)\,\mathrm{d}t = 0$ is varied to $\int y(t)\,\mathrm{d}t = \zeta$ to interpret the constant term, λ, of the good's price as $\partial\Pi_{\mathrm{SR}}/\partial\zeta$ (at $\zeta = 0$); the price decomposition (16.9) is a case of (24.7) with (24.8).

(3) Instead of obtaining $\mathbb{Y}^{\circ}$ from a formula for $\mathrm{N}(\cdot \mid \mathbb{Y})$ evaluated at $(0,0,0)$ as in the Proof of Lemma 25.1, one can calculate the polar directly: from (3.4) and (24.1), $(p,r,w) \in \mathbb{Y}^{\circ}$ if and only if the conditions $y \in Y_0$, $\check{k}(y) \le k$ and $\check{v}(y) \le v$ imply that $\langle p \mid y \rangle \le r \cdot k + w \cdot v$. This, in turn, is equivalent to: $r \ge 0$, $w \ge 0$, and $\langle p \mid y \rangle \le r \cdot \check{k}(y) + w \cdot \check{v}(y) + \delta(y \mid Y_0)$ for every y. Since $\check{k}$, $\check{v}$ and $\delta(\cdot \mid Y_0)$ all vanish at $y = 0$, the last inequality can be restated as: $p \in \partial\left(r\check{k} + w\partial\check{v} + \partial\delta(\cdot \mid Y_0)\right)(0)$. To obtain (25.4), this sum is subdifferentiated term by term (as is done above from (25.8) on).

26. Verification of assumptions for techniques with conditionally fixed coefficients

To apply the results of Sections 21 to 23 (which rule out duality gaps and ensure that both primal and dual programmes are soluble), one needs to verify the Production Set Assumptions of Section 21. This is next done for c.f.c. techniques, when the output space Y is the dual of a Banach lattice Y'.

Lemma 26.1 (Properties of production set with c.f.c.). *Assume that $\mathbb{Y}$ is given by (24.1), i.e., that $(y, -k, -v) \in \mathbb{Y}$ if and only if*

$$\check{k}(y) \le k, \ \check{v}(y) \le v \text{ and } y \in Y_0$$

where:[128] *$\check{k}\colon Y \to \mathbb{R}^{\Phi}$ and $\check{v}\colon Y \to \mathbb{R}^{\Xi}$ are sublinear maps (with $\check{k}(0) = 0$ and $\check{v}(0) = 0$) that are nondecreasing and nonnegative on Y_0, which is a convex cone in Y (and Φ and Ξ are finite sets). Then:*

(1) *$\mathbb{Y}$ satisfies PSAs 1, 2, 4, 5, 6 and 10.*

(2) *If $\check{k}_\phi$ and $\check{v}_\xi$ are $\mathrm{w}(Y, Y')$-lower semicontinuous (for each $\phi \in \Phi$ and $\xi \in \Xi$) and Y_0 is $\mathrm{w}(Y, Y')$-closed, then $\mathbb{Y}$ satisfies PSA 3 (i.e., it is also weakly* closed).*

(3) *If $\check{k}$ and $\check{v}$ are norm-continuous (on Y), then $\mathbb{Y}$ satisfies PSA 7 (and hence also PSA 11).*

(4) *Under the assumptions of Part 2, if the set $\left\{y \in Y_0 : \check{k}(y) \le k\right\}$ is bounded for each k,*[129] *then $\mathbb{Y}$ satisfies PSA 8.*
If additionally $\check{v}_\xi$ is norm-continuous (for each $\xi \in \Xi$),[130] *then $\mathbb{Y}$ satisfies PSA 9 (and hence also PSA 11).*

(5) *$\mathbb{Y}$ satisfies PSA 8 also when either (a) Y_0 is a vector subspace with $Y_0 \cap Y_+ = \{0\}$, or (b) for some $\xi \in \Xi$, the function $\check{v}_\xi$ is increasing on Y_0 (i.e., $\check{v}_\xi(y') < \check{v}_\xi(y'')$ whenever $y' < y''$, for y' and y'' in Y_0).*

---

[128] In other words, each $\check{k}_\phi$ or $\check{v}_\xi$ is a p.l.h. convex finite function.

[129] When each $\check{k}_\phi$ is weakly* l.s.c. and Y_0 is weakly* closed as in Part 2, this is equivalent to weak* compactness of $\left\{y \in Y_0 : \check{k}(y) \le k\right\}$.

[130] This assumption holds vacuously when $\Xi = \emptyset$ (i.e., when there are no variable inputs, as with the storage and hydro techniques (15.4) and (15.9)).

Proof. Part 1 is obvious: PSAs 1 and 2 hold (i.e., $\mathbb{Y}$ is a convex cone) because $\check{k}$ and $\check{v}$ are sublinear and Y_0 is a convex cone. PSA 4 holds because $\check{k}$ and $\check{v}$ are nondecreasing on Y_0 (and because $\mathrm{proj}_Y(\mathbb{Y}) = Y_0$ here). PSA 5 (with $K = \mathbb{R}^\Phi$, $V = \mathbb{R}^\Xi$) holds because $\check{k}$ and $\check{v}$ are nonnegative on Y_0. Finally, PSAs 6 and 10 are verified at $\widetilde{k} = \check{k}(y)$ and $\widetilde{v} = \check{v}(y)$.

Part 2 is also obvious: the l.s. continuity of $\check{k}$ and $\check{v}$ mean that their sublevel sets are closed.

For Part 3, note that $\mathrm{vmin}\,\mathbb{I}_{\mathrm{LR}}(y) = \{(\check{k}(y), \check{v}(y))\}$ if $y \in Y_0$ (and if not, then $\mathbb{I}_{\mathrm{LR}}(y) = \emptyset$). By their norm-continuity, $\check{k}$ and $\check{v}$ are bounded on some ball centered at the origin. It follows, by their p.l. homogeneity, that $\check{k}$ and $\check{v}$ are bounded on every ball in Y—i.e., PSA 7 holds.

For Part 4, given a $y \in Y_0$ with $\check{k}(y) \le k$, take the point $(y, -\check{v}(y))$. It is itself efficient (maximal) if the set

$$\{y' \in Y_0 : y' \ge y,\ \check{k}(y') \le k,\ \check{v}(y') = \check{v}(y)\} \tag{26.1}$$

has no element other than y. But even if it has, the method of Lemma 21.1 applies. This is because the set (26.1), after embedding it in $Y \times V$ by taking its Cartesian product with $\{-\check{v}(y)\}$, is here the set (21.2) with $v = \check{v}(y)$, and it is bounded (being contained in the set that is bounded by assumption)—so Part 1 of Corollary 21.2 applies. So PSA 8 holds. To verify PSA 9, note that every bounded $B \subset K$ is bounded from above by some $\overline{k}$ (since K is finite-dimensional). For each $k \in B$, every point of $\mathrm{vmax}\,\mathbb{Y}_{\mathrm{SR}}(k)$ has the form $(y, -\check{v}(y))$ for some $y \in Y_0$ with $\check{k}(y) \le \overline{k}$. So $\bigcup_{k \in B} \mathrm{vmax}\,\mathbb{Y}_{\mathrm{SR}}(k)$ is bounded (since $\{y \in Y_0 : \check{k}(y) \le \overline{k}\}$ is bounded by assumption, and since $\check{v}$, being p.l.h. and norm-continuous, is bounded on every bounded set).

In Part 5, $(y, -\check{v}(y))$ itself is always efficient: in Case (a), no point of Y_0 is greater than y; and in Case (b), if y' is a point of Y_0 greater than y, then $\check{v}(y')$ is greater than $\check{v}(y)$. □

27. Existence of optimal operation and plant valuation and their equality to marginal values for c.f.c. techniques

The foregoing analysis (Sections 19, 20, 21, 22, 23) is next specialized to the SRP programme and its dual for a technique with conditionally fixed coefficients. As in the preceding Section 26 (and in Sections 21, 22, 23), the output space Y is the dual of a Banach lattice Y' (and Y^* is the norm-dual of Y, with $Y' \subseteq Y^*$).

> **Notation:** The optimal solution sets for programmes (24.9)–(24.11) and (24.12)–(24.15) are denoted by $\hat{Y}(p, k, w)$ and $\hat{R}(p, k, w)$, respectively. The corresponding lowercase notation, $\hat{y}$ or $\hat{r}$, is used *only* when the solution is known to be unique.

Proposition 27.1 (Hotelling's Lemma and solubility of SRP programme with c.f.c.)**.** *Assume that the production set $\mathbb{Y}$ is given by (24.1), where the input requirement functions $\check{k}_\phi\colon Y \to \mathbb{R}$ and $\check{v}_\xi\colon Y \to \mathbb{R}$ are $\mathrm{w}(Y, Y')$-lower semicontinuous (for each $\phi \in \Phi$ and $\xi \in \Xi$), and the output constraint cone Y_0 is $\mathrm{w}(Y, Y')$-closed (so $\mathbb{Y}$ is weakly* closed). Then:*

(1) *For every $p \in Y^*$, $k \geq 0$ and $w \geq 0$,*[131]

$$\hat{Y}(p,k,w) = \partial_p \Pi_{\mathrm{SR}}(p,k,w). \tag{27.1}$$

Also, (21.7) applies, i.e., $\Pi_{\mathrm{SR}}(\cdot,k,w)$ can be restricted to Y' for the calculation of $\partial_p \Pi_{\mathrm{SR}}$ at a $p \in Y'$.

(2) *If $k \geq 0$ and $\left\{y \in Y_0 : \check{k}(y) \leq k\right\}$, the feasible set of the operation programme (24.11)–(24.9), is norm-bounded then $\Pi_{\mathrm{SR}}(p,k,w)$ is finite, for every $p \in Y^*$ and $w \geq 0$. If additionally $p \in Y'$ then*

$$\hat{Y}(p,k,w) \neq \emptyset \tag{27.2}$$

for every $w \geq 0$, i.e., the profit-maximizing operation problem (24.9)–(24.11) has a proper solution.

Proof. The assumptions on $\check{k}$, $\check{v}$ and Y_0 imply that, being given by (24.21), the proper convex function $C_{\mathrm{SR}}(\cdot,k,w)$ is l.s.c. for the weak* topology $\mathrm{w}(Y,Y')$.[132] So (27.1) follows from Lemma 19.21, i.e., from (C.12) and (C.2). Furthermore, (21.7) applies with $C_{\mathrm{SR}}(\cdot,k,w)$ as C. This proves Part 1.

For Part 2, $\Pi_{\mathrm{SR}}(p,k,w) > -\infty$ because, for every $k \in \mathbb{R}^{\Phi}_{+}$, the operation programme (24.9)–(24.11) is feasible. Since the feasible set is norm-bounded by assumption,[133]

$$\Pi_{\mathrm{SR}}(p,k,w) \leq \sup_{y \in Y_0} \left\{\langle p \,|\, y\rangle : \check{k}(y) \leq k\right\} - \inf_{y \in Y_0} \left\{w \cdot \check{v}(y) : \check{k}(y) \leq k\right\} < +\infty.$$

(The infimum is of course nonnegative if $\check{v} \geq 0$ on Y_0, but it is finite in any case because each $\check{v}_\xi$ is weakly* l.s.c.)

Solubility (27.2) can be deduced from Proposition 22.1; its assumptions can be verified by applying Parts 2 and 4 of Lemma 26.1. This requires assuming that $\check{v}$ is norm-continuous (as well as weakly* l.s.c.). But the norm-continuity of $\check{v}$ is actually unnecessary because Weierstrass's Theorem applies directly: a maximum point exists because (i) the maximand of (24.10) is weakly* u.s.c. (since $p \in Y'$), and (ii) the feasible set is weakly* compact and nonempty (since the point $y = 0$ is feasible). □

If $p \in Y'$ or $k \gg 0$ (Slater's Condition), then $\Pi_{\mathrm{SR}}(p,\cdot,w)$ is u.s.c. on $K = \mathbb{R}^{\Phi}$ or continuous at k, respectively (Lemmas 21.3 and 23.1). Under either assumption, there is no duality gap between the profit-maximizing operation and plant valuation programmes, (24.9)–(24.11) and (24.12)–(24.15). It follows that the optimal shadow prices for the fixed inputs are their profit-imputed marginal values; this is spelt out next.

Proposition 27.2 (Dual Hotelling Lemma and solubility of FIV programme with c.f.c.). *In addition to the assumptions of Proposition 27.1 on $\check{k}_\phi$, $\check{v}_\xi$ and Y_0 (viz., that each $\check{k}_\phi$ and $\check{v}_\xi$ is weakly* l.s.c. and that Y_0 is weakly* closed), assume that each $\check{v}_\xi$ is norm-continuous, and that the set $\left\{y \in Y_0 : \check{k}(y) \leq k\right\}$ is norm-bounded. Then:*

[131] Formally, (27.1) holds also when $k \not\geq 0$: in this case, $\hat{Y} = Y = \partial_p \Pi_{\mathrm{SR}}$ (the programme (24.9)–(24.11) is then infeasible, so every y is an improper solution, and $\Pi_{\mathrm{SR}}(\cdot,k,w) = -\infty$).

[132] When $\check{k}$ and $\check{v}$ are norm-continuous, the l.s. continuity of C_{SR} (on $Y \times K$) can also be deduced by using Lemma 26.1 to verify the assumptions of Lemma 21.5.

[133] Being also weakly* closed, the set $\left\{y \in Y_0 : \check{k}(y) \leq k\right\}$ is actually weakly* compact by the Banach-Alaoglu Theorem.

(1) *If* $p \in Y'$ *or* $k \gg 0$ *(i.e.,* $k_\phi > 0$ *for each* $\phi \in \Phi$*) then, for every* $w \geq 0$*,*

$$\hat{R}(p,k,w) = \widehat{\partial}_k \Pi_{\mathrm{SR}}(p,k,w)\,. \tag{27.3}$$

(2) *If* $k \gg 0$ *then* $\Pi_{\mathrm{SR}}(p,\cdot,w)$ *is continuous at* k*, and so*

$$\hat{R}(p,k,w) \neq \emptyset \tag{27.4}$$

for every $p \in Y^*$ *and* $w \geq 0$*. (This means that the fixed-input value minimization programme (24.12)–(24.15) has a proper solution, since its value* $\overline{\Pi}_{\mathrm{SR}}(p,k,w)$ *is finite.)*

Proof. The operation programme (24.9)–(24.11) is feasible, i.e., $\Pi_{\mathrm{SR}}(p,k,w) > -\infty$ for every $k \in \mathbb{R}^{\Phi}_{+}$. If $p \in Y'$ then $\Pi_{\mathrm{SR}}(p,\cdot,w)$ is u.s.c. by Lemma 21.3; its assumptions are verified by applying Parts 2 and 4 of Lemma 26.1. If $k \gg 0$ then $\Pi_{\mathrm{SR}}(p,\cdot,w)$ is continuous at k, by Part 1 of Lemma 23.1. In either case, $\Pi_{\mathrm{SR}}(p,k,w) = \overline{\Pi}_{\mathrm{SR}}(p,k,w)$ by Lemma 20.1. So (27.3) follows from Lemma 19.2 with Remark 19.8 (as in the Proof of Part 1 of Corollary 19.19). This proves Part 1.

For Part 2, since $k \gg 0$, (27.4) follows from (27.3) and Part 2 of Lemma 23.1. This means that the valuation programme (24.12)–(24.15) has a proper solution, provided that it is feasible, i.e., that $+\infty > \Pi_{\mathrm{SR}}(p,k,w) = \overline{\Pi}_{\mathrm{SR}}(p,k,w)$—which is the case here (see the Proof of Proposition 27.1). □

28. Linear programming with c.f.c. techniques

The original description of a c.f.c. technique's production set $\mathbb{Y}$ need not be in terms of input requirement functions as in (24.1). Indeed, a sublinear requirement function $\check{k}_\phi$ can arise from summarizing, in a single scalar constraint, a set of linear inequality constraints (i.e., a multi- or infinite-dimensional linear inequality constraint). For example, a capacity k_ϕ may constrain the output rate to a $y(t) \leq k_\phi$ at any time t, and this can be summarized as $k_\phi \geq \check{k}_\phi(y) := \sup_t y(t)$—as in (15.2) for the case of thermal generation. Another example is the storage capacity requirement $\check{k}_{\mathrm{St}}(y)$ of (15.6), which is used in (15.8) to summarize the continuum of reservoir constraints of (15.4).[134] In other words, the profit or cost optimization problem for a c.f.c. technique can typically be formulated as an LP from the start (as we do for peak-load pricing in Section 16). With continuous time, there is a continuum of decision variables and a continuum of capacity constraints, so the LP is doubly infinite. The sublinear representation (24.1) of $\mathbb{Y}$ provides the alternative framework of a nonlinear CP with a continuum of decision variables but with only a finite number of constraints. Its usefulness depends on the availability of tractable formulae for $\check{k}$ and $\check{v}$—such as (15.6)–(15.7), which make the CP workable in our study of pumped storage [21]. But a clear advantage of formulating the profit or cost problem as an LP is that routines such as the simplex algorithm can be applied (after discretization); such methods solve the primal LP and its (standard) dual simultaneously.

[134] Similarly, if a unit output requires a unit of a costlessly storable variable input, whose total amount available, v_ξ, can be spread as an input flow $\widetilde{v}_\xi(\cdot)$ over the period, then the output rate is constrained to a nonnegative $y(t) \leq \widetilde{v}_\xi(t)$ for some $\widetilde{v}_\xi(t) \geq 0$ with $\int \widetilde{v}_\xi(t)\,\mathrm{d}t = v_\xi$. This can be summarized in the single constraint $v_\xi \geq \check{v}_\xi(y) := \int y(t)\,\mathrm{d}t$.

Even if it is not an LP originally, the profit or cost problem for a c.f.c. technique can always be reformulated as an LP: a sublinear inequality constraint on y can be converted to an equivalent system of linear constraints by using the "convex variant" of Euler's Theorem on p.l.h. functions, stated here as (C.39). Each condition $\check{k}_\phi(y) \le k_\phi$ in (24.1) is thus rewritten as the system: $\langle \gamma \mid y \rangle \le k_\phi$ for every $\gamma \in \partial \check{k}_\phi(0)$. The same is done for each function $\check{v}_\xi$.

As for the dual (to the profit or cost problem), it can be reformulated as an LP by applying (C.40) to the $C_{\mathrm{LR}}(\cdot, r, w)$ of (24.3) to rewrite the subdifferential condition (24.15) as the following system of linear constraints on the dual variables (viz., either r or p or both): $\langle p \mid y \rangle \le r \cdot \check{k}(y) + w \cdot \check{v}(y)$ for every $y \in Y_0$.

Spelt out, the profit-maximizing plant operation programme (full, not reduced) is thus reformulated as the LP:

(28.1) Given p, k and $w \ge 0$

(28.2) maximize $\langle p \mid y \rangle - w \cdot v$ over y and v

(28.3) subject to: $\left\langle \gamma_\phi \mid y \right\rangle \le k_\phi$ for every $\gamma_\phi \in \partial \check{k}_\phi(0)$, for each $\phi \in \Phi$

(28.4) $\langle \iota_\xi \mid y \rangle - v_\xi \le 0$ for every $\iota_\xi \in \partial \check{v}_\xi(0)$, for each $\xi \in \Xi$

(28.5) $\langle \lambda \mid y \rangle \le 0$ for every $\lambda \in Y_0^\circ$.

An equivalent *sub*-system of these constraints is obtained by taking only an extreme point of $\partial \check{k}_\phi(0)$ as a γ_ϕ, i.e., by replacing $\partial \check{k}_\phi(0)$ with $\operatorname{ext} \partial \check{k}_\phi(0)$ in (28.3). Similarly, ι_ξ can be made to run only through $\operatorname{ext} \partial \check{v}_\xi(0)$, and λ to be a generator of the cone Y_0°. But even after the pruning, the LP (28.1)–(28.5) may be doubly infinite: the number of its decision variables is finite if and only if the space Y is finite-dimensional, and the number of constraints is finite if each $\partial \check{k}_\phi(0)$ or $\partial \check{v}_\xi(0)$ is a polytope and the cone Y_0° is finitely generated.

And the plant valuation programme is reformulated as the LP:

(28.6) Given p, k and $w \ge 0$

(28.7) minimize $r \cdot k$ over r

(28.8) subject to: $r \ge 0$

(28.9) $\langle p \mid y \rangle \le r \cdot \check{k}(y) + w \cdot \check{v}(y)$ for every $y \in Y_0$.

This LP has a finite number of variables, so it is generally semi-infinite (although the constraints can of course be whittled down to a finite system if Y_0 is finitely generated *and* both $\check{k}(y)$ and $\check{v}(y)$ are linear in y—but this is *not* the case with (15.2), (15.6) or (15.7)).

29. Conclusions

The long-run general equilibrium can be determined most efficiently through the short-run equilibrium, which itself is of central practical interest. Our method uses either the producer's plant operation and valuation programmes, which form a primal-dual pair, or an optimal-value function. The choice depends on the available description of the technology but, in engineering models with multiple outputs, this is usually a production set (which favours the use of programming). The primal programme in question can be either short-run profit maximization or short-run cost minimization, but the profit approach is much easier. This brings to the fore the equilibrium pricing of capital goods and natural resources. Such inputs divide into those which are fixed, or nearly fixed, even in the long run (e.g., river flows for hydroelectric generation) and those which are variable in the long run but are supplied at an increasing marginal cost (like water reservoirs). Correct valuation of such inputs is essential for efficient investment decisions and operating policies, as well as to other matters (compensation payments for, e.g., land or rivers). Their values, as the key to the transition from the short-run to the long-run solution, are fundamental to the approach. Thus the use of long-run general-equilibrium analysis puts valuation on a sound basis, and the short-run programmes provide a workable method for calculating these values.

Appendix A. Example of duality gap between short-run profit maximization and fixed-input valuation

Equality of the primal and dual optimal values is equivalent to semicontinuity of either value function w.r.t. its "own" parameters, i.e., Type One semicontinuity (Section 20). Therefore, any sufficient condition for continuity of the one value rules out a duality gap and implies that the other value is semicontinuous. It also implies that the other programme is soluble (Section 23). In this Appendix, "continuity" means Type One continuity (unless specified as Type Two).

Any result for the primal value can be transcribed for the dual value by swapping the two programmes. Below, we consider only those sufficient conditions for continuity which are put entirely and directly in terms of the primal programme. Such a criterion can be classified by the particular value whose continuity it guarantees, i.e., it is either a primal-value or a dual-value continuity criterion. In other words, it gives, in terms of the one programme, a condition that guarantees value continuity for either the same or the other programme of the pair.

There is a salient criterion in each class. A criterion of *primal*-value continuity (w.r.t. primal parameters) is Slater's Condition on the primal programme, together with its generalized forms: see [44, (8.12) and Theorem 18 (a)]. A useful criterion of *dual*-value continuity (w.r.t. dual parameters) can be based on compactness and continuity conditions on the primal constraints and the optimand: see [44, Theorem 18' (e)]. Its semicontinuity implication for the primal value, w.r.t. primal parameters, can be viewed as a version of a part of Berge's Maximum Theorem [6, VI.3: Theorem 2]; the basic semicontinuity result of [44, Example 4' after (5.13)] is simply a special case of Berge's. Our semicontinuity results are closely related, being applications of Berge's Theorem (Lemmas 21.3–21.5).

In the context of profit or cost as the primal value function, Slater's Condition takes the form spelt out in Section 23. Furthermore, in the case of short-run profit maximization with conditionally fixed coefficients, Slater's Condition boils down to strict positivity of the fixed-input bundle k; this guarantees continuity of $\Pi_{\mathrm{SR}}(p, \cdot, w)$ on a neighbourhood of k (Part 2 of Proposition 27.2). The alternative upper semicontinuity result for $\Pi_{\mathrm{SR}}(p, \cdot, w)$ on K (Lemma 21.3) requires a price system from the predual of the commodity space, i.e., a $p \in Y'$ (in addition to Production Set Assumptions 2, 3, 8 and 9, which hold whenever Parts 1, 2 and 4 of Lemma 26.1 apply).

Either condition (positive capacities or predual output price) rules out a duality gap between profit-maximizing operation and plant valuation (for a c.f.c. technique satisfying the relevant PSAs). Between them, the two sufficient conditions cover a lot of ground: although the alternation "$p \in Y'$ or $k \gg 0$" is not actually necessary for $\overline{\Pi}_{\mathrm{SR}}$ to equal Π_{SR} at (p, k, w), it comes close to being so with technologies such as pumped storage and hydroelectric generation. In the case of storage, if the reservoir capacity k_{St} is zero *and* the price system $p \in L^{\infty *}[0,T]$ has a singular a.k.a. purely finitely additive part $p_{\mathrm{FA}} \neq 0$ (in addition to a density a.k.a. countably additive part $p_{\mathrm{CA}} \in L^1$), then the operating profit is obviously zero, but the unit value of conversion capacity is positive. This example, spelt out next, shows also that the failure of Slater's Condition can lead to nonexistence of an exact dual solution. A similar example of a duality gap for the hydro technology is given in [24].[135]

Example A.1 (Duality gap between operation and valuation of an incomplete plant)**.** *Take the pumped-storage technology (15.4) and an output price system $p \in L^{\infty *}[0,T]$ with $p_{\mathrm{FA}} \neq 0$ and $p_{\mathrm{CA}} \in \mathrm{BV} \subset L^1$ (i.e., with a nonzero singular part and a density part of bounded variation). If additionally $k_{\mathrm{Co}} > 0$ but $k_{\mathrm{St}} = 0$ (i.e., the plant has a conversion capacity but no storage capacity), then the operating profit is zero, i.e., $\Pi_{\mathrm{SR}}^{\mathrm{PS}}(p; 0, k_{\mathrm{Co}}) = 0$. But the optimal stock price (the dual solution) is $\hat{\psi} = p_{\mathrm{CA}}$, and so the capacity value (the dual optimal value) is*

$$\overline{\Pi}_{\mathrm{SR}}^{\mathrm{PS}}(p; 0, k_{\mathrm{Co}}) = k_{\mathrm{Co}} \|p_{\mathrm{FA}}\|_{\infty}^{*} > 0 = \Pi_{\mathrm{SR}}^{\mathrm{PS}}(p; 0, k_{\mathrm{Co}}) . \tag{A.1}$$

If $p_{\mathrm{CA}} \in L^1 \setminus \mathrm{BV}$ (and still $k_{\mathrm{Co}} > 0$ but $k_{\mathrm{St}} = 0$), then the dual (stock-pricing) programme for ψ has no (exact) solution, but any sequence of ψ's in BV that converges to p_{CA} in the L^1-norm is a sequence of approximate dual optima. The infimal capacity value is still $k_{\mathrm{Co}} \|p_{\mathrm{FA}}\|_{\infty}^{*}$ (i.e., there is the same duality gap).

Comments (on Example A.1):

(1) It gives an example of a duality gap in *infinite* linear programming, since the SRP programme can be formulated as an LP: see (16.12)–(16.16).

(2) The example shows in a simple way why a duality gap must open at a point of the optimal value's discontinuity (of Type One). With the other parameters ($p \in L^{\infty *}$ and $k_{\mathrm{Co}} > 0$) kept fixed, $\overline{\Pi}_{\mathrm{SR}}$ and Π_{SR} are equal and vary continuously with k_{St} as long as it stays positive: every finite concave function on $\mathbb{R}_{++}$ is continuous, and $\overline{\Pi}_{\mathrm{SR}} = \Pi_{\mathrm{SR}}$ when $k_{\mathrm{St}} > 0$ because this is Slater's Condition. But at $k_{\mathrm{St}} = 0$, Π_{SR} can fail to be right-continuous and then, being concave, it also fails to be u.s.c.—which means that it drops at $k_{\mathrm{St}} = 0$.[136] By contrast, Type Two semicontinuity holds automatically, i.e., $\overline{\Pi}_{\mathrm{SR}}$ is always u.s.c. and hence it is actually right-continuous at $k_{\mathrm{St}} = 0$. So the discontinuity of Π_{SR} at $k_{\mathrm{St}} = 0$ implies that $\Pi_{\mathrm{SR}}(0) < \overline{\Pi}_{\mathrm{SR}}(0)$. See Figure 5.

[135] In the case of hydro with $p \geq 0$, $p_{\mathrm{FA}} \neq 0$ and $k_{\mathrm{St}} = 0$, if $k_{\mathrm{Tu}} > \mathrm{Sup}(e)$ then $\langle p \,|\, e \rangle < \langle p_{\mathrm{CA}} \,|\, e \rangle + k_{\mathrm{Tu}} \|p_{\mathrm{FA}}\|$; i.e., the optimal output is obviously equal to the inflow e, which yields a revenue lower than the value of hydro inputs (turbine and inflow).

[136] A finite concave function on a polyhedral set $Z \subseteq \mathbb{R}^n$ is l.s.c. on Z (so if it is u.s.c. on Z then it is continuous on Z): see [42, 10.2 and 20.5]. This applies to $Z = \mathbb{R}^n_+$ for every n (here, $n = 1$).

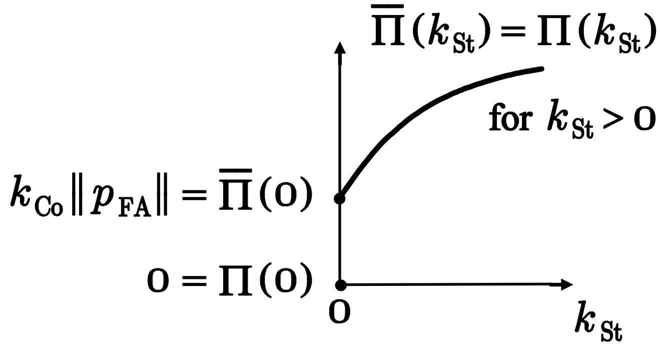

FIGURE 5. Capacity value and operating profit for the pumped-storage technique, $\overline{\Pi}_{\mathrm{SR}}$ and Π_{SR}, as functions of storage capacity k_{St} (for a fixed conversion capacity $k_{\mathrm{Co}} > 0$ and a fixed good's price $p \in L^{\infty *} \setminus L^1$). When $k_{\mathrm{St}} > 0$, Slater's Condition is met and so $\overline{\Pi} = \Pi$, but a duality gap opens at $k_{\mathrm{St}} = 0$, where $\overline{\Pi}$ is continuous but Π drops (Example A.1).

(3) Recall from Section 6 that the data (here, p and k) and a pair of solutions (here, y and r) with the *same* value (i.e., without a duality gap) can be permuted to form the data and solutions to another programme pair. As the example shows, this need not be so when there *is* a duality gap. Indeed, none of the other programme pairs need have a gap. In this example, the SRP programme pair does have a gap, but the LRC and the SRC programme pairs do not, since both cost functions are semicontinuous in the quantities (which means Type One semicontinuity). That is, C_{LR} is L^1-l.s.c. (and *a fortiori* $L^{\infty *}$-l.s.c.) in $y \in L^\infty$. (This can be shown either directly from the formulae for capacity requirements (15.6)–(15.7), or by applying Lemma 21.4.) The same is obviously true of C_{SR} as a function of (y, k), which is simply the 0-∞ indicator function of the closed set $\mathbb{Y}$. (There are no variable inputs with this technique, i.e., the SRC programme is merely a check of capacity sufficiency.) So permutation of p, k, y and r must fail to yield a cost-minimizing solution and its dual, and it does fail: (i) the LRC programme's solution has $k_{\mathrm{Co}} = 0$, unlike the SRP data in this example; and (ii) the OFIV (dual to SRC) programme's solution has $r_{\mathrm{Co}} = 0$, unlike the FIV (dual to SRP) programme's solution, which has $r_{\mathrm{Co}} = \|p_{\mathrm{FA}}\|_\infty^* > 0$. (In detail, the SRP primal-dual solution pair—given a nonconstant $p_{\mathrm{CA}} \in \mathrm{BV}$ and $p_{\mathrm{FA}} \neq 0$, $k_{\mathrm{St}} = 0$ and $k_{\mathrm{Co}} > 0$—is $y = 0$ and $r = (r_{\mathrm{St}}, r_{\mathrm{Co}}) = \left(\mathrm{Var}_{\mathrm{c}}^{+}(p_{\mathrm{CA}}), \|p_{\mathrm{FA}}\|_\infty^*\right) \gg 0$. But, given $y = 0$ and $r = (r_{\mathrm{St}}, r_{\mathrm{Co}}) \gg 0$, the LRC solution pair is obviously $(k_{\mathrm{St}}, k_{\mathrm{Co}}) = (0, 0)$ with any LRMC as p, i.e., with any $p \in r_{\mathrm{St}} \partial \check{k}_{\mathrm{St}}(0) + r_{\mathrm{Co}} \partial \check{k}_{\mathrm{Co}}(0) + \mathrm{const.}$ Similarly, given $y = 0$, $k_{\mathrm{St}} = 0$ and $k_{\mathrm{Co}} > 0$, the SRC dual solution is $r_{\mathrm{Co}} = 0$ with any $r_{\mathrm{St}} \geq 0$ and any $p \in r_{\mathrm{St}} \partial \check{k}_{\mathrm{St}}(0) + \mathrm{const.}$)

APPENDIX B. A NONFACTORABLE JOINT SUBDIFFERENTIAL

We identify a class of jointly convex functions of two variables (which can be vector variables) such that: (i) nondifferentiability in one of the variables implies nondifferentiability in the other, and (ii) the joint subdifferentials do *not* factorize into the Cartesian product of the partial subdifferentials. This means that a partial subgradient cannot be extended to a joint one by adjoining just any partial subgradient w.r.t. the other variable. But, as we also show, it can usually be extended by a suitable choice of the other partial subgradient.

Proposition B.1. *Assume that $C\colon Y\times K\to\mathbb{R}\cup\{+\infty\}$ is (jointly) positively linearly homogeneous, convex and lower semicontinuous (for the pairing of the space $Y\times K$ with $P\times R$). If additionally $(p',-r')$ and $(p'',-r'')$ are elements of $\partial_{y,k}C(y,k)$ with*[137]

$$\langle p'\,|\,y\rangle\neq\langle p''\,|\,y\rangle \tag{B.1}$$

then $r'\neq r''$ (so $\partial_k C(y,k)$ is not a singleton, i.e., $C(y,\cdot)$ is not Gateaux-differentiable at k). What is more, neither $(p',-r'')$ nor $(p'',-r')$ is in $\partial_{y,k}C(y,k)$, and so

$$\partial_{y,k}C(y,k)\neq\partial_y C(y,k)\times\partial_k C(y,k)\,.$$

Proof. By (C.41), which is a variant of Euler's Theorem,

$$C(y,k)=\langle p\,|\,y\rangle-\langle r\,|\,k\rangle \tag{B.2}$$

for every $(p,-r)\in\partial_{y,k}C(y,k)$. So (B.2) holds for both $(p',-r')$ and $(p'',-r'')$, but it therefore fails for $(p',-r'')$ and $(p'',-r')$ because of (B.1). So neither $(p',-r'')$ nor $(p'',-r')$ is in $\partial_{y,k}C(y,k)$, which shows that this set is not a Cartesian product. □

Example B.2. *Take the function $c\colon\mathbb{R}^2_+\to\mathbb{R}$ defined as in (2.7), i.e., $c(\mathsf{y},k)=w\mathsf{y}$ if $0\le\mathsf{y}\le k$ and $+\infty$ otherwise (given a number $w\ge 0$). With the scalar product $\langle\mathsf{p},-r\,|\,\mathsf{y},k\rangle:=\mathsf{p}\mathsf{y}-rk/T$ where $T>0$ is a given number, the joint subdifferential at a point with $\mathsf{y}=k>0$ is*

$$\partial_{\mathsf{y},k}c(\mathsf{y},k)=\left\{(\mathsf{p},-r)\in\mathbb{R}_+\times\mathbb{R}_-:p=w+\frac{r}{T},\ r\ge 0\right\}$$

(which, being a half-line not parallel to either axis of the plane $\mathbb{R}^2$, is not a Cartesian product).

When c serves as a convex integrand, this non-factorization is inherited by the integral functional

$$C(y,k):=\int_0^T c(y(t),k)\,\mathrm{d}t\quad\text{for } y\in L^\infty[0,T]\,.$$

Take a y and k with $0\ll y\le k$ and $\operatorname{meas}\{t\in[0,T]:y(t)=k\}>0$. When $L^1[0,T]\times\mathbb{R}$ is paired with $L^\infty[0,T]\times\mathbb{R}$ by the scalar product $\langle p,-r\,|\,y,k\rangle:=\int_0^T p(t)\,y(t)\,\mathrm{d}t-rk$, one has $(p,-r)\in\partial_{y,k}C(y,k)$ if and only if both $p=w+\kappa$ and $r=\int_0^T\kappa(t)\,\mathrm{d}t$ for some $\kappa\in L^1_+[0,T]$ with $\kappa(t)=0$ for a.e. $t\in[0,T]$ such that $y(t)<k$.

Besides this example, Condition (B.1) is met by some $(p',-r')$ and $(p'',-r'')$ from $\partial_{y,k}C(y,k)$ if: (i) Y is a vector lattice, P is a sublattice of the order dual $Y^\sim$, and y is strictly positive as a linear functional on $Y^\sim$, (ii) $\partial_y C(y,k)$ contains a p' and a p'' with $p'<p''$,[138] and (iii) Every partial subgradient $p\in\partial_y C(y,k)$ can be extended to a joint subgradient $(p,-r)\in\partial_{y,k}C(y,k)$.

Such extensibility can be proved in two ways. One method is to establish that the relevant partial conjugate of the bivariate convex function C is superdifferentiable in

[137] The minus sign in $(p,-r)$ is there to make r nonnegative when $C(y,\cdot)$ is nonincreasing on K.

[138] Then $\langle p'\,|\,y\rangle<\langle p''\,|\,y\rangle$, since $p'<p''$ and $y\gg 0$.

the non-conjugated variable—i.e., to introduce the saddle (convex-concave) function on $P \times K$ defined by $\Pi := C^{\#_1}$, then show that $\widehat{\partial}_k \Pi (p, k) \neq \emptyset$ for the given k and the given $p \in \partial_y C (y, k)$, and finally apply the Subdifferential Sections Lemma (i.e., Lemma C.5) to conclude that any $r \in \widehat{\partial}_k \Pi (p, k)$ extends p to a $(p, -r) \in \partial_{y,k} C (y, k)$. This can also be an effective method of calculating a suitable r. Without introducing Π, mere existence of such an r can also be proved by using the Hahn-Banach Extension Theorem, which can be stated as follows in terms of subgradients.

Theorem B.3 (Hahn-Banach)**.** *Assume that $C \colon Y \times K \to \mathbb{R} \cup \{+\infty\}$ is a (jointly) convex function, where Y and K are topological vector spaces (with P and R as the continuous duals). If $k \in \operatorname{int}_K \operatorname{dom} (C (y, \cdot))$, i.e., $C \left(y, \widetilde{k} \right) < +\infty$ for every $\widetilde{k}$ in some neighbourhood of k, then for every $p \in \partial_y C (y, k)$ there exists an r such that $(p, -r) \in \partial_{y,k} C (y, k)$.*

Proof. See, e.g., [37, Theorem 0.28]; although that formulation applies only when $(y, k) \in \operatorname{int}_{Y \times K} \operatorname{dom} C$, the same proof is valid under the weaker assumption made here. □

Theorem B.3 does not apply to the boundary points of the function's effective domain, which is

$$\operatorname{dom} C := \{(y, k) : C (y, k) < +\infty\} .$$

And indeed, at a boundary point, a partial subgradient may have no extension (to a joint one). But it is useful to identify those cases in which such extensions do exist. This is because the boundary points can be the points of greatest interest: e.g., when C is the SRC as a function of the output bundle y and the fixed-input bundle k, all the efficient combinations of y and k lie on the boundary of $\operatorname{dom} C$. However, if C has a *finite* convex extension C^{Ex}, defined on the whole space (or at least on a neighbourhood of $\operatorname{dom} C$), and $\operatorname{dom} C$ is the sublevel set of another finite convex function C^{Do}, then Theorem B.3 can be applied to both functions, C^{Ex} and C^{Do}. For the original function C, this yields a result that applies also to the domain's boundary points.

Corollary B.4. *Let $C \colon Y \times K \to \mathbb{R} \cup \{+\infty\}$ be a (jointly) convex function. Assume that:*

(1) *Its effective domain has the form*

$$\operatorname{dom} C = \left\{ (y, k) : C^{\mathrm{Do}} (y, k) \leq 0 \text{ and } k \in K_0 \right\} \tag{B.3}$$

where K_0 is a convex subset of K, and $C^{\mathrm{Do}} \colon Y \times K \to \mathbb{R}$ is a continuous convex function.

(2) *$k \in K_0$ and $C^{\mathrm{Do}} (y, k) \leq 0$, i.e., $(y, k) \in \operatorname{dom} C$.*

(3) *There exists a $y^{\mathrm{S}} \in Y$ with $C^{\mathrm{Do}} \left(y^{\mathrm{S}}, k \right) < 0$ (Slater's Condition).*

(4) *C (or, more precisely, its restriction to $\operatorname{dom} C$) has a continuous convex extension $C^{\mathrm{Ex}} \colon Y \times K \to \mathbb{R}$.*

Then for every $p \in \partial_y C (y, k)$ there exists an r such that $(p, -r) \in \partial_{y,k} C (y, k)$.

Proof. Every $p \in \partial_y C (y, k)$ has the form $p = p' + \alpha p''$ for some $p' \in \partial_y C^{\mathrm{Ex}} (y, k)$, $p'' \in \partial_y C^{\mathrm{Do}} (y, k)$ and a scalar $\alpha \geq 0$, with $\alpha = 0$ if $C^{\mathrm{Do}} (y, k) < 0$. This is because, since $C = C^{\mathrm{Ex}} + \delta (\cdot \mid \operatorname{dom} C)$,

$$\partial_y C (y, k) = \partial_y C^{\mathrm{Ex}} (y, k) + \partial_y \delta (y, k \mid \operatorname{dom} C)$$

$$= \partial_y C^{\mathrm{Ex}}(y,k) + \partial\delta\left(y \mid \left\{y' : C^{\mathrm{Do}}(y',k) \le 0\right\}\right)$$

$$= \partial_y C^{\mathrm{Ex}}(y,k) + \operatorname{cone}\left(\partial_y C^{\mathrm{Do}}(y,k)\right) \tag{B.4}$$

if $C^{\mathrm{Do}}(y,k) = 0$. When $C^{\mathrm{Do}}(y,k) < 0$, the term $\partial_y \delta$ (which is the outward normal cone to the sublevel set of $C^{\mathrm{Do}}(\cdot,k)$) is $\{0\}$, in which case the term denoting the cone generated by $\partial_y C^{\mathrm{Do}}$ must be deleted from (B.4). For additivity of ∂ (also with an application to a sum of the form $C+\delta$), see, e.g., [42, 23.8 and proof of 28.3.1], [44, Theorem 20] or [48, 5.38 and 7.2]. The relevant formula for the normal cone to a sublevel set is given in, e.g., [32, 4.3: Proposition 2], [42, 23.7.1] or [48, 7.8].

Since C^{Ex} and C^{Do} are continuous (everywhere on $Y \times K$), Theorem B.3 applies to both; so there exist r' and r'' with

$$(p', -r') \in \partial_{y,k} C^{\mathrm{Ex}}(y,k) \text{ and } (p'', -r'') \in \partial_{y,k} C^{\mathrm{Do}}(y,k). \tag{B.5}$$

It now suffices to set $r := r' + \alpha r''$. To see this, use again the formula for the normal cone and the additivity of ∂ (this time for joint subdifferentials) to obtain from (B.5) that

$$\begin{aligned}
(p,-r) &= (p' + \alpha p'', -r' - \alpha r'') \\
&\in \partial_{y,k} C^{\mathrm{Ex}}(y,k) + \partial_{y,k}\delta\left(y,k \mid \left\{y',k' : C^{\mathrm{Do}}(y',k') \le 0\right\}\right) \\
&\subseteq \partial_{y,k} C^{\mathrm{Ex}}(y,k) + \partial_{y,k}\delta\left(y,k \mid \left\{C^{\mathrm{Do}} \le 0\right\}\right) + \partial_{y,k}\delta\left(y,k \mid Y \times K_0\right) \\
&= \partial_{y,k} C^{\mathrm{Ex}}(y,k) + \partial_{y,k}\delta\left(y,k \mid \operatorname{dom} C\right) = \partial_{y,k} C(y,k).
\end{aligned}$$

The penultimate equality follows from (B.3); also, $\partial_{y,k}\delta(y,k \mid Y \times K_0) = \{0\} \times \partial_k \delta(k \mid K_0)$ on its l.h.s. $\square$

Comments: Extensibility of partial subgradients means that the obvious inclusions $\partial_{y,k} C_{\mathrm{SR}} \subseteq \partial_y C_{\mathrm{SR}} \times \partial_k C_{\mathrm{SR}}$ and $\partial_y C_{\mathrm{LR}} \subseteq \partial_y C_{\mathrm{SR}}$—or (9.1) and (11.7)—are "tight", each in its sense:

(1) $\partial_y C_{\mathrm{SR}}$ is equal to the projection of $\partial_{y,k} C_{\mathrm{SR}}$ onto Y if and only if every $p \in \partial_y C_{\mathrm{SR}}$ extends to some $(p,-r) \in \partial_{y,k} C_{\mathrm{SR}}$. A similar result applies to $\partial_k C_{\mathrm{SR}}$.

(2) With C_{LR} defined by (11.2), if every $p \in \partial_y C_{\mathrm{SR}}(y,k)$ extends to some $(p,-r) \in \partial_{y,k} C_{\mathrm{SR}}(y,k)$ then

$$\partial_y C_{\mathrm{SR}}(y,k) = \bigcup_{r \in -\partial_k C_{\mathrm{SR}}(y,k)} \partial_y C_{\mathrm{LR}}(y,r).$$

This follows from the second equivalence in (11.1), which is a case of the SSL (Lemma C.5). A similar result for C_{SR} and Π_{SR} shows that the inclusion (11.4) is tight.

Appendix C. Convex conjugacy and subdifferential calculus

C.1. **Semicontinuous envelope.** Let $C\colon Y \to \mathbb{R} \cup \{\pm\infty\}$ be a convex extended-real function on a real vector space Y that is paired with another one, P, by a bilinear form $\langle \cdot \mid \cdot \rangle : P \times Y \to \mathbb{R}$. The *effective domain* of C is the convex set

$$\operatorname{dom} C := \{y \in Y : C(y) < +\infty\}.$$

Given a locally convex topology $\mathcal{T}$ on Y that is *consistent* with P (i.e., makes P the continuous dual space), the *l.s.c. envelope* of C is the greatest lower semicontinuous (l.s.c.) minorant of C. Denoted by $\operatorname{lsc} C$, it can be determined pointwise by the formula

$$(\operatorname{lsc} C)(y) := \min\left\{C(y), \liminf_{y' \to y} C(y')\right\}$$

or globally by the formula $\operatorname{epi} \operatorname{lsc} C := \operatorname{cl} \operatorname{epi} C$, where cl means the $\mathcal{T}$-closure, and

$$\operatorname{epi} C := \{(y, \varrho) \in Y \times \mathbb{R} : C(y) \leq \varrho\}$$

is the epigraph of C. Note that $\operatorname{lsc} C$ depends on the dual space P but not on the consistent topology $\mathcal{T}$, by the Hahn-Banach Separation Theorem [18, 12A: Corollary 1]. Also, C is l.s.c. at y if and only if $C(y) = (\operatorname{lsc} C)(y)$.

A *proper convex* function is one that takes a finite value (somewhere) but does not take the value $-\infty$ (anywhere). A convex function taking the value $-\infty$ is peculiar: it may take finite values only on the algebraic boundary of its effective domain,[139] and it has no finite value at all if it is lower semicontinuous along each straight line: see, e.g., [42, 7.2 and 7.2.1], [44, Theorem 4] or [48, 5.12 with Proof].

C.2. **The conjugate function.** The Fenchel-Legendre *convex conjugate* of C is

$$C^{\#}(p) := \sup_{y \in Y} (\langle p \mid y \rangle - C(y)) \tag{C.1}$$

for $p \in P$; it is l.s.c. and either proper convex or an infinite constant ($+\infty$ or $-\infty$). Obviously

$$C^{\#}(p) \geq \langle p \mid y \rangle - C(y) \tag{C.2}$$

for every y and p; this is the Fenchel-Young Inequality.

The second convex conjugate, $C^{\#\#}$, is the pointwise supremum of all the affine minorants of C with coefficients in P (supremum of those functions of the form $\langle p \mid \cdot \rangle - \varrho$, with $p \in P$ and $\varrho \in \mathbb{R}$, that nowhere exceed C), i.e.,

$$C^{\#\#}(y) = \sup_{p \in P,\, \varrho \in \mathbb{R}} \{\langle p \mid y \rangle - \varrho : \langle p \mid y' \rangle - \varrho \leq C(y') \text{ for every } y' \in Y\}. \tag{C.3}$$

So $C^{\#\#}$ is l.s.c. on Y and

$$C^{\#\#} \leq \operatorname{lsc} C \leq C. \tag{C.4}$$

Furthermore, $C^{\#\#} = \operatorname{lsc} C$ unless $\operatorname{lsc} C$ takes the value $-\infty$ (and hence has no finite value).[140] In the latter case, $C^{\#\#} = -\infty$ (everywhere on Y) and $\operatorname{lsc} C = -\infty$ on the convex set $\operatorname{cl} \operatorname{dom} C$, but $\operatorname{lsc} C = +\infty$ on the complement set. So if C is l.s.c. at y then: (i) $C^{\#\#}(y)$ and $C(y)$ can differ only by being oppositely infinite, and (ii) $C^{\#\#}(y) = C(y)$ if and only if either $C(y) < +\infty$ or both $C(y) = +\infty$ and $\operatorname{lsc} C > -\infty$ everywhere on

[139] In precise terms, $C(y) = -\infty$ for every y in the intrinsic core (a.k.a. the relative algebraic interior) of $\operatorname{dom} C$ if $C(y') = -\infty$ for some y' (and C is convex).

[140] When additionally Y is finite-dimensional, if $\operatorname{lsc} C$ takes the value $-\infty$, then so does C itself. This follows from [42, 7.5]; it is stated in, e.g., [44, Example 1”].

Y. Also, $C^{\#\#} = C$ (everywhere on Y) if and only if C is either l.s.c. proper convex or an infinite constant.[141] Applied to $C^{\#}$ (instead of C), this shows that

$$C^{\#\#\#} = C^{\#} \tag{C.5}$$

(which can also be seen directly from (C.1) and (C.4): $\left(C^{\#\#}\right)^{\#} \geq C^{\#}$ because $C^{\#\#} \leq C^{\#}$, but also $\left(C^{\#}\right)^{\#\#} \leq C^{\#}$).

For a bivariate convex function C, its *partial* second conjugate (i.e., its second conjugate taken w.r.t. just one variable y, with the other variable k kept fixed) lies always between the *total* second conjugate (i.e., the second conjugate w.r.t. both variables) and the original function itself. Formally, the partial first and second conjugates w.r.t., say, the first variable of a bivariate convex function C on $Y \times K$ (where K is another vector space) is defined by

$$C^{\#_1}(p,k) := (C(\cdot,k))^{\#}(p) := \sup_{y \in Y} \left(\langle p \mid y \rangle - C(y,k)\right) \tag{C.6}$$

for every $p \in P$ and $k \in K$. This $(C^{\#_1})$ is a saddle (convex-concave) function on $P \times K$: it is convex (like C) in the "conjugated" first variable, but (unlike C) it is concave in the non-conjugated second variable. The partial second conjugate (w.r.t. the first variable) is the bivariate convex function

$$C^{\#_1\#_1}(y,k) := (C(\cdot,k))^{\#\#}(y). \tag{C.7}$$

Remark C.1 (Inequality between partial and total second conjugates)**.** *Assume that $C\colon Y \times K \to \mathbb{R} \cup \{\pm\infty\}$, where Y and K are vector spaces paired with P and R. Then*

$$C^{\#\#} \leq C^{\#_1\#_1} \leq C \tag{C.8}$$

on $Y \times K$. (In other words, for each $k \in K$, if C_k means the function on Y defined by $C_k(y) := C(y,k)$ for every y, then $\left(C^{\#\#}\right)_k \leq (C_k)^{\#\#} \leq C_k$ on Y.)

Proof. The second inequality of (C.8) is a case of (C.4), without the middle term. As for the first inequality of (C.8), this follows from a comparison, for the partial and total second conjugates, of their representations as suprema of affine minorants: by (C.3) applied to $C(\cdot,k)$ and to C,

$$C^{\#_1\#_1}(y,k) = \sup_{p \in P,\, \alpha \in \mathbb{R}} \left\{\langle p \mid y \rangle - \alpha : \langle p \mid \cdot \rangle - \alpha \leq C(\cdot,k)\right\} \tag{C.9}$$

$$C^{\#\#}(y,k) = \sup_{p \in P,\, r \in K,\, \beta \in \mathbb{R}} \left\{\langle p, -r \mid y, k \rangle - \beta : \langle p, -r \mid \cdot, \cdot \rangle - \beta \leq C(y')\right\}. \tag{C.10}$$

By setting α equal to $\langle r \mid k \rangle + \beta$, it follows that the supremum in (C.9) is not less than that in (C.10).[142] $\square$

[141] In [42] and [44], C is called "closed" when $C = C^{\#\#}$, and $\operatorname{cl} C$ serves as an alternative notation for $C^{\#\#}$. This is abandoned in [45], and rightly so: $\operatorname{cl} C$ can be misinterpreted as $\operatorname{lsc} C$, especially since others—e.g., [37]—do use $\operatorname{cl} C$ instead of $\operatorname{lsc} C$ (to have $\operatorname{epi} \operatorname{cl} C := \operatorname{cl} \operatorname{epi} C$).

[142] In other words, the α in (C.9) is allowed to vary with k in any way (subject to the stated inequality), whilst the corresponding term in (C.10) is $\langle r \mid k \rangle + \beta$, which is additionally linear in k.

C.3. **Subgradients.** A $\mathcal{T}$-continuous *subgradient* (a.k.a. topological subgradient) of C at a $y \in Y$ is any $p \in P$ such that

$$C(y + \Delta y) \geq C(y) + \langle p \mid \Delta y \rangle \tag{C.11}$$

for every $\Delta y \in Y$. The set of all subgradients (at y) is the *subdifferential* $\partial C(y)$. In other words,

$$p \in \partial C(y) \Leftrightarrow y \text{ maximizes } \langle p \mid \cdot \rangle - C \tag{C.12}$$

$$\Leftrightarrow C^{\#}(p) = \langle p \mid y \rangle - C(y). \tag{C.13}$$

So the graph of the subdifferential correspondence ($\partial C \subseteq Y \times P$) consists of those points (y, p) at which the Fenchel-Young Inequality holds as an equality.

Any linear, not necessarily $\mathcal{T}$-continuous, functional p meeting (C.11) is an *algebraic* subgradient of C at y, and the set of *all* such subgradients is the *algebraic subdifferential* $\partial^{\mathrm{a}} C(y)$, with $P \cap \partial^{\mathrm{a}} C(y) = \partial C(y)$ by definition. The two subdifferentials are identical, for every C, when $\mathcal{T}$ is the *strongest locally convex topology*, $\mathcal{T}_{\mathrm{SLC}}$, on Y. This is because every linear functional on Y is $\mathcal{T}_{\mathrm{SLC}}$-continuous, i.e., the $\mathcal{T}_{\mathrm{SLC}}$-continuous dual is equal to the *algebraic dual* Y^{a} (what is more, $\mathcal{T}_{\mathrm{SLC}}$ is obviously $\mathrm{m}(Y, Y^{\mathrm{a}})$, the Mackey topology for this pairing).

Directly from the subgradient inequality (C.11), if C' and C'' are convex functions with values in $\mathbb{R} \cup \{+\infty\}$, i.e., not taking the value $-\infty$, then

$$\partial (C' + C'')(y) \supseteq \partial C'(y) + \partial C''(y). \tag{C.14}$$

Equality holds for proper convex functions under a continuity assumption: if, in addition to C' and C'' being convex with values in $\mathbb{R} \cup \{+\infty\}$, there exists a point of Y at which both C' and C'' are finite and at least one (C' or C'') is continuous, then

$$\partial (C' + C'')(y) = \partial C'(y) + \partial C''(y) \tag{C.15}$$

for *every* $y \in Y$. See, e.g., [44, Theorem 20 (i) under (a)] or [48, 5.38 (b)]. Applied to the case of 0-∞ indicator functions of convex subsets of Y, (C.15) gives the outward normal cone to the intersection of sets Z' and Z'' as the sum of their normal cones, i.e.,

$$\begin{aligned} \mathrm{N}(y \mid Z' \cap Z'') &:= \partial\delta(y \mid Z' \cap Z'') = \partial\delta(y \mid Z') + \partial\delta(y \mid Z'') \\ &=: \mathrm{N}(y \mid Z') + \mathrm{N}(y \mid Z'') \end{aligned} \tag{C.16}$$

for *every* $y \in Y$ if $Z' \cap \operatorname{int} Z'' \neq \emptyset$. This is stated in, e.g., [32, 4.3: Proposition 1].

Also directly from (C.11), for every $\alpha > 0$,

$$\partial(\alpha C)(y) = \alpha \partial C(y) \tag{C.17}$$

and this holds for $\alpha = 0$ as well if (and only if) $\partial C(y) \neq \emptyset$, i.e., if C is *subdifferentiable* at y.

For C to be subdifferentiable at y, it is necessary that C be l.s.c. at y and actually that $C^{\#\#}(y) = C(y)$; in this case $\partial C^{\#\#}(y) = \partial C(y)$. In other words,

$$p \in \partial C(y) \Leftrightarrow \left(p \in \partial C^{\#\#}(y) \text{ and } C^{\#\#}(y) = C(y)\right) \tag{C.18}$$

from (C.13) and (C.4).

Lower semicontinuity is not generally sufficient for subdifferentiability, but continuity is. In precise terms, if a proper convex function $C\colon Y \to \mathbb{R} \cup \{+\infty\}$ is continuous and finite at some point of Y, then it is subdifferentiable (and continuous) at every interior point of its effective domain, i.e., $\partial C(y)$ is nonempty and, also, $\mathrm{w}(P,Y)$-compact (weakly compact) for every $y \in \operatorname{int}\operatorname{dom} C$: see, e.g., [32, 4.2: Proposition 3], [44, Theorem 11 (a)] or [48, 5.35 (a)]. Furthermore, every algebraic subgradient is then $\mathcal{T}$-continuous, i.e., $\partial^{\mathrm{a}} C(y) = \partial C(y) \neq \emptyset$ or, equivalently,

$$\emptyset \neq \partial^{\mathrm{a}} C(y) \subseteq P \tag{C.19}$$

for every $y \in \operatorname{int}\operatorname{dom} C$: see, e.g., [18, 14B: Proof of Theorem] or [37, Corollary 2 to Theorem 0.27, and p. 60].

C.4. **Continuity of convex functions.** Any continuous function is bounded from above (by a finite number) on a neighbourhood of any point where its value is either finite or $-\infty$. With *convex* functions, this obvious necessary condition is also sufficient for continuity. In precise terms, if $C: Y \to \mathbb{R} \cup \{\pm\infty\}$ is convex then the following conditions are equivalent to one another:

(1) C is continuous at some $y \in Y$ with $C(y) < +\infty$.
(2) There exists an open set $N \subseteq Y$ and a $\varrho \in \mathbb{R}$ such that $C(y) \leq \varrho$ (or, equivalently, the epigraph of C has a nonempty interior in $Y \times \mathbb{R}$).
(3) C is continuous on $\operatorname{int}\operatorname{dom} C$, which is nonempty.

See, e.g., [18, 14A], [32, 3.2: Theorem 1], [44, Theorem 8] or [48, 5.20]. In particular, this shows that continuity (of a convex function) is a property that "propagates" from any single point to the whole interior of the effective domain (Part 1 $\Rightarrow$ Part 3). Also, the sufficiency of local boundedness for continuity can be combined with a Baire category argument to deduce continuity from mere lower semicontinuity for a convex function on a Banach space (or, more generally, on a barrelled space). The result has two variants (which are very similar, but not identical): see, e.g., [44, Corollary 8B] and [18, p. 84 and Exercise 3.50].

Another "automatic continuity" result, limited to finite-dimensional spaces, is that a finite convex function C on a polyhedral set $Z \subseteq \mathbb{R}^n$ is upper semicontinuous on Z (so if C is also l.s.c. on Z then it is actually continuous on Z). More generally, a convex function $C: \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ is u.s.c. on any locally simplicial (not necessarily convex or closed) subset, Z, of $\operatorname{dom} C$. See [42, 10.2 and 20.5].

C.5. **Concave functions and supergradients.** All of these concepts and results can be reoriented to concave functions. In particular, when $\Pi: K \to \mathbb{R} \cup \{\pm\infty\}$ is a concave function on a space K paired with another space R, its *effective domain* (in the concave sense) is the convex set

$$\widehat{\operatorname{dom}}\Pi := \{k \in K : \Pi(k) > -\infty\}$$

and the *concave conjugate* of Π is

$$\Pi_{\#}(r) := \inf_{k \in K} (\langle r \mid k \rangle - \Pi(k)) \tag{C.20}$$

for $r \in R$. The second concave conjugate meets the inequality

$$\Pi_{\#\#}(k) \geq \operatorname{usc}\Pi(k) \geq \Pi(k) \tag{C.21}$$

where $\operatorname{usc}\Pi$ is the least upper semicontinuous (u.s.c.) majorant of Π; and $\operatorname{usc}\Pi(k)$ and $\Pi_{\#\#}(k)$ differ in only one case: if $k \notin \operatorname{cl}\widehat{\operatorname{dom}}\Pi$ and $\operatorname{usc}\Pi(k'') = +\infty$ for some k'', then $\operatorname{usc}\Pi(k) = -\infty$ but $\Pi_{\#\#} = +\infty$ (on K). So if Π is u.s.c. *proper* concave (i.e., takes a finite value but does not take the value $+\infty$), then $\Pi_{\#\#} = \Pi$ (everywhere). Hence

$$\Pi_{\#\#\#} = \Pi_{\#}. \tag{C.22}$$

A *supergradient* of Π at a $k \in K$ is any $r \in R$ such that

$$\Pi(k + \Delta k) \leq \Pi(k) + \langle r \mid \Delta k \rangle \tag{C.23}$$

for every $\Delta k \in K$. The set of all supergradients (at k) is the *superdifferential*, $\widehat{\partial}\Pi(k)$, i.e.,

$$r \in \widehat{\partial}\Pi(k) \Leftrightarrow k \text{ maximizes } \Pi - \langle r \mid \cdot \rangle \tag{C.24}$$

$$\Leftrightarrow \Pi_{\#}(r) = \langle r \mid k \rangle - \Pi(k). \tag{C.25}$$

Also,

$$r \in \widehat{\partial}\Pi(k) \Leftrightarrow \left(r \in \widehat{\partial}\Pi_{\#\#}(k) \text{ and } \Pi_{\#\#}(k) = \Pi(k)\right). \tag{C.26}$$

The concave and convex cases are linked by the rules:

$$\Pi_{\#}(r) = -(-\Pi)^{\#}(-r) \tag{C.27}$$

$$\Pi_{\#\#} = -(-\Pi)^{\#\#} \tag{C.28}$$

$$\operatorname{usc}\Pi = -\operatorname{lsc}(-\Pi) \tag{C.29}$$

$$\widehat{\partial}\Pi = -\partial(-\Pi). \tag{C.30}$$

C.6. **Subgradients of conjugates.** The subdifferential correspondences of mutual conjugates are inverse to each other.[143]

Theorem C.2 (Inversion Rule)**.** *Assume that $C\colon Y \to \mathbb{R} \cup \{\pm\infty\}$ is convex, and Y is paired with P. Then, for every $y \in Y$ and $p \in P$*

$$p \in \partial C(y) \Leftrightarrow \left(y \in \partial C^{\#}(p) \text{ and } C^{\#\#}(y) = C(y)\right). \tag{C.31}$$

For a concave function Π (on a space K paired with R), this becomes

$$r \in \partial\Pi(k) \Leftrightarrow (k \in \partial\Pi_{\#}(r) \text{ and } \Pi_{\#\#}(k) = \Pi(k)). \tag{C.32}$$

Proof. This follows from the Fenchel-Young Inequality and from the case of equality therein as a characterization of the subdifferential: apply (C.12)–(C.13) twice, to C and to $C^{\#}$ (in place of C), to see that the conditions $p \in \partial C(y)$ and $y \in \partial C^{\#}(p)$ are equivalent when $C^{\#\#}(y) = C(y)$. It remains to show that this equality holds when $p \in \partial C(y)$. And this is because, by (C.2) and by (C.12)–(C.13) applied to $C^{\#}$, $C^{\#\#}(y) \geq \langle p \mid y \rangle - C^{\#}(p) = C(y) \geq C^{\#\#}(y)$ by (C.4). □

[143] This is given in, e.g., [4, 4.4.4], [42, 23.5 (a) and (a*)] and [44, Corollary 12A].

The Inversion Rule and the First-Order Condition (C.12) are next combined in a derivative property of conjugate functions. In convex programming, this yields the derivative property of the optimal value (in the same way as is shown here for the case of profit or cost programmes and their duals, in Section 19).

Corollary C.3 (Derivative Property of the Conjugate)**.** *Assume that* $C\colon Y \to \mathbb{R}\cup\{\pm\infty\}$ *is convex (and* Y *is paired with* P*). Then, for every* $y \in Y$ *and* $p \in P$*,*

$$y \text{ maximizes } \langle p \,|\, \cdot \rangle - C \Leftrightarrow \left(y \in \partial C^{\#}(p) \text{ and } C^{\#\#}(y) = C(y)\right). \tag{C.33}$$

When C *is lower semicontinuous proper convex on* Y*, this means that*

$$\partial C^{\#}(p) = \operatorname{argmax}\left(\langle p \,|\, \cdot \rangle - C\right) \tag{C.34}$$

for every $p \in P$*.*[144]

Proof. The equivalence (C.33) follows from the FOC (C.12) and the Inversion Rule (C.31). And (C.34) follows from (C.33) because $C^{\#\#} = C$ in this case. □

The convex conjugate of the 0-∞ indicator $\delta(\cdot \mid Z)$ of a set $Z \subseteq Y$ (i.e., of the function equal to 0 on Z and $+\infty$ on $Y \setminus Z$) is the *support function* of Z, i.e.,

$$\delta^{\#}(p \mid Z) = \sup_{Z} \langle p \,|\, \cdot \rangle \tag{C.35}$$

and the Derivative Property (C.34) gives its subdifferential at a $p \in P$ as

$$\partial\delta^{\#}(p \mid Z) = \operatorname*{argmax}_{Z} \langle p \,|\, \cdot \rangle \tag{C.36}$$

if Z is nonempty, convex and closed. This is stated in, e.g., [42, 23.5.3] and [44, p. 36, lines 1–7]. Similarly, the *inf-support function* of a set $Z \subseteq R$ is the concave conjugate of $-\delta(\cdot \mid Z)$, i.e.,

$$\inf_{Z} \langle \cdot \,|\, k \rangle = (-\delta)_{\#}(k \mid Z) \tag{C.37}$$

for every $k \in K$ (the space paired with R). Its superdifferential at k is

$$\widehat{\partial}(-\delta)_{\#}(k \mid Z) = \operatorname*{argmin}_{Z} \langle \cdot \,|\, k \rangle \tag{C.38}$$

if Z is nonempty, convex and closed.

Comment (proper and improper solutions): As in [45], $\operatorname{argmax}_Z f$ means the set of all maximum points of a function f on a set Z—provided that $\sup_Z f > -\infty$. Points of $\operatorname{argmax}_Z f$ maximize f *properly* (i.e., either to a finite value or to $+\infty$). When $f = -\infty$ on Z, any point of Z maximizes f on Z, but $\operatorname{argmax}_Z f := \emptyset$. In other words, when a programme is infeasible, it is convenient to regard any point as an *improper solution*, as in [44, p. 38]. But note that in a dual pair of solutions with *equal* values both solutions are always proper (i.e., are feasible) or, equivalently, their common value is finite. To see this, let the primal programme be to maximize a concave $f\colon X \to \mathbb{R}\cup\{-\infty\}$; then the dual is to minimize a certain convex $g\colon Y \to \mathbb{R}\cup\{+\infty\}$ such that $f(x) \le g(y)$ for every

[144] This is given in, e.g., [4, 4.4.5], [42, 23.5 (b) and (a*)] and [44, Corollary 12B]. It holds formally also when C is the constant $-\infty$ (but not when C is $+\infty$ because $\operatorname{argmax}(-\infty) := \emptyset$ by convention, whilst $\partial(-\infty)(p) := Y$).

x and y (where X and Y are vector spaces). If $\overline{x}$ maximizes f, $\overline{y}$ minimizes g and there is no duality gap, then $+\infty > f(\overline{x}) = g(\overline{y}) > -\infty$ (so $\overline{x} \in \operatorname{argmax} f$ and $\overline{y} \in \operatorname{argmin} g$).[145]

The support function of a nonempty set Z is sublinear—i.e., it is convex and positively linearly homogeneous (p.l.h.) or, equivalently, it is p.l.h. and subadditive. Conversely, every l.s.c. sublinear function $C\colon Y \to \mathbb{R} \cup \{+\infty\}$ is the support function of a nonempty, convex and closed set, viz., $\partial C(0)$—i.e.,

$$C(y) = \sup_{p \in \partial C(0)} \langle p \mid y \rangle \tag{C.39}$$

$$\text{where} \quad \partial C(0) := \{p : \langle p \mid y \rangle \le C(y)\}. \tag{C.40}$$

See, e.g., [32, 4.1: Proposition 1], [42, 13.2.1] or [48, 6.22]. By (C.36), it follows that

$$\partial C(y) := \left\{ p \in \partial C(0) : \langle p \mid y \rangle = \sup_{\partial C(0)} \langle \cdot \mid y \rangle = C(y) \right\} \tag{C.41}$$

which is stated in, e.g., [32, 4.2.1: Example 3], [42, 23.5.3] and [44, p. 36, lines 1–7]. This is a variant of Euler's Theorem on homogeneous functions.

C.7. **Subgradients of partial conjugates.** In the case of *partial* conjugacy, between a bivariate convex function C and a saddle (convex-concave) function Π, the Inversion Rule not only applies to the relevant partial derivatives but also extends to the *total* derivatives (Corollaries C.6 and C.8 below). Namely, when Π and C are differentiable, their gradient maps can be obtained from each other by transposition of that pair of variables, p and y, w.r.t. which Π and C are mutual conjugates. When Π and C are nondifferentiable, the rule applies to their subdifferential correspondences—i.e., to the "saddle differential" $\partial_p \Pi \times \widehat{\partial}_k \Pi$ and the joint subdifferential $\partial_{y,k} C$ (which does not usually factorize into $\partial_y C \times \partial_k C$). This rule is based on a key lemma, useful also by itself,[146] which identifies the section of the joint subdifferential $\partial_{y,k} C$ through a $p \in \partial_y C$ as $-\widehat{\partial}_k \Pi$, the partial subdifferential of $-\Pi$ w.r.t. the argument k that it shares with C (Lemma C.5).

These relationships between a saddle function Π and its bivariate convex "parent" C are spelt out below. First, since Π is the partial conjugate of C w.r.t. one variable, the total (bivariate) conjugate of C is the partial conjugate of $-\Pi$ w.r.t. the other variable.

Lemma C.4 (Total conjugacy by stages)**.** *Assume that $C\colon Y \times K \to \mathbb{R} \cup \{\pm\infty\}$ and let the spaces Y and K be paired with P and R. Then, in the notation of (C.6),*

$$C^{\#} = \left(-C^{\#_1}\right)^{\#_2}$$

on $P \times R$. In other words, if

$$\Pi(p,k) = C^{\#_1}(p,k) := \sup_y \left(\langle p \mid y \rangle - C(y,k)\right) \tag{C.42}$$

[145] This argument assumes that the maximand f is nowhere $+\infty$ and that the minimand g is nowhere $-\infty$. These sensible conditions are met when the perturbed primal constrained maximand, F, is a u.s.c. proper concave function on a space $X \times A$ paired with $B \times Y$ (where A and B are the spaces of primal and dual perturbations). This is because: (i) $f(x) = F(x,0) < +\infty$ for every x, and (ii) the perturbed dual constrained maximand, $G(b,y) := -F_{\#}(-b,y)$, is then l.s.c. proper convex, and so $g(y) := G(0,y) > -\infty$ for every y: see, e.g., [44, (4.17)].

[146] For example, it yields the extension (11.1) of the Wong-Viner Theorem.

for every $p \in P$ and $k \in K$, then

$$C^{\#}(p, -r) = (-\Pi)^{\#_2}(p, -r) := \sup_k (\Pi(p, k) - \langle r \mid k \rangle)$$

for every $p \in P$ and $r \in R$.

Proof. For every $(p, r) \in P \times R$

$$\begin{aligned} C^{\#}(p, -r) &= \sup_{y,k} (\langle p \mid y \rangle - \langle r \mid k \rangle - C(y, k)) = \sup_k \left(- \langle r \mid k \rangle + \sup_y (\langle p \mid y \rangle - C(y, k)) \right) \\ &= \sup_k (\Pi(p, k) - \langle r \mid k \rangle) \end{aligned}$$

as required. □

Comment ("staged" conjugacy and alternative proofs of the inequality between partial and total second conjugates): Also the second conjugate can be obtained in stages, i.e.,

$$C^{\#\#} = C^{\#_1\#_1\#_2\#_2}.$$

That is, the total second conjugate of C is equal to the partial second conjugate, w.r.t. one variable, of the partial second conjugate of C w.r.t. the other variable. This gives another proof of the first inequality in (C.8): $C^{\#\#} = C^{\#_1\#_1\#_2\#_2} \le C^{\#_1\#_1}$ (by (C.4) applied to the function $C^{\#_1\#_1}(y, \cdot)$ on K, in place of C). Similarly, in terms of the partial second concave conjugate of $\Pi := C^{\#_1}$ w.r.t. the second variable, $C^{\#\#} = \left(\left(C^{\#_1} \right)_{\#_2\#_2} \right)^{\#_1} \le C^{\#_1\#_1}$ (because $\Pi_{\#_2\#_2} \ge \Pi$).

The "staged" conjugacy is next used to "slice" the joint subdifferential of the bivariate convex function along one of the "axes" (the p-axis): the section of the set $\partial C(y, k) \subseteq P \times R$ through any $p \in \partial_y C(y, k)$ is found to be $-\widehat{\partial}_k \Pi(p, k) \subseteq \partial_k C(y, k) \subseteq R$.

Lemma C.5 (Subdifferential sections)**.** *Assume that $C\colon Y \times K \to \mathbb{R} \cup \{+\infty\}$ is proper convex, and that $\Pi\colon P \times K \to \mathbb{R} \cup \{\pm\infty\}$ is the partial convex conjugate of C, i.e., (C.42) holds for each k in K (which is paired with a space R). Then the following conditions are equivalent to each other:*

1. *$(p, -r) \in \partial C(y, k)$.*
2. *$p \in \partial_y C(y, k)$ and $r \in \widehat{\partial}_k \Pi(p, k)$.*

Also, either condition implies that both $C(y, k)$ and $\Pi(p, k)$ are finite.

Proof. Since $C^{\#} = (-\Pi)^{\#_2}$ by Lemma C.4, and since $\Pi := C^{\#_1}$ by (C.42), one has by (C.1)

$$\langle p \mid y \rangle - C(y, k) \le \Pi(p, k) \tag{C.43}$$

$$- \langle r \mid k \rangle + \Pi(p, k) \le C^{\#}(p, -r) \tag{C.44}$$

as well as

$$\langle p \mid y \rangle - \langle r \mid k \rangle - C(y, k) \le C^{\#}(p, -r) \tag{C.45}$$

for every p, y, r and k. By (C.13), Condition 1 is equivalent to equality in (C.45), which holds if and only if equalities hold in both (C.43) and (C.44). Finally, the pair of equalities is equivalent to Condition 2, again by (C.13).

It remains to show that the equivalent Conditions, 1 and 2, imply that $C(y,k)$ and $\Pi(p,k)$ are finite (as is also $C^{\#}(p,-r)$). For a start, note that, by assumption, C does not take the value $-\infty$, and neither does $C^{\#}$ (since C is not the constant $+\infty$). But both C and $C^{\#}$ can take the value $+\infty$. As for Π, it can take both infinite values, although for no p can the concave function $\Pi(p,\cdot)$ be the constant $-\infty$.[147]

Assume, say, Condition 1—i.e., that equality holds in (C.45). Since $C(y,k)$ is either finite or $+\infty$, and since so is $C^{\#}(p,-r)$, both $C(y,k)$ and $C^{\#}(p,-r)$ are actually finite (since they add up to $\langle p \mid y\rangle - \langle r \mid k\rangle$, which is finite). Given this, the inequalities (C.44) and (C.43) show that $\Pi(p,k)$ is also finite.

It is equally easy to argue from Condition 2: if equalities hold in (C.43) and (C.44), then

$$\Pi(p,k) = \langle p \mid y\rangle - C(y,k) < +\infty$$
$$\Pi(p,k) = C^{\#}(p,-r) + \langle r \mid k\rangle > -\infty$$

so $\Pi(p,k)$ is finite;[148] and hence so are $C(y,k)$ and $C^{\#}(p,-r)$. □

Finally, the Inversion Rule is applied to the partial subdifferential ($\partial_y C$) that is the range of the variable (p) indexing the sections of the joint subdifferential (∂C) in Lemma C.5. The result shows that, up to a sign change, the saddle-differential and the joint-subdifferential correspondences ($\partial_p \Pi \times \widehat{\partial}_k \Pi$ and $\partial_{y,k} C$) are partial inverses of each other: their graphs are identical.

Corollary C.6 (Partial Inversion Rule)**.** *Under the assumptions of Lemma C.5, the following conditions are equivalent to each other:*[149]

(1) $(p,-r) \in \partial C(y,k)$.
(2) $y \in \partial_p \Pi(p,k)$ *and* $r \in \widehat{\partial}_k \Pi(p,k)$, *and* $C(\cdot,k)$ *is finite and lower semicontinuous at* y.

Also, either condition implies that both $C(y,k)$ *and* $\Pi(p,k)$ *are finite.*

Proof. By Lemma C.5, if $(p,-r) \in \partial C(y,k)$ then, in addition to $r \in \widehat{\partial}_k \Pi(p,k)$ and $C(y,k) < +\infty$, one has $p \in \partial_y C(y,k)$. By the Inversion Rule (C.31) and (C.4), this implies that $y \in \partial_p \Pi(p,k)$ and that $C(\cdot,k)$ is l.s.c. at y. So Condition 1 implies Condition 2.

For the converse, since $C(y,k) < +\infty$ and $C(\cdot,k)$ is l.s.c. at y, one has $C(y,k) = C^{\#_1\#_1}(y,k)$. So if $y \in \partial_p \Pi(p,k)$ then $p \in \partial_y C(y,k)$ by the Inversion Rule (C.31). And if additionally $r \in \widehat{\partial}_k \Pi(p,k)$, then $(p,-r) \in \partial C(y,k)$ by Lemma C.5. □

Comments (on the PIR and SSL):

(1) Finiteness of $C(y,k)$ can be dropped from Condition 2 (and the proof of its equivalence to Condition 1 simplifies) if either (i) $C(\cdot,k)$ is assumed to be l.s.c.

[147] What is more, for every $k \in K$ either (i) $\Pi(\cdot,k) = -\infty$ (everywhere on P), or (ii) $\Pi(\cdot,k)$ does not take the value $-\infty$ (anywhere on P). The latter is the case for some k (since $C(\cdot,k) \neq +\infty$ for some k); and so $\Pi(p,\cdot) \neq -\infty$ for every $p \in P$.

[148] That $\Pi(p,k) > -\infty$ can also be deduced from $r \in \widehat{\partial}_k \Pi(p,k)$, since $\Pi(p,\cdot) \neq -\infty$.

[149] This is in, e.g., [4, 4.4.14], [41, Lemma 4], [42, 37.5] and [45, 11.48].

on the whole space Y (and not just at the particular point y), or (ii) Y is finite-dimensional. This is because, in either case, the assumption (of Lemmas C.5 and C.6) that $C(\cdot,k) > -\infty$ on Y implies that $\operatorname{lsc}(C(\cdot,k)) > -\infty$ on Y (when Y is finite-dimensional, this follows from [42, 7.5]). Therefore $\operatorname{lsc}(C(\cdot,k)) = C^{\#_1\#_1}(\cdot,k)$ on Y, and so the Inversion Rule (C.31) shows that $p \in \partial_y C(y,k)$ if and only if both $y \in \partial_p \Pi(p,k)$ and $C(\cdot,k)$ is l.s.c. at y. Thus Corollary C.6 reduces immediately to Lemma C.5.

(2) There is a structural difference between the Subdifferential Sections Lemma and the Partial Inversion Rule. The SSL turns the condition $(p,-r) \in \partial_{y,k} C$ into a pair of conditions like $p \in \partial_y C$ and $r \in \widehat{\partial}_k \Pi$—which involve two functions but use partial subdifferentials w.r.t. the *same* variables as in the joint subdifferential. The PIR turns the condition $(p,-r) \in \partial_{y,k} C$ into the pair of conditions $y \in \partial_p \Pi$ and $r \in \widehat{\partial}_k \Pi$. These use a single function Π, but only one of its arguments (k) is the same as in the original function C: the other argument (y) is replaced by its dual (p) in inverting $\partial_y C$ into $\partial_p \Pi$. This step requires the semicontinuity of C w.r.t. y—and this is why the PIR is not purely algebraic like the SSL.

Remark C.7. *Under the assumptions of Lemma C.5,*

$$\widehat{\partial}_k \Pi(p,k) \subseteq -\partial_k C(y,k) \quad \text{when } p \in \partial_y C(y,k) \tag{C.46}$$

i.e., when y yields the supremum defining Π in (C.42).

Proof. Since

$$\partial C(y,k) \subseteq \partial_y C(y,k) \times \partial_k C(y,k) \tag{C.47}$$

$\partial_k C(y,k)$ contains the section of $\partial C(y,k)$ through any $p \in \partial_y C(y,k)$. And this section is $-\widehat{\partial}_k \Pi(p,k)$ by Lemma C.5. □

Comments:

(1) A simpler proof of (C.46) comes straight from the definition (C.42):

$$\Pi(p, k + \Delta k) \ge \langle p \mid y \rangle - C(y, k + \Delta k) \text{ for every } \Delta k$$

with equality at $\Delta k = 0$. In other words, the graph of the convex function $-\Pi(p,\cdot)$ lies below that of $C(y,\cdot) + \text{const.}$, touching it at k. It follows that $-\widehat{\partial}_k \Pi(p,k)$ is a subset of $\partial_k C(y,k)$, although this "envelope argument" does not show it ($-\widehat{\partial}_k \Pi$) to be a section of $\partial C(y,k)$ through p.

(2) The inclusion (C.47) is usually "tight" in the sense that $\partial_y C \times \partial_k C$ is the smallest Cartesian product set encasing ∂C: by (C.47) itself, $\partial_y C$ and $\partial_k C$ contain the projections of ∂C (onto P and R), and the reverse inclusions can be obtained by using the Hahn-Banach Extension Theorem (Theorem B.3 or Corollary B.4).

For a saddle function S with a (bivariate) convex parent I, the following useful variant of Corollary C.6 transposes the saddle differential correspondence ∂S into $\partial I^{\#}$ instead of ∂I (i.e., into the subdifferential correspondence of I's total conjugate instead of I itself).

Corollary C.8 (Dual Partial Inversion Rule)**.** *Assume that $I\colon Y \times V \to \mathbb{R} \cup \{+\infty\}$ is proper convex and (jointly) lower semicontinuous for the pairing of the space V with W*

(and Y with P), and that $-S\colon Y \times W \to \mathbb{R} \cup \{\pm\infty\}$ is $I^{\#_2}$ (the partial convex conjugate of I), i.e.,

$$S(y,w) = \inf_u \left(I(y,u) - \langle w \mid u \rangle \right) \tag{C.48}$$

for every $y \in Y$ and $w \in W$. Then the following conditions are equivalent to each other:

(1) *$(y,u) \in \partial I^{\#}(p,w)$.*

(2) *$p \in \partial_y S(y,w)$ and $-u \in \widehat{\partial}_w S(y,w)$.*

Also, either condition implies that both $I(p,w)$ and $S(y,w)$ are finite.

Proof. Since $I^{\#\#} = I$ by the assumption that I is l.s.c., the Inversion Rule (C.31) shows that Condition 1 is equivalent to: $(p,w) \in \partial I(y,u)$. And this is equivalent to Condition 2 by the Partial Inversion Rule (Corollary C.6) and the first Comment thereafter. □

Comment (on another derivation of DPIR): By Lemma C.4, the convex function $I^{\#}$ is a partial conjugate of the saddle function S; and when this relationship can be inverted to represent S as a partial conjugate of $I^{\#}$, the equivalence of $\partial I^{\#}$ and ∂S follows from the Partial Inversion Rule alone. But this argument requires $S(\cdot,w)$ to be l.s.c. on Y, and this is a condition that S can actually fail at some points (even when I is l.s.c.). Corollary C.8 obviates the need to ensure that S is l.s.c. in y.

References

[1] Afriat, S. N. (1971): "Theory of maxima and the method of Lagrange", *SIAM Journal on Applied Mathematics*, 20, 343–357.

[2] Aliprantis, C., and O. Burkinshaw (1985): *Positive operators*. New York-London: Academic Press.

[3] Anderson, E. J., and P. Nash (1987): *Linear programming in infinite-dimensional spaces*. New York-Chichester-Brisbane-Toronto-Singapore: Wiley.

[4] Aubin, J. P., and I. Ekeland (1984): *Applied nonlinear analysis*. New York: Wiley.

[5] Bair, J., and R. Fourneau (1975): *Etude géometrique des espaces vectoriels* (Lecture Notes in Mathematics, vol. 489). Berlin-Heidelberg-New York: Springer-Verlag.

[6] Berge, C. (1963): *Topological spaces*. Edinburgh: Oliver and Boyd.

[7] Bewley, T. (1972): "Existence of equilibria in economies with infinitely many commodities", *Journal of Economic Theory*, 4, 514–540.

[8] Birkhoff, G. (1967): *Lattice theory*. Providence, RI: American Mathematical Society.

[9] Boiteux, M. (1964): "Peak-load pricing", in *Marginal cost pricing in practice* (Chapter 4), ed. by J. R. Nelson. Engelwood Cliffs, NJ: Prentice Hall. (A translation of "La tarification des demandes en pointe: application de la théorie de la vente au cout marginal", *Revue Général de l'Electricité*, 58 (1949), 321–340.)

[10] Boutacoff, D. (1989): "Emerging strategies for energy storage", *Electric Power Research Institute Journal*, vol. 14, No. 5, 4–13.

[11] Chvatal, V. (1983): *Linear programming*. New York: Freeman.

[12] Craven, B. D. (1978): *Mathematical programming and control theory*. London: Chapman and Hall.

[13] Diewert, W. E. (1974): "Applications of duality theory", in *Frontiers of quantitative economics*, Volume II (Chapter 3), ed. by M. D. Intriligator and D. A. Kendrick. Amsterdam: North-Holland.

[14] Diewert, W. E. (1982): "Duality approaches to microeconomic theory", in *Handbook of mathematical economics*, Volume II (Chapter 12), ed. by K. J. Arrow and M. D. Intriligator. Amsterdam: North-Holland.

[15] Drèze, J. H. (1964): "Some postwar contributions of French economists to theory and public policy", *American Economic Review*, 54 (supplement, June 1964), 1–64.

[16] Foran, J. (1991): *Fundamentals of real analysis*. New York-Basel-Hong Kong: Dekker.

[17] Henderson, J. M., and R. E. Quandt (1971): *Microeconomic theory*. New York-London: McGraw-Hill.
[18] Holmes, R. B. (1975): *Geometric functional analysis and its applications*. Berlin-Heidelberg-New York: Springer.
[19] Horsley, A., and A. J. Wrobel (1988): "Subdifferentials of convex symmetric functions: An application of the Inequality of Hardy, Littlewood and Pólya", *Journal of Mathematical Analysis and Applications*, 135, 462–475.
[20] Horsley, A., and A. J. Wrobel (1993): "Uninterruptible consumption, concentrated charges, and equilibrium in the commodity space of continuous functions", STICERD Discussion Paper TE/96/300, LSE (presented at ESEM93, Uppsala).
[21] Horsley, A., and A. J. Wrobel (1996): "Efficiency rents of storage plants in peak-load pricing, I: pumped storage", STICERD Discussion Paper TE/96/301, LSE. (This is a fuller version of Ref. [27].)
[22] Horsley, A., and A. J. Wrobel (1996): "Comparative statics for a partial equilibrium model of investment with Wicksell-complementary capital inputs", STICERD Discussion Paper TE/96/302, LSE.
[23] Horsley, A., and A. J. Wrobel (1999): "Efficiency rents of hydroelectric storage plants in continuous-time peak-load pricing", in *The current state of economic science* (Volume 1, pp. 453–480), ed. by S. B. Dahiya. Rohtak: Spellbound Publications.
[24] Horsley, A., and A. J. Wrobel (1999): "Efficiency rents of storage plants in peak-load pricing, II: hydroelectricity", STICERD Discussion Paper TE/99/372, LSE. (This is a fuller version of Ref. [23].)
[25] Horsley, A., and A. J. Wrobel (2000): "Localisation of continuity to bounded sets for nonmetrisable vector topologies and its applications to economic equilibrium theory", *Indagationes Mathematicae* (New Series), 11, 53–61.
[26] Horsley, A., and A. J. Wrobel (2002): "Boiteux's solution to the shifting-peak problem and the equilibrium price density in continuous time", *Economic Theory*, 20, 503–537.
[27] Horsley, A., and A. J. Wrobel (2002): "Efficiency rents of pumped-storage plants and their uses for operation and investment decisions", *Journal of Economic Dynamics and Control*, 27, 109–142.
[28] Horsley, A., and A. J. Wrobel (2005): "Continuity of the equilibrium price density and its uses in peak-load pricing", *Economic Theory*, 26, 839–866.
http://www.springerlink.com/index/10.1007/s00199-004-0568-3
[29] Horsley, A., and A. J. Wrobel (2005): "Demand continuity and equilibrium in Banach commodity spaces", CDAM Research Report LSE-CDAM-2005-01.
http://www.cdam.lse.ac.uk/Reports/reports2005.html
[30] Horsley, A., and A. J. Wrobel (2005): "Efficiency rents of a hydroelectric storage plant with a variable head", forthcoming CDAM Research Report, LSE.
[31] Horsley, A., A. J. Wrobel, and T. Van Zandt (1998): "Berge's Maximum Theorem with two topologies on the action set", *Economics Letters*, 61, 285–291.
[32] Ioffe, A. D., and V. M. Tihomirov (1979): *Theory of extremal problems*. Amsterdam-New York-Oxford: North-Holland.
[33] Kantorovich, L. W., and P. G. Akilov (1982): *Functional analysis*. Pergamon Press: Oxford.
[34] Klein, E., and A. C. Thompson (1984): *Theory of correspondences*. New York-Chichester-Brisbane-Toronto: Wiley.
[35] Koopmans, T. C. (1957): "Water storage policy in a simplified hydroelectric system", in *Proceedings of the First International Conference on Operational Research*, pp. 193–227. London-Baltimore. (Also in Koopmans, T. C. (1970): *Scientific papers of Tjalling C. Koopmans*, pp. 282–316. Berlin-Heidelberg: Springer.)
[36] Laurent, P.-J. (1972): *Approximation et optimisation*. Paris: Hermann.
[37] Levin, V. L. (1985): *Convex analysis in spaces of measurable functions and its applications to mathematics and economics* (in Russian). Moscow: Nauka.

[38] Marsh, W. D. (1980): *Economics of electric utility power generation*. Oxford-New York: Oxford University Press-Clarendon Press.
[39] Meyer-Nieberg, P. (1991): *Banach lattices*. Berlin-Heidelberg-New York: Springer.
[40] Moore, T. (1994): "Storing megawatthours with SMES", *Electric Power Research Institute Journal*, vol. 19, No. 5, 24–33.
[41] Rockafellar, R. T. (1968): "A general correspondence between dual minimax problems and convex programs", *Pacific Journal of Mathematics*, 25, 597–611.
[42] Rockafellar, R. T. (1970): *Convex analysis*. Princeton, NJ: Princeton University Press.
[43] Rockafellar, R. T. (1970): "Conjugate convex functions in optimal control and the calculus of variations", *Journal of Mathematical Analysis and Applications*, 32, 174–222.
[44] Rockafellar, R. T. (1974): *Conjugate duality and optimization*. Philadelphia, PA: SIAM.
[45] Rockafellar, R. T., and R. J. B. Wets (1997): *Variational analysis*. Berlin-Heidelberg-New York: Springer.
[46] Rudin, W. (1973): *Functional analysis*. New York: McGraw-Hill.
[47] Takayama, A. (1985): *Mathematical economics*. Cambridge-London-New York: Cambridge University Press.
[48] Tiel, J. van (1984): *Convex analysis*. Chichester-New York-Brisbane: Wiley.

(Anthony Horsley and Andrew J. Wrobel) Department of Economics, London School of Economics, Houghton Street, London WC2A 2AE, United Kingdom
E-mail address: `LSEecon123@mac.com`