

Discussion Papers

A Simple Regulatory Incentive Mechanism Applied to Electricity Transmission Pricing and Investment

Opinions expressed in this paper are those of the author(s) and do not necessarily reflect views of the institute.

IMPRESSUM

© DIW Berlin, 2017

DIW Berlin
German Institute for Economic Research
Mohrenstr. 58
10117 Berlin

Tel. +49 (30) 897 89-0
Fax +49 (30) 897 89-200
<http://www.diw.de>

ISSN electronic edition 1619-4535

Papers can be downloaded free of charge from the DIW Berlin website:
<http://www.diw.de/discussionpapers>

Discussion Papers of DIW Berlin are indexed in RePEc and SSRN:
<http://ideas.repec.org/s/diw/diwwpp.html>
<http://www.ssrn.com/link/DIW-Berlin-German-Inst-Econ-Res.html>

A Simple Regulatory Incentive Mechanism Applied to Electricity Transmission Pricing and Investment

Mohammad Reza Hesamzadeh,^{*} Juan Rosellon,[†] Steven A. Gabriel,[‡] Ingo Vogelsang[§]

18 October 2017

Abstract

The informationally simple approach to incentive regulation applies mechanisms that translate the regulator's objective function into the firm's profit-maximizing objective. These mechanisms come in two forms, one based on subsidies/taxes, the other based on constraints/price caps. In spite of a number of improvements and a good empirical track record simple approaches so far remain imperfect. The current paper comes up with a new proposal, called H-R-G-V, which blends the two traditions and is shown to apply well to electricity transmission pricing and investment. In particular, it induces immediately optimal pricing/investment but is not based on subsidies. In the transmission application, the H-R-G-V approach is based on a bilevel optimization with the transmission company (Transco) at the top and the independent system operator (ISO) at the bottom level. We show that H-R-G-V, while not perfect, marks an improvement over the other simple mechanisms and a convergence of the two traditions. We suggest ways to deal with remaining practical issues of demand and cost functions changing over time.

Keywords: Electricity transmission, incentive regulation, multi-level optimization.

JEL codes: D24, L51, L94, Q40

^{*}Electricity Market Research Group (EMReG), KTH Royal Institute of Technology, Stockholm, Sweden (mrhesamzadeh@ee.kth.se)

[†]CIDE, Department of Economics, Carretera Mexico-Toluca 3655 Col. Lomas de Santa Fe 01210 Mexico, D.F., Mexico (juan.rosellon@cide.edu), DIW Berlin, EVU (jrosellon@diw.de), and Universidad Panamericana, Mexico.

[‡]Applied Math and Statistics and Scientific Computation Program, University of Maryland, College Park, USA (sgabriel@umd.edu)

[§]Department of Economics, Boston University, Boston, USA (vogelsang@bu.edu)

1 Introduction and Motivation

The purpose of regulatory incentive mechanisms is to influence the profit function of the regulated firm in such a way that profit maximization leads to goal fulfillment of the regulator, where the objective is usually assumed to be the maximization of welfare in the form of social surplus. Such mechanisms are desirable, because the regulator typically is less informed about costs and demands facing the firm and because the regulator can only do limited monitoring and enforcement. The mechanisms should therefore be easily enforceable. In particular, such mechanisms can induce socially optimal electricity transmission investments.

Two basic types of approaches have been developed in the academic literature. The informationally sophisticated or Bayesian approach has dominated the theoretical literature, starting with Baron and Myerson (1982) and Laffont and Tirole (1986). Informationally demanding mechanisms capture uncertainty and asymmetric information by a subjective probability distribution of types of firms, where the regulator only knows the distribution, while the firm also knows its own type. The mechanisms are called Bayesian, because regulators start with a subjective *a priori* type distribution of firms and they use Bayesian updating to reach posterior distributions.

The main drawbacks of the Bayesian approach are (1) that regulators cannot be monitored well by the public, because the distribution of firm types is based on subjective probabilities, which makes it somewhat arbitrary (Koray and Saglam, 2005), and (2) because real-world functional forms for type distributions, costs and demands are largely unknown it is hard to get realistic quantitative results, which would make it highly complex for an application.

However, the sophisticated approach provides strong insights into the incentive properties of regulation. For example, (1) firms have to receive an information rent in order to be induced to reveal their type, or (2) it is impossible to reach a first-best outcome, or (3) the less the regulator can commit to future policies the weaker incentives should be. In sum, this approach is not very practical but highly insightful.

In contrast, the informationally simple or non-Bayesian approach does not use a formal

probabilistic model but applies mechanisms that translate the regulator's objective function into the firm's profit-maximizing objective. These mechanisms are typically quite practical, easy to understand and are at least partially based on observable or even verifiable data. They come in two forms, one based on subsidies/taxes the other based on constraints/price caps. The subsidy approach goes back to Loeb and Magat (Loeb and Magat, 1979), while the constraint/price cap approach goes back to Vogelsang and Finsinger (Vogelsang and Finsinger, 1979). In spite of a number of improvements and a good empirical track record, simple approaches so far remain imperfect relative to well-defined social welfare maximization. In particular, regulators generally do not have the power to grant subsidies or to impose taxes, while constraint-based mechanisms are unable to reach optimal outcomes in an environment with demand and cost functions changing over time.

The current paper comes up with a new proposal, the H-R-G-V mechanism, which blends the mechanisms in the L-M tradition with those in the V-F tradition and is shown to be effective in electricity transmission pricing and investment. In particular, it is not based on subsidies and it gets rid of the issue of slow or no convergence that has plagued the constraint-based mechanisms. The H-R-G-V mechanism requires the regulator to have only local information of demand but no information on costs. While local demand information may be hard to come by in other industries, Gans and King (2000) point out that such information is readily available for electricity transmission networks based on nodal pricing. At the same time electricity transmission pricing and investment are very timely issues because of the restructuring of electricity from traditional fossil fuel-based generation to renewables. We therefore provide a detailed application of the H-R-G-V mechanism to electricity transmission pricing and investment, using simulations to show optimality properties of the H-R-G-V mechanism.

This paper is structured as follows. In Section 2 we provide a short history of simple regulatory incentive mechanisms. Section 3 then motivates two-part tariffs as a bridge between subsidy-based and constraint-based mechanisms. Section 4 shows the main theoretical properties of the H-R-G-V mechanism, contrasting it in particular to a two-part tariff version of

Sappington's and Sibley's ISS mechanism. While Sections 2 to 4 assume a very simplified context of a monopoly firm facing only end-users as customers and without specifically modeling investments, Section 5 is an application of H-R-G-V to electricity network investment, where transmission is an intermediate input placed in the electricity value chain between generation and distribution companies, while the sale of electricity to end-users is not explicitly modeled. A more general application of H-R-G-V to electricity network investment is presented in Section 6. Section 7 provides numerical results. Section 8 concludes the paper and discusses some further thoughts.

2 A short history of simple regulatory incentive mechanisms¹

In 1979, Loeb and Magat (1979) published their famous Loeb-Magat (L-M) mechanism, which offered marginal cost pricing and cost-minimizing incentives under a subsidy scheme that handed over consumer surplus as a subsidy to the firm. In the same year, Vogelsang and Finsinger (1979) published their V-F mechanism, which was a discrete dynamic adjustment process that did not require subsidies and led to Ramsey pricing with zero profits without requiring the regulator to have demand and service-specific cost information.² While being simple and while dealing well with asymmetric information between the regulator and the regulated firm, both mechanisms had severe drawbacks limiting their practical application. The L-M mechanism required the regulator to know the demand function for the firm's services and required potentially huge subsidies, represented by the area between the price line and the demand curve. Since demand is not easily observable, there can be severe disputes about the size of the subsidy. In contrast to L-M, the V-F mechanism required no demand information but provided poor cost-minimizing incentives and took time to converge to Ramsey prices, meaning that it could be subject to strategic manipulation (Sappington, 1980) and would perform badly under changing cost and demand

¹For an exhaustive history of both Bayesian and non-Bayesian mechanisms see Armstrong and Sappington (2007).

²Ramsey prices maximize social welfare subject to the constraint that the regulated firm at least breaks even. Such prices vary inversely with the price elasticities of the services provided by the firm.

conditions (Neu, 1993) and (Fraser, 1995). Subsequent mechanisms were partially able to deal with these problems.

The main new development starting from L-M as the basis was the incremental surplus subsidy (ISS) scheme by Sappington and Sibley (1988). It gave the firm a subsidy (or charged a tax) amounting to the change in consumer surplus over last period minus last period's accounting profit based on market revenues and expenses, while the firm collected current period market revenues and had to pay current period expenses.³

$$ISS_t = V(p_t) - V(p_{t-1}) - \pi_{t-1} \quad (1)$$

Here $V(p)$ stands for consumer surplus and π for the firm's market profit under linear pricing.⁴ As a result, in each period the firm's after-subsidy profits Π_t^{ISS} would equal the change in total surplus ΔW_t over last period. It is represented by area ACB in Figure 1.

$$\Pi_t^{ISS} = ISS_t + \pi_t = \Delta V_t + \Delta \pi_t = \Delta W_t \quad (2)$$

This mechanism still used subsidies but much less so than L-M, while converging to marginal cost prices in a single period. The regulator still had to know demand, however only in the neighborhood of last period's and current prices. The cost-minimizing incentives were viewed as fairly weak (Blackmon, 1992, 1994), certainly not as strong as under L-M but definitely stronger than under V-F. Gans and King (2000) suggested the ISS as the basis for electricity transmission pricing and investment. They argued that in the environment of electricity dispatch

³For simplicity we assume at this point that there is only a single output, but, as can be seen below, all our arguments extend to the multi-product case.

⁴Because of their better observability Sappington and Sibley (1988) use current expenses instead of partially unobservable economic costs. Thus, π may not be strictly interpreted as profit.

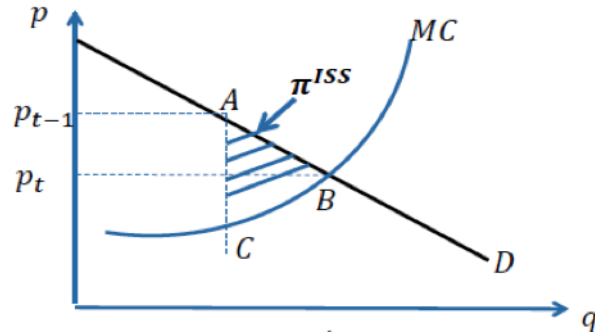


Figure 1: The firm's profit under ISS in period t

the demand for transmission services is observable.

However, the drawback of this mechanism comes through the tax/subsidy part, which is usually neither a feasible nor a necessarily desirable policy for regulators. In contrast, V-F uses a price index approach, according to which the price index of a multi-product firm's services should be lowered by last period's (excess) profit margin. In particular, they use a Laspeyres price index.

$$\frac{\sum_{i=1}^M p_{i,t} q_{i,t-1}}{\sum_{i=1}^M p_{i,t-1} q_{i,t-1}} \leq 1 - \frac{\pi_{t-1}}{\sum_{i=1}^M p_{i,t-1} q_{i,t-1}} \quad (3)$$

Here p_t and q_t are prices and quantities of outputs t where $t = 1, \dots, M$. The key idea is that the firm should be able to reduce its price level if it makes (excess) profits. In a stationary environment under a Laspeyres price index the welfare increase in each period is always greater than profit (Vogelsang and Finsinger, 1979). Thus welfare increases as long as the firm makes a positive profit. Total welfare reaches its maximum when the firm is no longer making profit (Ramsey pricing condition). Thus, the prices converge to Ramsey prices over time. Because the firm at least has to break even, this contrasts with L-M and ISS, which lead to marginal cost prices even if they do not cover costs. The main drawbacks are: (1) that V-F has weak cost-reducing incentives, (2) that there is the strategic issue that a firm expecting V-F regulation

may increase prices beyond monopoly prices or may wastefully increase costs prior to regulation (Sappington, 1980), and (3) there is the issue that the mechanism only approaches the Ramsey welfare optimum over time and may never converge if the cost and/or demand functions change. The latter becomes a fundamental problem for a real-world application. Regulation therefore needs to update with changes in inflation, cost, demand, etc.

The main new developments, starting from V-F as the basis, were various forms of price caps (based on Littlechild (1983)), which strengthened the cost-reducing incentives, while allowing for some profits to the firm even in the long run. While under V-F prices on average have to be reduced by last period's profit margin, Littlechild's formula adjusts them by the rate of inflation and by an exogenously determined adjustment factor 'X' that should reflect long-run productivity changes. Using a general inflation index I one gets

$$\frac{\sum_{t=1}^M p_{t,t} q_{t,t-1}}{\sum_{t=1}^M p_{t,t-1} q_{t,t-1}} \leq (1 + I - X) \quad (4)$$

The key new idea (based on Baumol (1982)) is to adjust prices to inflation minus expected productivity increases instead of basing them on last period's profit. Prices under Littlechild (1983) also converge to the Ramsey price structure, however, with $\pi \geq 0$.

Since prices here do not depend on cost, any reduction in cost will generate more profit for the firm. This implies strong incentives for cost reduction. However, because the world changes in unforeseen ways, 'X' has to be revisited from time to time. Such a predetermined adjustment of X every few years creates a potential 'ratcheting problem'. As a result there will be room for strategic behavior and no full cost-reducing incentive.

Price caps were further refined through two-part tariffs meant to improve poor investment incentives under linear price caps (Vogelsang, 1989) and (Vogelsang, 2001). The purpose of Vogelsang (2001) is to provide compatibility between investment and usage of an electricity transmission network. Optimal usage requires congestion pricing, but congestion pricing alone

leads to incentives for the regulated firm to create congestion by withholding investment.

In the context of Vogelsang (2001) investment occurs by a monopoly transmission company (Transco). An independent system operator (ISO) calculates congestions prices and the Transco receives the network merchandising surplus. If investments were sufficient then linear price caps would need to exceed marginal congestion costs by a lot in order to cover total costs (by about 300% according to (Perez-Arriaga et al., 1996)). However, the explicit use of two-part tariffs in wholesale price caps can assure cost coverage and induce balanced network expansion and network utilization. If under the two-part tariff congestion charges p increase (decrease) the fixed fee ϕ decreases (increases).

$$\frac{\sum_{t=1}^M p_{t,t} q_{t,w} + \phi_t N_w}{\sum_{t=1}^M p_{t,t-1} q_{t,w} + \phi_{t-1} N_w} \leq (1 + I - X) \quad (5)$$

Here the index w stands for "weight". In the case of a Laspeyres index the weight is set to q_{t-1} . Setting $I - X = 0$ and assuming the number of consumers N to be given, under Vogelsang (2001) the regulated two-part tariff fixed fee is defined by the constraint $\phi_t \leq \phi_{t-1} + (\mathbf{p}_{t-1} - \mathbf{p}_t) \mathbf{q}_w / N$, where the bold variables are vectors.

While this appears to be only a constraint on the fixed fee, it implicitly defines the range of allowed prices \mathbf{p} . If we set $\mathbf{q}_w = \mathbf{q}_{t-1}$ the weighted price change, $(\mathbf{p}_{t-1} - \mathbf{p}_t) \mathbf{q}_{t-1}$, is a first-order (Slutzky) approximation to consumer surplus change that always underestimates that change. Thus, as long as $\mathbf{p} \geq \mathbf{marginal\ cost}$ and as long as N is given, a price reduction is profitable for the firm and provides a consumer welfare improvement larger than any loss suffered by the firm before applying the fixed fee.

The Vogelsang (2001) mechanism was specifically developed for electricity transmission pricing, where short-term nodal prices would not allow for the cost coverage necessary for transmission expansion investments. Hogan et al. (2010) (in the following: H-R-V) refine Vogelsang (2001), but use the same formula for the two-part tariff. They show the applicability of the

mechanism in a realistic electricity transmission context. In addition, Rosellon with various co-authors has applied and simulated this approach.⁵ They showed that H-R-V generally yielded socially superior outcomes than other applied mechanisms but was still far from perfect. Quantity weights of the underlying price index have played a large role in this context. Averaged Laspeyres/Paasche weights, mimicking a linear demand curve, turned out to dominate most other approaches.

3 Two-part tariffs as a bridge between subsidy-based and constraint-based regulatory incentive mechanisms

As mentioned in Section 2, Gans and King (2000) suggest using the ISS to induce efficient electricity transmission investment. They note that the ISS would provide the firm with a reward (penalty) equal to the social surplus increase (decrease) in each period. They further claim that the ISS would be capable of efficiently alleviating market power by generators. All this is based on the assumption that the relevant information is at the disposal of the regulator. Gans and King (2000) suggest that this information can be readily inferred by the Independent System Operator (ISO) from demand bids and generator bids for short-term generation dispatch. A question that has been raised with this approach is to what extent such short-run bidding behavior can be used as a guide for long-term investments. Because it is based on current expenses rather than on economic costs, a further issue is that the approach could require extra-ordinarily high fixed fees or subsidies when new transmission lines would be added. The remaining drawbacks of the Hogan et al. (2010) and Gans and King (2000) approaches are therefore the use of subsidies/taxes by Gans and King and slow convergence by H-R-V.

The H-R-G-V approach suggested here combines Gans and King (or ISS) with H-R-V and eliminates these drawbacks. Under H-R-G-V, the firm faces a constraint on the fixed fee of two-part tariffs but can implicitly set the market price for network usage by setting network

⁵See, in particular, Rosellon and Weigt (2011) and Schill et al. (2015).

capacity. The constraint on the fixed fee for the current period is given by the fixed fee of the previous period plus the change in consumer surplus over the previous period, where the fixed fee is aggregated over all users. Under this approach the transmission company (Transco) can choose variable prices (or congestion prices: indirectly, via investment), while the revenue from the fixed fee at time t is set as $\Phi_t = \Phi_{t-1} + \Delta V$ or $\phi_t N_t = \phi_{t-1} N_{t-1} + \Delta V$, where ϕ is the fixed fee paid by individual users and N is the number of users. Under the assumption that N is constant we can write $\phi_t = \phi_{t-1} + \frac{\Delta V}{N}$.⁶

H-R-G-V marks a change from the Slutsky approximation $(\mathbf{p}_{t-1} - \mathbf{p}_t)\mathbf{q}_{t-1}$ that was the basis of H-R-V to the actual consumer surplus change that was the basis of ISS. As a result the firm receives as total revenue the market revenue plus the change in consumer surplus between the starting period '0' and the current period 't'. It has to pay its costs so that its overall profit equals market profit plus the portion of consumer surplus gained (or lost) over the base period. To see the relationship between the ISS, H-R-V (or Vogelsang (2001)) and H-R-G-V, we can look at the net profit (after subsidy or fixed fee) that the firm receives in period t . In the following π represents profits from variable fees, while capital Π is total profit after applying the mechanism. For the ISS we have:

$$\Pi_t^{ISS} = ISS_t + \pi_t = \Delta V_t + \Delta \pi_t = \Delta W_t \quad (6)$$

For H-R-V, assuming a binding H-R-V constraint and $I-X = 0$, it follows that

$$\Pi_t^{H-R-V} = \pi_t + N_t \left[\frac{(p_{t-1} - p_t)q_{t-1}}{N} + \phi_{t-1} \right] \quad (7)$$

Assuming further that N is given, it follows that

$$\Pi_t^{H-R-V} = \Pi_{t-1} + \Delta \pi_t + (p_{t-1} - p_t)q_{t-1} \leq \Pi_{t-1} + \Delta W_t \quad (8)$$

⁶The assumption that N is constant is not innocuous. If the firm reduces the price and therefore increases the fixed fee consumers with low valuations of the service may be excluded although they would pay a lower price p . This would lead to a deadweight loss problem that we will consider below in the conclusions.

Also for H-R-G-V when N is given, it follows that

$$\Pi_t^{H-R-G-V} = \Pi_{t-1} + \Delta\pi_t + \Delta V_t = \Pi_{t-1} + \Delta W_t \quad (9)$$

Thus, H-R-G-V differs from ISS by adding last period's total profit to the firm's total ISS reward and differs from H-R-V by replacing the Slutsky approximation with the accurate consumer surplus change. The H-R-G-V mechanism therefore represents a blend of the two types of mechanisms taking from ISS the consumer surplus approach and from H-R-V the price-cap approach.

4 Main theoretical properties of the H-R-G-V mechanism

We will now analyze some of the properties of H-R-G-V relative to the ISS and H-R-V and bring out, why, from a practical perspective, H-R-G-V may dominate the others. Our evaluation of the mechanisms will be guided by some general objectives. These are: (1) static price efficiency and optimal capacity utilization, (2) static production efficiency, and least-cost operation, and (3) optimal capacity investment. The regulated firm should: (4) neither depend on subsidies nor should it earn excessive returns. There should: (5) be little room for strategic behavior. Finally, the mechanism should: (6) be verifiable, the latter in the sense that its application should depend on observable and verifiable data so that its application could, in principle, be defended in court.

(1) *Static price efficiency*: In the context of myopic profit maximization the firm's total profit under the H-R-G-V mechanism can be written as $\Pi_t^{H-R-G-V} = \pi_t + V_t - V_{t-1} + \Phi_{t-1}$. Thus, the two leftmost right-hand side terms $\pi_t + V_t$ that are relevant for myopic profit maximization amount to total surplus, because the other terms are not affected by current period maximization. This shows the relationship to L-M, where the firm also maximizes total surplus, resulting in marginal cost prices. Because the firm has to pay 100% of its costs, it also is incentivized to minimize costs for given output. Since H-R-G-V is a dynamic adjustment process, however, the issue of potential strategic behavior of the firm is highly relevant. It can be shown that, just like

the case of ISS, from period '1' onward the firm will not use strategic behavior under H-R-G-V.

In order to simplify mathematical proofs for H-R-G-V, we will first look at the two-part tariff version of the ISS with a fixed fee $\phi_t = \Delta V/N$ where N is given. Under this form of ISS the firm will $\max_{\{p_t\}_{t=1}^{\infty}} \sum_{t=1}^{\infty} (\Delta V_t + \Delta \pi_t)(1+r)^{-t}$ where r is the applicable discount rate. Since this is identical to the original ISS with the exception that the regulator observes the firm's cost rather than expenses, the Proposition 1(i) of Sappington and Sibley (1988) applies. Intuitively, summed over all periods and without discounting the firm could in aggregate earn at most ΔW^* , which is the maximal change in total surplus compared to the beginning surplus in period '0' or the area ACB in Figure 1 above: $S_{ACB} = \Delta W^* = W_{max} - W_0$. Any discounting over time will make the profit less than S_{ACB} . Therefore, the firm will make an immediate adjustment to lower price to marginal cost so that $p_0 > MC$, $p_1 = p_2 = \dots = p_{\infty} = MC$.

In contrast to the ISS, the lack of strategic behavior under H-R-G-V follows from the fact that the firm is in any period $t > 0$ rewarded with the period 't-1' profit plus the increase in social surplus in period 't' and that it can, by repeating the period 't' behavior, perpetuate this profit. Under H-R-G-V, without discounting, for any sequence of periods $(0, \dots, T)$ the firm's cumulative profit summed over all periods is $\pi_0 + \sum_{t=1}^T (\Delta W_0^t)$ where $\Delta W_0^t = W_t - W_0$.

This sum is maximized if $\Delta W_0^t = \Delta W^*$ for all t . Thus, it is best for the firm to immediately maximize social surplus and then continue to do so by its subsequent decisions. Contrary to the case of the ISS, not even a discounting argument needs to be made. The myopic maximization would strictly dominate any strategic considerations even at a zero discount rate.

(2) *Static production efficiency and cost-reducing incentives*: Sappington and Sibley (1988) establish static production efficiency for the ISS in their Proposition 1(iii). Their proof also holds for H-R-G-V. However, cost-reducing incentives under the ISS are socially sub-optimal (Blackmon (1992) and Blackmon (1994)). In contrast, since the firm under H-R-G-V always has to pay its full costs, it has full cost-reducing incentives, stronger than under the ISS. Also, contrary to the ISS, the regulator does not even have to check the firm's expenses, only consumer surplus needs to be checked.

(3) *Optimal investment*: Just like for Gans and King (2000) a question with H-R-G-V is to what extent the short-run bidding behavior of demand determination can be used as a guide for long-term investments. Rewarding the regulated firm ex post by using the consumer surplus based on actual demands will induce the firm's managers to do everything to learn about all factors influencing demand and to use their best predictions. The firm has no incentive to misrepresent its best demand predictions. Ex post, if the prediction was wrong the firm will make less profit than it would have under the correct prediction. However, users may strategically shade their demands in auctions to avoid high fixed fees (Henze, 2016). That is the reason we have used a uniform fixed fee for all users and not a discriminatory fixed fee. If the number of users is large enough, demand shading should not be a problem. Also, discrete bids may not fully describe the shape of the demand curve. Furthermore, there is some moral hazard possibility for the regulated firm from the X-factor adjustments, if any, because those adjustment factors have to be set by the regulator.

(4) *No subsidies, no excess profits*: While Proposition 1(ii) of Sappington and Sibley (1988) makes sure that under the two-part tariff version of the ISS the firm will not earn excessive profits, the firm under H-R-G-V will generally earn higher profits than under the ISS and it can even earn higher profits than under profit-maximizing linear prices. Thus, the regulator would have to make sure before the process starts that the firm earns no excess profits at that time. As suggested by Sharkey (1979), the regulator may want to ask the firm for an entry fee to be able to supply under H-R-G-V (although the net gain from average-cost prices to marginal cost prices may not be too large).

(5) *No room for strategic behavior*: As shown above for static price and production efficiency, the regulated firm has incentives to behave efficiently in all periods after the H-R-G-V mechanism has been introduced. In that sense there will be no strategic behavior. The firm will, however, use strategic pricing behavior during the period before the regulation is introduced, provided it knows about the introduction and can still increase its prices at that time. This strategic incentive holds equally for ISS and H-R-G-V. In this case the Lerner Index, and therefore the

price in period 0, will be larger than under simple myopic profit maximization. Note that this argument on strategic behavior does not hold after the process has started because then the firm always suffers the consumer surplus loss from any price increase.

(6) *Verifiability*: While H-R-V is based only on verifiable data, both the ISS and H-R-G-V require the observation of part of the demand curve.

Summing up, the main differences in performance between H-R-G-V, H-R-V and the two-part tariff version of ISS is that H-R-G-V allows the firm to earn higher profits but that it provides strong incentives for cost reduction and innovation (compared to the ISS) and converges immediately to optimal pricing (compared to H-R-V). Its only significant drawbacks over an ideal scheme are (1) some danger of excess profits and (2) the necessity to observe the relevant portion of the demand curve. The latter requirement may make electricity transmission the preferred application of the H-R-G-V mechanism as claimed by Gans and King (2000).

5 Detailed application of the H-R-G-V mechanism to electricity transmission services

5.1 Adaptation of the necessary context

The modeling approach so far has been entirely stylized and generic and therefore could have been applied to any regulated monopoly. The real world is substantially more complex, something that is particularly true for the electricity transmission industry. In the following, we therefore adapt the context towards a real-world application to electricity transmission.

Contrary to the modeling approach in Sections 2 to 4, electricity transmission service is an intermediate input with a derived demand rather than an end-user demand. Electricity transmission links electricity generators with distribution companies/loads. In order to continue with the consumer surplus approach we therefore assume that both these production stages are provided competitively.⁷ As a result of this assumption, the relevant gross consumer surplus is the area

⁷Alternatively, loads could be regulated at competitive prices.

between the end-user demand and the sum of loads' and generators' marginal cost curves. Under perfect competition (and neglecting generators' fixed costs) this is equal to the sum of the generators' profits and the loads' willingness-to-pay (WtP) for electricity.⁸ In order to arrive at the net consumer surplus relevant for the fixed fee of the H-R-G-V mechanism, we have to deduct the variable fees (which is the merchandising surplus) paid by the loads for transmission services.

A second complication of the electricity transmission industry is that the price setting is not usually done by the Transco as the transmission supplier but by an independent system operator (ISO) who collects the supply information of the generators and the demand information of the loads and sets prices by maximizing the difference between the loads' willingness to pay and the generators' supply bids (which are marginal costs). Thus, the ISO finds the static welfare maximizing prices (for given capacities). As a result in our model, the Transco only makes the investment decisions that automatically imply the prices and via the regulatory constraint also the fixed fees.

A third complication is that transmission networks have a very complicated cost function. In particular, as will become clear below, several constraints have to be added to the regulated firm's and the social planner's maximization problem.

In this context, one has to keep in mind (1) that although electricity transmission resembles transportation services, it is subject to specific physical laws, Kirchhoff's law, in particular, (2) that after capacity investment has been done, there are virtually no direct transport cost (except for line losses that we neglect), but there are opportunity costs via congestion of lines. These opportunity costs are shadow-valued via the nodal price differences of electricity at different nodes. If there is no line congestion, arbitrage eliminates the nodal price differences.

⁸The relevant gross consumer surplus concept for electricity transmission services is the end-users' willingness to pay $WtP^{end-users}$ for electricity minus the cost of all electricity services but transmission. Assume that these costs consist of generation costs $C^{generators}$ and distribution costs C^{loads} . If generation and distribution are competitive industries with zero profits and we define their revenues as $R^{generators}$ and R^{loads} and if loads buy the electricity from generators and sell it to end-users, we have $WtP^{loads} = WtP^{end-users} - R^{generators} - C^{loads}$ and $\pi^{generators} = R^{generators} - C^{generators}$. Hence, $WtP^{loads} + \pi^{generators} = WtP^{end-users} - R^{generators} - C^{loads} + R^{generators} - C^{generators} = WtP^{end-users} - C^{loads} - C^{generators}$.

5.2 The welfare-maximizing investment of the profit-maximizing Transco under H-R-G-V

We propose an independent regional Transco owning the transmission network. The Transco does the transmission investment, bears the investment costs and collects the revenues from its investments over periods t ($t = 1, \dots, T$). The Transco revenue in each period has two parts. The first part is the network merchandising surplus and the second part is a fixed fee.⁹ The fixed fee is a charge to loads to fund the transmission expansion costs. The Transco is subject to a proposed regulatory constraint. The proposed regulatory constraint is defined on the fixed-fee part of the Transco revenue. Under H-R-G-V, this constraint in period t has the form:

$$\Phi_t = \Phi_{t-1} + \Delta LS_t + \Delta GS_t \quad (10)$$

where Φ_t is the fixed fee in planning period t , $\Delta LS_t = LS_t - LS_{t-1}$ is the change in the load surplus from planning period $t - 1$ to t and $\Delta GS_t = GS_t - GS_{t-1}$ is the change in the generator surplus from planning period $t - 1$ to t . The interaction between a profit-maximizing Transco and the regulator can then be formulated as the bilevel optimization problem (11), where the upper-level problem (11a)-(11c) concerns the Transco's optimization and the lower-level problem (11d) the ISO's optimization.

⁹Perez-Arriaga et al. (1995) show that the merchandising surplus recovers only up to approximately 30% of total cost of transmission grid based on the data from Argentina, Central America, Chile, Spain and England and Wales. Accordingly, Rubio-Oderiz and Perez-Arriaga (2000) propose that the revenue from merchandising surplus be complemented with a fixed fee to recover the remaining costs. The second part of the Transco revenue plays the role of this additional complementary charge.

$$\underset{\Phi_t, \mathbf{K}_t}{\text{Maximize}} \left(\sum_t MS(x_t, \rho_t) + \Phi_t - IC(\mathbf{K}_t) \right) \quad (11a)$$

Subject to

$$\Phi_1 = 0 \quad (11b)$$

$$\Phi_t = \Phi_{t-1} + \Delta LS_t + \Delta GS_t \quad \forall t \geq 2 \quad (11c)$$

$$\underset{x_t \in X_t(\mathbf{K}_t) : \rho_t}{\text{Maximize}} \sum_t (U_t - C_t) \quad (11d)$$

In the above equations x_t is the vector of dispatch quantities, $X_t(K_t)$ is the feasible set depending on capacity K_t , decided in the upper-level problem, and ρ_t is the vector of shadow prices. U_t is a concave utility function of loads and C_t is a convex cost function of generators. Assuming (without loss of generality) that no investment happens in period 1, the fixed fee in this period can be set to zero. $MS(x_t, \rho_t)$ is the merchandising surplus in period t and x_t and ρ_t are calculated by the ISO. $IC(K_t)$ is the cost of investing capacity of K_t in the transmission system. K_t is the vector of transmission investments at different locations. Optimization problem (11d) models the ISO maximizing gross consumer surplus.¹⁰

Objective function (11a) and regulatory constraints (11b)-(11c) show how the merchandising surplus and fixed fee can be traded off against each other. Transmission investment generally reduces the merchandising surplus of the network and would then decrease the profit of the Transco. Using the regulatory constraints (11b)-(11c), the Transco can counter the decrease of merchandising surplus by increasing its fixed fees, because the investment lowers the shadow prices of transmission capacities and increases the difference between load willingness-to-pay and generators' marginal costs. Similarly, if the Transco does not invest in the transmission network, the merchandising surplus will generally increase (due to higher levels of congestion in the system). However, the Transco needs to reduce its fixed fee to meet the H-R-G-V constraint

¹⁰Here we use $U - C$, which economically speaking is equal to the $MS + LS + GS$ used before. However, mathematically speaking, for calculating surpluses we need locational marginal price (LMP), and the LMP is the Lagrange multiplier which does not exist explicitly in the ISO list of primal variables.

(11b)-(11c). By this mechanism, the Transco is incentivized to expand the transmission system even if the merchandising surplus is decreased. As long as the Transco can compensate the decrease in merchandising surplus and its investment cost by higher fixed fees, we can expect to observe expansion. On the other hand, the fixed fee changes at the same time with the change in load and generator surplus. Accordingly, the regulatory model intertemporally incentivizes transmission investments that relieve congestion from the network, increase load and generator surplus, improve the Transco profit and lower the market prices.

It is straightforward to show that for any sequence of periods $t = 1, \dots, T$, the profit-maximizing Transco in (11) invests in transmission network capacity such that the social welfare is maximized. Substituting Φ_t in the objective function (11a) leads to:

$$(11) \equiv \underset{K_t \geq 2}{\text{Maximize}} \left(MS_1 + \sum_{t \geq 2} MS_t + GS_t + LS_t + \Phi_{t-1} - GS_{t-1} - LS_{t-1} - IC_t \right) \quad (12a)$$

Subject to

$$\underset{x_t \in X_t(K_t) : \rho_t}{\text{Maximize}} \sum_t (U_t - C_t) \quad (12b)$$

The maximand in (12) can be rewritten as:

$$\begin{aligned} (12a) &= MS_1 + (MS_2 + GS_2 + LS_2 + \overbrace{\Phi_1 - GS_1 - LS_1}^{-(GS_1+LS_1)} - IC_2) + \\ &(MS_3 + GS_3 + LS_3 + \underbrace{\Phi_2 - GS_2 - LS_2}_{-(GS_1+LS_1)} - IC_3) + (MS_4 + GS_4 + LS_4 + \underbrace{\Phi_3 - GS_3 - LS_3}_{-(GS_1+LS_1)} - IC_4) + \dots \\ &= MS_1 - (T-1)(GS_1 + LS_1) + \sum_{t \geq 2} MS_t + GS_t + LS_t - IC_t \end{aligned} \quad (13)$$

Then optimization problem (12) can be written as:

$$(12) \equiv \underset{K_t \geq 2}{\text{Maximize}} \left(MS_1 - (T-1)(GS_1 + LS_1) + \sum_{t \geq 2} \overbrace{MS_t + GS_t + LS_t - IC_t}^{W_t} \right) \quad (14a)$$

Subject to

$$\underset{x_t \in X_t(K_t) : \rho_t}{\text{Maximize}} \sum_t (U_t - C_t) \quad (14b)$$

Since the sum of the terms after the summation sign represents W_t , the optimization problem (14) clearly shows that from period two onwards the Transco invests in K_t such that the total surplus ($MS_t + GS_t + LS_t$) is maximized. In other words, the Transco maximizes the social welfare by its transmission investment decisions.

5.3 The H-R-G-V mechanism for a two-node system: An analytical discussion

We now employ a two-node example system to illustrate the properties discussed above. Such a network considered in Figure 2 involves one link. There is one generator located at node 1 and one load located at node 2. The inverse demand curve is linear and downward sloping. The generator at node 1 can produce more than enough power at a constant marginal cost, which for simplicity and without loss of generality is assumed to be zero. We consider two periods $t-1$ (status quo capacity) and t (expanded capacity). Period $t-1$ provides necessary information for setting up the regulatory constraint and the H-R-G-V regulation is applied to period t . The power flowing through the transmission line is f_t and the nodal price difference across the line is μ_t . We assume that the transmission line constraint is binding (capacity of the line is equal to f_t) and the investment cost $C(f_t)$ is continuous and differentiable. Under welfare maximization, the expanded capacity f_t meets the condition that $\mu_t = \frac{\partial C}{\partial f_t}$.¹¹ On the other hand, the profit-

¹¹For a given network configuration, the welfare optimal capacity is where the expected marginal value of the constraint on the network flow limit is equal to the marginal cost of adding capacity (Biggar and Hesamzadeh, 2004).

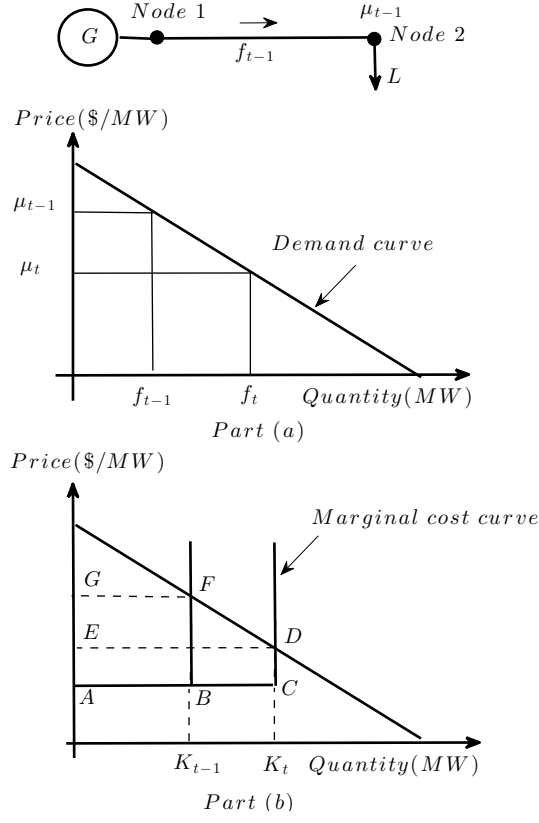


Figure 2: The analytical two-node system

maximization problem of the Transco in period t can be written as:

$$\underset{\Phi_t, f_t}{\text{Maximize}} \quad (\mu_t f_t + \Phi_t - C(f_t)) \quad (15a)$$

Subject to

$$\Phi_t = \Phi_{t-1} + \Delta LS_t : (\rho_t) \quad (15b)$$

In (15a), $\mu_t f_t$ models the congestion rent of the line between nodes 1 and 2 in period t , Φ_t is the fixed fee in period t and $C(f_t)$ is the investment cost. Constraint (15b) is the H-R-G-V regulatory constraint and ρ_t is the Lagrange multiplier associated with this constraint. Based on this constraint, the fixed fee in period t is equal to the sum of the fixed fee in the previous period

$t - 1$ and the change in the load surplus in period t . Note that in this analysis the generator is left out because its marginal cost is zero. To derive the Lagrangian associated with the optimization problem (15), the Lagrange multiplier ρ_t is introduced.

From Figure 2 part (a), $\Delta LS_t = \frac{1}{2}(\mu_{t-1} - \mu_t)(f_{t-1} + f_t)$. The Lagrangian associated with (15) is $\mathcal{L} = \mu_t f_t + \Phi_t - C(f_t) + \rho_t(\Phi_{t-1} + \frac{1}{2}(\mu_{t-1} - \mu_t)(f_{t-1} + f_t) - \Phi_t)$. We have $\frac{\partial \mathcal{L}}{\partial \mu_t} = \frac{1}{2}f_t + \frac{1}{2}\mu_t \frac{\partial f_t}{\partial \mu_t} + \frac{1}{2}\mu_{t-1} \frac{\partial f_t}{\partial \mu_t} - \frac{1}{2}f_{t-1} - \frac{\partial C}{\partial f_t} \frac{\partial f_t}{\partial \mu_t} = 0$. This can be written as:

$$-(\mu_t - \frac{\partial C}{\partial f_t}) \frac{\partial f_t}{\partial \mu_t} = f_t - f_{t-1} + \frac{\partial f_t}{\partial \mu_t} (\mu_{t-1} - \frac{\partial C}{\partial f_t}) \quad (16)$$

Because of the linear demand curve, equality (16) implies $\mu_t = \frac{\partial C}{\partial f_t}$ independent of μ_{t-1} . But this is the condition of welfare optimal capacity for the two-node system. In other words, independent of f_{t-1} and μ_{t-1} , the regulated Transco invests in welfare-maximum capacity in period t . For the case of generation-transmission investment, the arguments are quite intuitive. Assume that the status-quo generation and transmission capacity is K_{t-1} and the social-welfare maximizing capacity is K_t (see Fig. 2 part b). The Transco revenue in period $t - 1$ is area $ABFG$ (network merchandising surplus). In period t , the merchandising-surplus revenue is area $ACDE$ and the fixed-fee revenue (resulting from load-surplus change) is area $EDFG$. Accordingly, the total revenue of Transco in period t is area $ACDFG$. The change in the revenue of Transco is area $BCDF$ which is exactly the change in social welfare resulting from optimal generation-transmission investment.

The loop-flow nature of power in meshed networks does not allow us to derive an analytical solution for general settings. However, in the next two sections we show in numerical simulations that the proposed H-R-G-V mechanism results in welfare-maximizing investment in the electricity transmission network.

6 A more general application of the H-R-G-V mechanism to electricity transmission pricing and investment

For the more general case of meshed networks the profit-maximizing Transco is modeled via the bilevel program shown in (11) above. This nonconvex optimization problem is hard to solve for realistically-sized networks. A conceptual presentation of the optimization problem (11) is set out below.

$$\begin{aligned} & \text{Maximize } \textit{Total congestion rent} + \textit{Fixed fee} \\ & \quad - \textit{Total transmission investment cost} \end{aligned} \tag{17a}$$

Subject to:

$$(A) \textit{ Transmission investment constraints} \tag{17b}$$

$$(B) \textit{ H-R-G-V regulatory constraint} \tag{17c}$$

$$(C) \textit{ The ISO optimization problem} \tag{17d}$$

For comparison purposes, the following two types of Transcos are also modeled and simulated.

1. *Transco without regulation:* In this case, the Transco is not regulated with respect to its transmission investment decisions but restricted to the nodal prices that emerge from the ISO's maximization problem. This case can be modeled by removing the regulatory constraints (11b)-(11c) from optimization problem (11) and setting $\Phi_t = 0$. We assume that in the no-regulation case, the cost of transmission investment has to be fully recovered by the network merchandising surplus.
2. *Transco with cost-plus regulation:* In this case, the Transco receives not only the merchandising surplus but it can charge an extra fixed fee based on its cost of transmission capacity

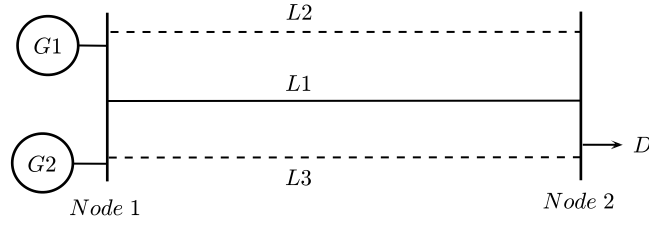


Figure 3: The single line diagram of the 2-node system

expansion. This case can be modeled by replacing the regulatory constraint (11b)-(11c) by $\Phi_t = \Phi_{t-1} + (1+r) \sum_m \bar{C}_m (z_{m,t} - z_{m,t-1})$ where $r \in \mathbb{R}_+$ is the cost mark-up set by the regulator.

The economic theory behind no-regulation and cost-plus regulation are studied in detail in Schill et al. (2015) and Joskow (2006).

7 Numerical example: The 2-node example system

To show the operation of the mechanism, the optimization problem described in (11) is applied to a 2-node system example. Four periods are considered. Period 1 models the status-quo case. The 2-node example has two nodes, one existing line connecting these two nodes and two candidate transmission lines for investment planning. The single-line diagrams for the 2-node system is shown in Figure 3.

In considered case studies the period length ψ is taken as 500 hours and the discount rate is 0.05. The rest of system data is reported in Tables 1 to 3 for the 2-node system. The maximum load and maximum value of load at each node is increasing each planning period by 10%, an example of change of demand function is illustrated in Figure 4. Available generation capacity remains the same at each investment planning period.¹²

Different regulatory regimes lead to different transmission expansion strategies. The H-R-

¹²In Table 2 reactance is a form of opposition that electric components exhibit to the passage of alternating current because of capacitance or inductance.

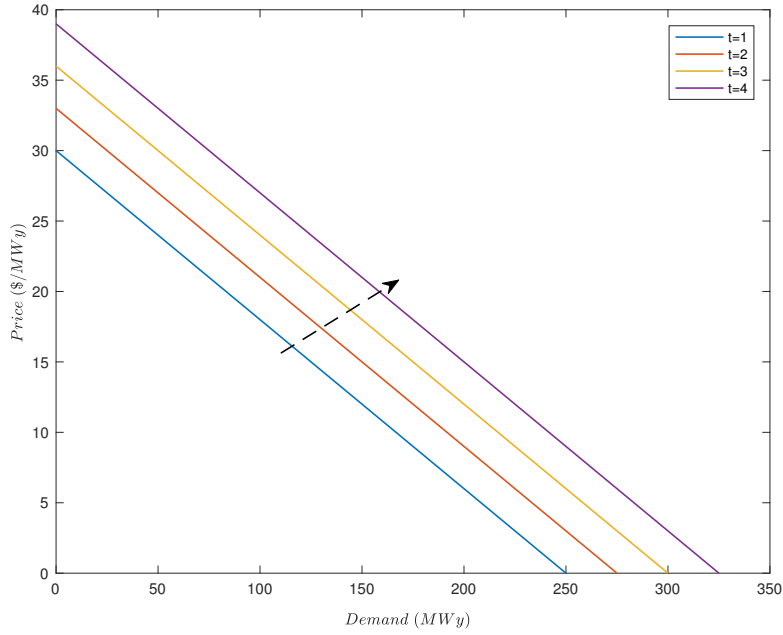


Figure 4: Example of change of demand function between each planning year

Table 1: Data of generators in the 2-node system

| Generator | Node | Short-run marginal cost (\$/MWh) | Capacity (MW) |
|-----------|------|----------------------------------|---------------|
| G1 | 1 | 10 | 50 |
| G2 | 1 | 30 | 200 |

G-V mechanism incentivizes the Transco to invest in the transmission network such that social welfare is maximized via the profit maximization of the Transco.

The results of transmission investment under H-R-G-V for the 2-node system are reported in Table 4, where $t = 1$ refers to the status quo and is equal for all three regimes. Since investments are assumed to be binary, investments only show up as 0 or 1. The results under the H-R-G-V mechanism were compared to the results of the transmission investment under cost-plus regulation and under no regulation. Tables 5 and 6 present the results for the 2-node system under cost-plus regulation and under no regulation, respectively. It should be noted that the load surplus is the surplus before the deduction of the fixed fees and social welfare is calculated as

Table 2: Data of transmission lines in the 2-node system, Cct: Circuit

| Line | Reactance (p.u.) | Capacity | Transmission Investment Cost (\$/Cct.) |
|------|------------------|----------|--|
| L1 | 0.4 | 100 | - |
| L2 | 0.2 | 100 | 38 000 |
| L3 | 0.48 | 100 | 40 000 |

Table 3: Data of loads in the 2-node system

| Load | Maximum price (\$/MWh) at $t = 1$ | Maximum Capacity (MW) at $t = 1$ |
|------|-----------------------------------|----------------------------------|
| D1 | 30 | 250 |

the aggregate discounted sum over the whole planning period.

For the 2-node system, under the H-R-G-V, the Transco invests in both lines $L2$ and $L3$ in period 2. This is exactly the social-welfare maximizing investment. However, under the no regulation case, the Transco just invests in line $L2$ in order to maximize the congestion rent in the network. This is obviously not the welfare maximizing investment. Under the cost-plus regulation, the Transco invests in the most expensive line $L3$ which is understandable given the regulatory constraint of the cost-plus regulation mechanism.

8 Conclusions and further thoughts

The proposed H-R-G-V is a regulatory incentive mechanism which induces efficient investment in electricity transmission networks. It is non-Bayesian, practical and partially based on observable and verifiable data.

In its application to electricity transmission, the mechanism is mathematically modeled as a bilevel program, which links the Transco, the regulator and the ISO. The proposed mathematical formulation considers transmission network topology change in transmission investment decisions. For the simulations, the nonconvex bilevel program modeling of the H-R-G-V mechanism is reformulated as a mixed-integer linear program (MILP). This MILP model is solved to ε -global optimality using a commercially available branch-and-cut solver.

To illustrate the operation of the mechanism, the 2-node example system is analyzed and

Table 4: Investment results under the H-R-G-V regulatory mechanism in the 2-node system.

Tran. Inv.: Transmission Investment

| | $t = 1$ | $t = 2$ | $t = 3$ | $t = 4$ |
|---|---------|-----------------|---------|---------|
| (1,2):L3 | 0 | 1 | 1 | 1 |
| (1,2):L2 | 0 | 1 | 1 | 1 |
| Fixed Fee (M\$) | 0 | 17.49 | 21.94 | 26.57 |
| Generator Surplus (M\$) | 0.5 | 4.11 | 7.94 | 11.97 |
| Load Surplus (M\$) | 2.5 | 16.38 | 17.01 | 17.61 |
| Congestion Rent (M\$) | 7.5 | 0 | 0 | 0 |
| Tran. Inv. Cost (\$) (considering present value) | 0 | 78 000 (74 286) | 0 | 0 |
| Tran. Profit (M\$) | 7.5 | 17.42 | 21.94 | 26.58 |
| Social Welfare (M\$) | 10.50 | 20.42 | 24.95 | 29.58 |

Table 5: Investment results under no regulatory mechanism in the 2-node system. Tran. Inv.:

Transmission Investment

| | $t = 1$ | $t = 2$ | $t = 3$ | $t = 4$ |
|---|---------|---------|----------------|---------|
| (1,2):L3 | 0 | 0 | 0 | 0 |
| (1,2):L2 | 0 | 0 | 1 | 1 |
| Fixed Fee (M\$) | - | - | - | - |
| Generator Surplus (M\$) | 0.5 | 0.48 | 0.45 | 0.43 |
| Load Surplus (M\$) | 2.5 | 2.62 | 10.88 | 11.27 |
| Congestion Rent (M\$) | 7.5 | 9.32 | 11.43 | 14.86 |
| Tran. Inv. Cost (\$) (considering present value) | 0 | 0 | 38 000 (34467) | 0 |
| Tran. Profit (M\$) | 7.5 | 9.32 | 11.40 | 14.86 |
| Social Welfare (M\$) | 10.50 | 12.42 | 22.73 | 26.56 |

Table 6: Investment results under cost-plus regulatory mechanism in the 2-node system, with

$r = 0.3$. Tran. Inv.: Transmission Investment

| | $t = 1$ | $t = 2$ | $t = 3$ | $t = 4$ |
|---|---------|---------|-----------------|---------|
| (1,2):L3 | 0 | 0 | 1 | 1 |
| (1,2):L2 | 0 | 0 | 0 | 0 |
| Fixed Fee (k\$) | - | - | 47.17 | 47.17 |
| Generator Surplus (M\$) | 0.5 | 0.48 | 0.45 | 0.43 |
| Load Surplus (M\$) | 2.5 | 2.62 | 10.88 | 11.27 |
| Congestion Rent (M\$) | 7.5 | 9.32 | 11.43 | 14.86 |
| Tran. Inv. Cost (\$) (considering present value) | 0 | 0 | 40 000 (36 281) | 0 |
| Tran. Profit (M\$) | 7.5 | 9.32 | 11.44 | 14.91 |
| Social Welfare (M\$) | 10.50 | 12.42 | 22.72 | 26.56 |

simulated when transmission investment is driven by demand growth. While the mechanism performed as expected in the simulations, some limitations and extensions deserve a discussion.

When introducing two-part tariffs in Section 3 above, a strong assumption was made, namely that the number of users N is given and is not affected by the pricing decisions of the regulated firm. It is well known that, in the presence of heterogeneous users, the number of users will be affected by changes in the variable and fixed fees. Specifically, if the firm reduces its variable prices it will be rewarded by higher fixed fees. If all consumers were homogeneous (with the same demand curves) under H-R-G-V, they would receive the same surplus after as before the price change. However, with heterogeneous users some will gain surplus, while others will lose surplus. Thus, some users could drop out. Under the H-R-G-V mechanism, the regulated firm will, in this case, have to reduce the fixed fee for all users, because of the lost surplus from the drop-outs. Thus, the Transco will internalize the welfare effects of the drop-outs in its investment decisions. In future work, it may be worth constructing numerical examples for equilibrium outcomes considering the case of including the possibility of fixed fees differentiated by user groups. In contrast, if the regulated firm increases its variable prices negative fixed fees could result. In this case, consumers may subscribe to the service without any substantial usage, just to collect the negative fixed fee. Thus, negative fixed fees may be a reason for regulators to restrict consumers from entering.

While the ISS and H-R-G-V mechanisms yield optimal prices for any period after the first, this property and efficient investment have formally been shown only in a stationary environment and, in simulations, only for small demand changes. However, for example, if costs were to increase, the firm would, by increasing prices, suffer a fixed fee reduction equal to the consumer surplus loss. Or, if demand shifted outward substantially, the firm could receive a large windfall profit from the consumer surplus increase even if there was no price reduction. This certainly affects the firm's rents, but does it make the firm deviate from optimal behavior, assuming the firm anticipates the demand or cost change? We have so far unsuccessfully tried to create examples of cost and/or demand changes that would lead to strategic deviations of the firm from optimal

adaptation to the new situation. However, the effects on profits could be enormous. For example, if the demand curve shifts outward the firm would, without a price change, be entitled to the whole new consumer surplus area generated by the shift. This, among others, is a reason, why the H-R-G-V mechanism requires some RPI-X type formula of dynamic adjustment (Littlechild, 1983).¹³ Here the inflation ('RPI') and productivity ('X') adjustment itself can take care of the cost changes, but demand shifts are harder to capture.¹⁴ If the regulator can observe a demand shift then he/she could reduce fixed fees by the amount of the demand shift on consumer surplus (at before shift prices) and let the firm capture the full benefits from the demand shift and the price change on the new demand curve only.¹⁵

So far we have assumed certainty and full information. We conjecture that uncertainty with asymmetric information can be taken care of by the mechanism as it is. This would clearly hold if the Transco were fully informed about costs and demands, while the regulator were not. The Transco receives the fixed fee based on the actual realized consumer surplus, and therefore has every incentive to make the right investment decision. This is similar to an investing firm in an unregulated setting. It receives whatever the market price is at the time. If the Transco were not fully informed but better informed than the regulator, those incentives would still persist. It would be desirable, though, to confirm this through further analysis.

Another potential complication could come from lumpy capacity investments in a growing market. Specifically, lumpy capacity investment means that (under price rationing) there will be periods with excess capacity and periods where capacity is strained. Gans and King (2002) argue that the ISS will efficiently cope with lumpy investments and we conjecture that the same holds for H-R-G-V. In our view, adaptations to changing environments should be no more difficult

¹³Because a multiplicative adjustment of the H-R-G-V constraint can distort investment incentives the RPI-X formula may have to be applied additively. Alternatively, profit sharing with the users of the transmission system in the form of reduced fixed fees can provide undistorted investment incentives, as long as the Transco generates enough profits to finance its investments.

¹⁴For an analysis and survey covering the determination of the X-factor see Bernstein and Sappington (1999).

¹⁵This assumes that the demand shift itself is not the result of the regulated firm's marketing and investment policy. For example, if the firm invests in a new link it may be entitled to the change in consumer surplus from demand created by this new line.

for H-R-G-V than for ISS and easier than for H-R-V. Thus, overall the mechanism improves on both. Whether this judgment extends to other network industries than electricity transmission depends on the ability to observe their demands.

Effects of transmission investments on electricity reliability and on increased competition between electricity generators may have to be dealt with as separate services provided under the price cap.

The mechanism is very specific about incentives of a Transco and treats other stakeholder incentives as a 'black box'. We have not specifically considered the effect of the mechanism on optimal generation and distribution grid investment. In our view, H-R-G-V will efficiently address those, provided generators, distribution grids and load-serving entities do not have distorted incentives and externalities are taken care of through separate instruments.¹⁶ In practice, the Transco will have to perform the function of a platform bringing together generators and loads. In this context, studying the impact of the mechanism on strategic behavior of generators and loads would be a good extension of the current work.

Acknowledgement

The authors would like to thank Dina Khastieva for her helps in numerical results of this paper.

Juan Rosell acknowledges support from project no. 232743 from the Sener-Conacyt Fondo de Sustentabilidad Energética. Rosell finished the research contained in this paper while he was a Research Fellow at DIW Berlin, the holder of the Mexico Studies Chair at the USC Sol Price School of Public Policy, at the University of Southern California in Sacramento, California, as well as a Nonresident Fellow of the Center for Energy Studies at the Baker Institute for Public Policy, Rice University, Houston, Texas.

¹⁶Note that under market power in generation, the prices used for calculating the merchandising surplus will be distorted.

References

- Baron, D. P. and Myerson, R. B. (1982). Regulating a monopolist with unknown costs. *Econometrica*, 50(4):911–930.
- Baumol, W. (1982). Productivity-incentive clauses and rate adjustment for inflation. *Public Util. Fortn.; (United States)*, 110:2.
- Biggar, D. R. and Hesamzadeh, M. R. (August 2004). *The Economics of Electricity Markets*. Wiley-IEEE, 1st edition.
- Blackmon, B. G. (1992). The incremental surplus subsidy and rate-of-return regulation. *Journal of Regulatory Economics*, 4(2):187–196.
- Blackmon, G. (1994). *Incentive Regulation and the Regulation of Incentives*. Springer US, 1st edition.
- Fraser, R. (1995). The relationship between the costs and prices of a multi-product monopoly: The role of price-cap regulation. *Journal of Regulatory Economics*, 8(1):23–31.
- Gans, J. S. and King, S. P. (2000). Options for electricity transmission regulation in Australia. *Australian Economic Review*, 33(2):145–160.
- Henze, B. (2016). *Laboratory experiments on the regulation of European network industries*. PhD thesis, School of Economics and Management, Tilburg University.
- Hogan, W., Rosellon, J., and Vogelsang, I. (2010). Toward a combined merchant-regulatory mechanism for electricity transmission expansion. *Journal of Regulatory Economics*, 38(2):113–143.
- Joskow, P. (2006). Incentive regulation in theory and practice: Electricity transmission and distribution networks. Technical report, MIT.

- Koray, S. and Saglam, I. (2005). The need for regulating a Bayesian regulator. *Journal of Regulatory Economics*, 28(1):5–21.
- Laffont, J.-J. and Tirole, J. (1986). Using cost observation to regulate firms. *Journal of Political Economy*, 94(3):614–641.
- Littlechild, S. C. (1983). Regulation of British telecommunications' profitability. Report, The Secretary of State, Department of Industry, London.
- Loeb, M. and Magat, W. A. (1979). A decentralized method for utility regulation. *The Journal of Law & Economics*, 22(2):399–404.
- Neu, W. (1993). Allocative inefficiency properties of price-cap regulation. *Journal of Regulatory Economics*, 5(2):159–182.
- Perez-Arriaga, I., Rubio, F., Puerta, J., Arceluz, J., and Marin, J. (1995). Marginal pricing of transmission services: an analysis of cost recovery. *Power Systems, IEEE Transactions on*, 10(1):546–553.
- Perez-Arriaga, I. J., Rubio, F. J., Puerta, J. F., Arceluz, J., and Marin, J. (1996). *Marginal Pricing of Transmission Services: An Analysis of Cost Recovery*, pages 59–76. Springer Netherlands, Dordrecht.
- Rosellon, J. and Weigt, H. (2011). A combined merchant-regulatory mechanism for electricity transmission expansion in europe. *Energy Journal*, 32(1):119 – 148.
- Rubio-Oderiz, F. and Perez-Arriaga, I. (2000). Marginal pricing of transmission services: a comparative analysis of network cost allocation methods. *Power Systems, IEEE Transactions on*, 15(1):448–454.
- Sappington, D. (1980). Strategic firm behavior under a dynamic regulatory adjustment process. *Bell Journal of Economics*, 11(1):360–372.

- Sappington, D. E. M. and Sibley, D. S. (1988). Regulating without cost information: The incremental surplus subsidy scheme. *International Economic Review*, 29(2):297–306.
- Schill, W.-P., Egerer, J., and Rosellon, J. (2015). Testing regulatory regimes for power transmission expansion with fluctuating demand and wind generation. *Journal of Regulatory Economics*, 47(1):1–28.
- Sharkey, W. W. (1979). A decentralized method for utility regulation: A comment. *The Journal of Law & Economics*, 22(2):405–407.
- Vogelsang, I. (1989). Two-part tariffs as regulatory constraints. *Journal of Public Economics*, 39(1):45 – 66.
- Vogelsang, I. (2001). Price regulation for independent transmission companies. *Journal of Regulatory Economics*, 20(2):141–165.
- Vogelsang, I. and Finsinger, J. (1979). A regulatory adjustment process for optimal pricing by multiproduct monopoly firms. *Bell Journal of Economics*, 10(1):157–171.