Economic Foundations for 21st Century Freight Rail Rate Regulation*

John W. Mayo
Georgetown University
mayoj@georgetown.edu

Robert D. Willig
Princeton University
willig@princeton.edu

Abstract: The Staggers Rail Act of 1980 made a substantial break from an almost century-old policy of pervasively regulating the prices for freight rail services provided in the United States. In particular, rather than regulators establishing prices, Staggers permits shippers and railroads to voluntarily negotiate rates, terms and conditions, with regulation providing a fallback if negotiation fails or is too onerous for what is at stake on the shipment. While largely deregulating the rate-setting process, the statute requires that rates be "reasonable" in the event of a complaint filed by a shipper and where a railroad is found to be market dominant – a requirement that necessitates that regulators determine a method for assessing whether a given rate or set of rates is, in fact, reasonable.

In the wake of the passage of Staggers, the Interstate Commerce Commission (ICC) [now the Surface Transportation Board (STB)] established the method for determining whether a rate is "reasonable," in which case the rate is allowed to stand; or whether the rate is unreasonable and must be reduced. The system, known as Constrained Market Pricing (CMP), was established largely on a bedrock of economic theory and has been in place now for almost thirty-five years. Yet, absent a refresher, the passage of time creates the prospect that the economic foundations of CMP will fade. The purpose of this paper is to provide that refresher. Along the way we seek to re-establish the fundamental economic appeal of CMP, reflect on criticisms and alternatives that have been proffered, and offer refinements for rail regulation moving forward.

November 2018

*We dedicate this paper to the memory of William J. Baumol, who was the inspiration and co-author of its better thoughts, as well as a prime intellectual mover behind the creation and adoption of modern regulation of freight railroad services.

Economic Foundations for 21st Century Freight Rail Rate Regulation

I. Introduction

The Staggers Rail Act of 1980 made a substantial break from an almost century-old policy of pervasively regulating the prices for freight rail services provided in the United States. In particular, rather than regulators establishing prices, Staggers permits shippers and railroads to voluntarily negotiate rates, terms and conditions, with regulation providing a fallback if negotiation fails or is too onerous for what is at stake on the shipment. While largely deregulating the rate-setting process, the statute requires that rates be "reasonable" in the event of a complaint filed by a shipper and where a railroad is found to be market dominant – a requirement that necessitates that regulators determine a method for assessing whether a given rate or set of rates is, in fact, reasonable.

In the wake of the passage of Staggers, the Interstate Commerce Commission (ICC) [now the Surface Transportation Board (STB)] established the method for determining whether a rate is "reasonable," in which case the rate is allowed to stand; or whether the rate is unreasonable and must be reduced. The system, known as Constrained Market Pricing (CMP), was established largely on a bedrock of economic theory and has been in place now for almost thirty-five years. Yet, absent a refresher, the passage of time creates the prospect that the economic foundations of CMP will fade. The purpose of this chapter is to provide that refresher. Along the way we seek to re-establish the fundamental economic appeal of CMP, reflect on criticisms and alternatives that have been proffered, and offer refinements for rail regulation moving forward.

The chapter is organized as follows. We begin in the next section with a short tutorial on the fundamental economic characteristics of the rail industry. These characteristics necessarily have shaped, and will continue to shape, the consequent policy framework for the industry. Next, we turn to a discussion of the modern economic policy dictates of CMP. Our penultimate section discusses some of the criticisms of, and alternatives to, CMP that have surfaced. Finally,

we close with a discussion of opportunities for and challenges to freight rail regulation looking forward.

II. Foundations: the Economic Characteristics of the Freight Rail Industry

In 2020 the (de) regulatory framework established by the passage of the Staggers Act will be forty years old. As amply documented elsewhere, the Staggers Act and the light-touch regulatory decisions adopted first by the ICC and later by the STB have proven strikingly successful.[1] Prices today are lower than they were in 1980, output has expanded dramatically, innovation has flourished and private sector investment has grown to tens of billions of dollars per year. Yet, despite these striking aggregate statistics, certain characteristics of the industry make it a natural target for government intervention. First, there are substantial economies of scale in the provision of some rail services, whether along particular routes or for specific types of freight, which result from the heavy fixed costs associated with rail infrastructure and operations. To transport even small amounts of freight, a railroad must generally incur substantial costs of track, right-of-way, locomotive power, crew, and certain facilities-costs that do not rise proportionately with traffic volume. In such cases, fixed costs per ton of freight fall as traffic volume increases, and rail services are most economically supplied by a single carrier. Some rail services, in other words, are natural monopolies, inasmuch as other modes of transport such as trucks and barges are not close substitutes and supply by two or more rail carriers involves an unnecessary increase in the resources employed in the transport process.

While economies of scale and scope are endemic to freight railroad technology, a railroad may well have far less price-setting discretion than a textbook monopolist. A rail carrier's ability to control price for a shipment or shipments will depend on whether its services are shielded from competition by barriers to entry from within the industry narrowly conceived (freight rail services) and from competitive alternatives from its relevant market broadly conceived (transportation services on the same or substitute route, including practically feasible trucking and barge services). In the railroad industry, extensive capital sums must be sunk in

[1] See Mayo (2015) and the references cited therein.

way and structures and in a variety of ancillary facilities in order to create new rail lines. These sunk costs generally suffice to deter the entry of new rail lines. Despite these barriers, however, rival services and rival sources of supply (including trucks, barges, and alternative rail routes) may impose effective competitive constraints upon many, if not most, rail activities.

Another pertinent feature of the railroad industry is that there are substantial economies of scope which result from the common costs of rail operations. Outlays on rails, ties, rights-of-way, yard facilities, locomotives, and train crews are among the many common costs of rail operations incurred in carrying a variety of types of freight between a variety of origins and destinations. These costs confer economies of scope on carriers offering a multiplicity of transportation services: a carrier that provides an array of services can do so at a lower total cost than a set of carriers producing each service separately.

In the presence of substantial economies of scale and scope, there remain cases where competition is weak. The resulting monopoly power is the basic justification for regulation of rail rates and defines the basic task with which regulation must grapple. Ironically, it is these same features of the industry that render the industry particularly difficult to regulate in the public interest. In particular, the presence of substantial economies of scale and scope in the rail industry creates a number of problems for government regulation. Perhaps the most troubling is the fact that for purposes of assessing the reasonableness of a set of prices it is impossible to allocate, in any nonarbitrary way, a share of fixed and common costs to any one of a railroad's many activities. There is simply no way to subdivide those costs in a mechanical fashion that is unique and has any foundation in economic logic. The significance of this problem is much magnified by the fact that a substantial share of total railroad costs is fixed and common. In addition, if the regulator attempts to force rates to equal marginal costs, overall revenues will endemically fall short of overall costs. For rail systems that are characterized by scale and scope economies, rates must generally lie above the costs economically attributable to individual services if the firm’s revenues are to cover its total costs.

The prevalence of common costs and the financial infeasibility of marginal-cost pricing rule out any sensible mechanical or formula-based procedure for regulatory determination of

rail rates. In particular, compensatory rates cannot be determined by the regulator on the basis of cost data alone since the financial viability of any price depends also on the quantity of rail services customers are willing to buy at that price. Rational determination of prices must be based on both cost and demand conditions. But that, in turn, precludes the use of a price-setting formula since no formula can possibly capture the subtleties of demand behavior. For instance, demand varies from one prospective customer to another and among services; demands for various services respond differently to changes in the applicable prices and to market and general economic conditions; and demand is affected by hereogeneous competitive conditions.

Another relevant feature of the rail industry is that the supply of some services involves elements of monopoly power while the supply of other services is subject to strong competitive pressures. This creates real challenges for government regulation since, in the presence of economies of scope, it can be highly inefficient to separate the competitive and monopolistic services and have them provided by different suppliers. Alternatively, attempts to comprehensively regulate the portfolio of both competitive and monopolistic services offered by the rail carrier are likely to introduce substantial disincentives for innovative behavior.[2] Even if regulators attempt to narrow the scope of regulation to only noncompetitive services, distortions are likely to arise.[3] In particular, regulating the services over which a supplier has monopoly power may interfere with the efficient supply of competitive services. It is also important to avoid inducing anticompetitive behavior in the supply of the competitive services, such as cross-subsidies which permit underpricing of the competitive services at the expense of the customers of the other services.[4]

Finally, the efficient supply of most rail services requires cooperative behavior on the part of many firms. Rail transportation often involves interline service, with individual shipments traversing the tracks of more than one railroad. Similarly, more than one railroad is often involved in reloading and employing empty cars returning from deliveries. Other efficient

[2] See Mayo and Sappington (2016), reprinted as chapter 10 in this volume.
[3] Id.
[4] Brennan (1990).

cooperative arrangements involve the use of terminal, switching, and yard facilities. The challenge is to determine when it is appropriate to regulate cooperative relationships. Should railroads be free to negotiate agreements over responsibilities and the division of revenues? What if a railroad owns facilities that are indispensable for moving certain shipments?

III. The Economic Determination of "Reasonable" Rates in the Rail Industry

Faced with the various dilemmas posed by the fundamental economic characteristics of the freight rail industry, regulators are tasked with the job of establishing a method to protect shippers from exploitation of railroad monopoly power, where that exists, without distorting the otherwise efficient and competitive operations of railroads where there is no such monopoly power. In the corresponding legal and regulatory parlance, regulators are tasked with establishing a method by which a rate or set of rates can be determined to be "reasonable" or "unreasonable." Arguably, the presence of choice and competition, available to most but not all customers, is a powerful driver of reasonable rates. In recognition of this, the regulator's posture in developing its rate regulation methodology in the post-Staggers' era has been to focus on a set of "captive customers" that may be subject to the exercise of substantial monopoly power.[5] For these customers, the challenge is to establish a methodology to reveal, and constrain, unreasonable prices without creating excessive social losses from a large set of potentially unintended consequences, including: (1) extending the coverage of regulation to services in which customer choice and competition provide adequate customer protections and incentives for efficient contracts resulting from bilateral negotations; (2) misincentives in the amount and placement of rail investments; (3) disincentives to entrepreneurship and productivity advances; (4) pricing that creates unnecessary suppression of the outputs of the carriers and their customers; and (5) regulatory impediments to railroads' earning enough revenues to financially sustain themselves and to invest efficiently.

[5] "Captive customers" are defined by the STB to be those customers facing a set of rates and associated revenue (R) that exceed the variable cost (VC) of supply by eighty percent (i.e., R/VC > 180), or for whom it is determined that no viable substitute for the focal rail carrier's services exists. For a more detailed description, see Mayo and Sappington (2016).

In pursuit of these objectives several basic and salient tenets emerge from economic theory. In particular, economic theory indicates that two policy options, while alluring, should be avoided assiduously. First, the model of perfect competition is not a reliable foundation for establishing the reasonableness of freight rail rates. In this model, the free interplay of competition among existing firms and the process of free entry and exit of firms will lead to an industry equilibrium in which prices stabilize at levels that are equal to representative firms' marginal costs. These prices have the attractive feature that they are consistent with allocative efficiency, and are therefore often referred to as first-best prices. Accordingly, it is tempting to use this model as a benchmark to declare that "reasonable" rates are equal to the firm's marginal costs. The allure of this benchmark should fade, however, once it is recalled that the rail industry is subject to considerable economies of scale and scope. Indeed, "reasonable" prices set equal to a firm's marginal costs will fail to allow the firm to recover its total costs, with the result of firm bankruptcy and a failure of supply. In short, such prices are not a reasonable standard for "reasonable" rates in the rail industry.

Second, benchmarks of "reasonable" rates that derive from the allocation of fixed costs across services and customers demonstrably produce inefficient and distortionary rates. This point has been forcefully and repeatedly made over the past forty years so we will not belabor the full-blown arguments here other than to underscore the unambiguous conclusion from professionally accepted economic theory that any attempt to ensure that a rail carrier is made financially solvent through the use of allocations of unattributable joint and common costs that are independent of shippers' demands and willingness to pay will lead to gross economic inefficiencies.[6]

Fortunately, economic theory not only rules out inappropriate regulatory methods, but also can conversely provide a foundation upon which to develop a sound 21st century regulatory framework for the freight rail industry. That is, relevant and sound economic

[6] See our simple example, *infra*, as well as numerous articles, including, Braeutigam (1980), Baumol and Willig (1983), Sweeney (1982), Baumol, Koehn and Willig (1987), Davidoff and Hermalin (2004) and Mayo and Sappington (2016).

principles for adjudicating the "reasonableness" of rail rates do exist. To see this, first note that while cost calculations or estimations alone do not provide a sound economic benchmark for pricing, the combination of demand-side principles and cost principles offered from established economic theory can provide a foundation for guideposts and upper bounds on economically efficient and financially sustaintable rates, characteristics that appropriately provide economic content to the rubric of "reasonable rates."

From the demand side, because economies of scale and scope are pervasive in the rail industry, prices set equal to the rail carriers' marginal cost of providing transportation will fail to be compensatory for the firm. The question that arises in this circumstance is how might rates be adjusted above the marginal costs of providing the range of services sufficiently for coverage of the total costs of service, inclusive of common and fixed and all capital costs, while minimizing output distortions and losses to consumer surplus?

This problem (in a slightly different context) was addressed by Frank Ramsey in 1927, and the features of its solutions have now stood the test of time since then.[7] The answer, he demonstrates, is to raise prices on particular services above their respective incremental costs so that the resulting margins are inversely proportional to the price elasticities of demand for those services. The resulting set of rates (now referred to as Ramsey prices) allow a multiproduct firm to remain economically viable while minimizing the damage to consumers, and society more generally, from output reduction associated with the price elevations above marginal costs.

The logic of this inverse elasticity rule is straightforward. The elasticity of demand provides a quantitative interpretation of the traditional concept of value of service, which has played an important role in public utility pricing. Consumers who place relatively high value on a service will have demands for it that are relatively inelastic, and vice versa. If a rise in the price of a service would lead to no significant reduction in quantity demanded (that is, if demand is inelastic) then the service must be worth at least the higher price to its consumers

[7] Ramsey (1927). A clear restatement of Ramsey's seminal contribution which was set in the context of optimal taxation, making it applicable to pricing of a regulated firm is found in Baumol and Bradford (1970).

(that is, the value of the service must be high). Conversely, if a rise in the price of a service would lead consumers to curtail their demand substantially (that is, if demand is quite elastic), then the service must be worth little more to its consumers than the original price (that is, the value of the service must be low).

In view of this correspondence between value of service and demand elasticity, the inverse elasticity rule associated with Ramsey pricing can be restated in terms of a familiar and long-used principle in railroad pricing. Services with relatively high values to their customers should contribute relatively large net revenues to aid in the recovery of unattributable fixed and common costs. Conversely, services with relatively low values to their customers should contribute relatively small net revenues to aid in the recovery of unattributable fixed and common costs. All factors that influence a rail carrier's elasticities of demand are relevant for the carrier's Ramsey prices. These factors may include the value of the commodity shipped, the urgency with which the shipment is to be made, the extent of intermodal, intermodal and interport competition, as well as the substitutability of other commodities for the one shipped at its destination.

Under Ramsey pricing, the burden for recovering the firm's unattributable joint and common costs are borne by those customers who, when faced with the prospect of higher prices, would reveal through their continued patronage of the rail carrier that the specific shipment (or group of shipments) is of sufficiently high value that they will continue to be purchased despite the higher prices. While it may seem uncomfortable that higher prices are being set for customers whose circumstances dictate that they must continue to purchase despite higher prices or have no other transport options, the alternative of collecting the necessary revenues from customers who will halt their patronage in the face of higher prices is not feasible. Moreover, those customers who would halt their patronage reveal a lower value for such shipments; a basis for providing them with lower prices. Collectively, by permitting the firm to establish Ramsey prices, economic efficiency is promoted by providing the firm with revenues required to recover its total costs, while minimizing the necessary output reductions that occur as a consequence of prices that exceed their marginal cost. In short, in situations in

which purely allocatively efficient (marginal cost) pricing is infeasible, Ramsey pricing provides an economically efficient set of rates.[8]

As a matter of formal economic theory, Ramsey Prices are economically efficient as they maximize consumer and social real income (i.e. social welfare in terms of consumers' utility or consumer and producer surplus), subject to the stipulation that the prices generate revenues sufficient to cover the total costs of production, under the condition of increasing returns to scale that make marginal cost pricing financial insufficient. It is important to recognize that while the formal treatment of Ramsey Pricing is illuminating economic theory, it is not usually treated as the source of quantitative formulas to set actual market prices. Precise estimates of the needed elasticities of demand are rarely available, and they are apt to change in the market faster than they can be reliably reestimated. Consequently, the theory of Ramsey pricing is generally regarded as more of a source of conceptual guidance for the conduct of regulation than as a source of real-time quantification for prices.

The practical realities of the demand for freight rail services, however, provide private market incentives for rail carriers to negotiate efficient arrangements with substantial demanders of their services, in the absence of regulatory micro management. In particular, customers and their corresponding demands for rail services are substantially heterogeneous in nature. In such situations, suppliers exercise market judgment, learn from experience, negotiate with customers, try out pilot experiments and generally grope their way through decentralized decision-making to determine compensation based on both costs and volumes and values of service. The result, achieved through the voluntary interaction of the sellers of rail services and buyers of rail services, is the proliferation of individual tailored contracts.[9] The contracts are apt to charge differential prices to different customers in reaction to the heterogeneous

[8] This result is widely known across both the economics profession and within the rail policy community. See, for instance, Baumol and Bradford (1970), Baumol and Willig (1983), Sherman (1989), Braeutigam (1989), Train (1991) and Kaserman and Mayo (1995). Within the rail policy community, the appeal to Ramsey pricing in the post-Staggers' era was acknowledged by policymakers as early as 1985. See *Coal Rate Guidelines, Nationwide*, Interstate Commerce Commission, August 8, 1985.

[9] For a detailed discussion of the emergence and proliferation of such individual contracts in the U.S. freight rail industry, see Macher and Mayo (2018).

values that shippers experience from the particular services offered by their suppliers. It can also be expected that the individual contracts will include heterogeneous non-price terms that respond to particular needs of the customers, including, for example, assurances of deliveries' timeliness, reliability, and adaptability to the customer's dynamic circumstances.

With such individually tailored contracts, customers with relatively low values, but values that still exceed the supplier's marginal costs of production, are beneficially served at relatively low prices that lie between their values and the marginal costs of serving them. These customers receive positive net value from their purchases equal to the difference between their value of consumption and the price they pay, and the supplier receives a contribution to its profitability equal to the difference between that price and its marginal cost. Prices to other customers who experience greater value from their consumption are apt efficiently to provide relatively greater margins above marginal costs to the supplier.

In addition, it can be expected that the individual deals will include individuated volume discounts whose quantity break points and degrees of discount will themselves be heterogeneous. As a matter of economic logic and formal theory, mutually beneficial volume discounts are always a feasible addition to uniform pricing that is above marginal costs.[10] To see this, imagine as a simple example that without a volume discount the customer's price is $10/unit, the marginal cost of production is $7/unit, and the customer's demand is 9000 units. Then the supplier could offer the volume discount deal under which the first 9000 units carry the price of $10/unit and any additional units are available at the price of just $8/ unit. The customer may be stimulated to purchase more than 9000 units under this deal, to the customer's benefit,

[10] See, Willig (1978).

while the additional volume earns additional contribution to the supplier's profits with the still positive markup of $1 above marginal cost.

Thus it is plain that a volume discount deal exists that is mutually beneficial, and it should also be noted that the details of such a volume discount are individual to the customer – since the effective volume break points and the depth of the discounts depend on the customer's level of demand and on the customer's price elasticity of demand. Such individuated volume discounts are not only consistent with workable competition, but they are also necessary reactions to the forces of any such competition. An incumbent who did not make heterogeneous deals with its heterogeneous sizable customers, including individuated terms and volume discounts, would find its business diverted by active rivals or entrants who would make those mutually beneficial deals in a sufficiently competitive or contestable market.

Note too that while there may be a superficial equity appeal to a set of rates that are all elevated uniformly above marginal costs, the market's Ramsey-like prices improve economic efficiency, and raise total output and real social incomes relative to uniform mark-ups. In particular, customer demands with value greater than marginal cost but less than the needed average markup above marginal cost are excluded by such average-markup pricing. This, in and of itself, represents an inefficient loss of social real income (social welfare) since the lost utilization of the product would have been beneficial to the customer by more than its marginal cost of supply. Further, the loss of this demand means less sales volume and thus a higher average markup above marginal cost is needed to cover total cost from the remaining volume of sales whose customers must pay more. Moreover, the consequent higher prices likely exclude more customer demands with values that exceed marginal costs but that are not sufficient to bear the higher prices, causing additional loss of social welfare and additional loss of volume with yet further negative consequences. In sharp contrast, setting prices with

relatively small markups above marginal costs to customers with relatively low value demands allows those demands to be met with net benefits to social welfare and with positive contributions to total costs that allow other customers to pay less towards the recovery of total costs than they would have to if that volume were not gainfully transacted. Customers with relatively high value demands must be charged prices with relatively high markups above marginal costs in order to complete the recovery of the total cost of supply.

In sum, even though regulators are generally unable to determine specific Ramsey prices and efficient volume discounts for the market, Ramsey-like price schedules, referred to in the rail industry as "differential pricing," emerge organically in the market as a consequence of private market negotiations of rail carriers and shippers. The crucial learning for policy is that it is unnecessary and generally counterproductive for regulation to attempt to force pricing through fully allocated costs, uniform markups, or even quantified application of Ramsey Pricing formulae. Instead it is necessary to understand that differential price schedules, that reflect values of service and volume considerations, achieved as part of the buyer-seller interaction process, have the potential to promote the same efficiency results as Ramsey pricing.[11]

While negotiations between a rail carrier and a shipper with enough business to warrant them have the strong potential to achieve efficient Ramsey-like results, such a salutary outcome is more assured if any substantial inbalance in the bargaining power of the parties is counterbalanced by a regulated "reasonable" fall-back. If the rail carrier is market dominant so that the shipper is "captive" in that its logistics have no practical alternative to rail services, its bargaining position may be far out of balance inasmuch as the carrier's serving assets have other gainful uses while the shipper's assets may lose

[11] For a rigorous treatment and demonstration, see Baumol, Bailey and Willig (1977).

much of their value without reasonable rail service. In such circumstances, absent some regulatory restraint, short-sighted exploitation of monopoly power could lead to the carrier inequitably and inefficiently extracting much of the value of the shipper's business, with resulting repression of investment and long-term business prospects. In so doing, the carrier would likely be charging the shipper more than its long-run costs of serving the shipper, thereby bolstering its overall profits or cross-subsidizing other shippers with better logistical alternatives.

Under such circumstances of a market dominant rail carrier and a captive shipper, the existence of a well-designed "reasonable" regulated fall-back for a complaining shipper has two critical functions. First, it can offer the shipper a way out of what might otherwise become a punishing and at least long-run highly inefficient position of facing excessive pricing. Second, the known option of a well-designed regulatory "reasonable" fall-back for the captive shipper upon complaint will influence its negotiations with its market dominant carrier. The regulated fall-back available to the shipper can mitigate or eliminate what would otherwise be the substantial inbalance in the negotiating parties' bargaining strengths. The expectations about the effects of the well-designed regulatory fall-back can create a de facto protective bound to the outcome of the negotiations, without weakening the ability of the negotiations to reflect efficiently the character of the shipper's demands and needs as well as the costs to the carrier of supplying them. Thus, a well-designed regulatory fall-back for the pricing offered to a complaining captive shipper by a market dominant rail carrier can create appropriate protection against long-run inefficiently excessive rates, and by so doing can facilitate and stimulate negotiations to efficient contracts that are even better for the shipper than the regulatory fall-back.

What economic properties of a regulatory fall-back of "reasonable" rates would qualify it as "well-designed" to play the key roles of providing appropriate protection

to a captive shipper while promoting long-run efficiency of outcomes, whether it actually becomes the operative price or rather the de facto bound on the outcome of efficient negotiations? Four, in particular, emerge:

1. The fall-back should protect captive shippers from pricing that would provide cross-subsidies to other shippers or that would itself yield supra-competitive overall profits for the rail carrier.
2. The fall-back should protect the long-run sustainability of a socially efficient shipper – i.e. a shipper whose business could cover its long-run costs including the long-run costs of its needed transportation logistics.
3. The fall-back should avoid creating misincentives for the rail carrier subject to regulation, unlike caps on prices that are related to expended costs that consequently promote overspending and unlike fully allocated cost pricing that discourages serving shippers with under-average willingness-to-pay.
4. The fall-back should be consistent with Ramsey pricing in that generally (idiosyncratic exceptions aside) it would not confine pricing to below the level that would be part of the welfare optimal set of prices that provide coverage of the carrier's total costs (including a competitive return on capital).

The standard for regulatory "reasonable" rates that has the requisite economic properties is based on stand-alone cost - the cost that would be incurred by a hypothetical efficient *de novo* rail carrier to provide the service or services at issue. Formally, consider a customer or group of customers purchasing a set T of services with quantities given by the vector y_T from among a possibly larger set of N services offered by the supplier in quantities given by the vector y_N. The *stand-alone cost* of serving the customer (or group of customers) is the total cost that would be incurred by an efficient supplier of y_T were it to produce only those services without simultaneously producing any other services or additional quantities of any

services included in T.[12] Under the standard for regulatory "reasonable" rates based on stand-alone cost, or the "stand-alone cost test," the prices a customer is asked to pay are "unreasonable" if the revenues they generate from the customer exceed the stand-alone costs of the services the customer is to be provided.

Where the provision of the services a customer employs is implemented more efficiently in combination with the services purchased by a group of other customers, then the applicable and more protective form of the stand-alone cost test could aggregate those services and compare the total payments of the entire group of customers with the stand-alone costs of the totality of the services purchased by the group. Equivalently, the revenue paid by a particular customer is to be compared with the customer's net stand-alone cost: the stand-alone cost of provision of the group's total services, less the total revenues paid at contemporaneous prices by the other customers in the group. If the revenues paid by the total group exceed the stand-alone cost of the production of the group's total services, or equivalently if the particular complaining customer pays revenue in excess of its net stand-alone cost, then at least some of the prices involved are shown to be "unreasonable."[13] In short, under this standard of rate reasonableness, no buyer or group of buyers should be required to pay more for their purchases than the stand-alone cost of those purchases. Prices necessitating customer payments in excess of the services' stand–alone costs can be said to be "unreasonable."[14]

---

[12] Properly calculated stand-alone costs are determined from a long-run, forward-looking perspective. This follows since they represent the costs that a new entrant into the relevant market would bear, with no preset rigidities and with the ability to choose the current best available technology and the most efficient inputs. The capital expenditures that the entry plan requires must be seen as engendering capital costs, including competitive rates of return on capital investment that are comparable to those earned by firms outside the industry that experience equivalent levels of risk. Recognition that these costs must be expected to be covered by revenues if there is to be entry is necessary for the entrant to compete successfully in competitive capital markets for the financing it needs initially and over its life span. For more background on the economic theory of stand-alone costs see Willig (1979).

[13] Here, the prices charged to the specific complaining captive shipper are shown to be unreasonable under this standard if the associated revenues exceed the shipper's net stand-alone cost and if the payments made by the other shippers in the group do not exceed the stand-alone costs of their total services (excluding the services provided to the complaining shipper).

[14] Clearly, the stand-alone cost is unnecessary and inappropriate where there is competition. In a competitive market, the price offered by competitors will set a market ceiling. However, for any shippers

In accordance with the first of the aforementioned properties of a well-designed regulatory fall-back, passing the stand-alone cost test implies the absence of cross-subsidization.[15] Any customer that receives a set of services for no greater expenditures than their associated stand-alone cost is not harmed relative to the alternative of self-supply (or supply by a third party producing only those services), and is thus not paying extra to support other customers by means of cross-subsidies. Indeed, to the extent that the payments are less than the associated stand-alone costs, the customer derives pecuniary benefits that come from being part of the rail carrier's larger customer pool and the associated economies of scale and scope from serving that customer group. Cross-subsidies are properly of public policy concern because they generally lead to a misallocation of resources by encouraging inefficient investment by the recipients and repressed investment by the donors. Cross-subsidies may be of special concern to shippers because they are perceived to be unfair.[16] It has been shown that if all services and combinations of services satisfy the stand-alone cost test (i.e., they are not the source of cross-subsidies), then in the presence of zero economic profits (i.e., no monopoly profits overall) all services and combinations of services will also provide revenues that exceed their associated incremental costs (i.e., they are not the recipient of cross-subsidies).[17]

The stand-alone cost test exhibits the second of the aforementioned properties of a well designed regulatory fall-back. Absent regulatory protection, a captive shipper could be charged so much for its needed rail services that it would be unable to sustain itself financially over the long run, even if it could survive in the short run by neglecting maintenance, efficient

---

that are truly captive, which is to say that the rail carrier faces no effective direct, indirect, or potential competition for their freight, the stand-alone cost provides an economically rational ceiling.

[15] See Faulhaber (1975a) for early recognition and explanation of the relationships between cross subsidization and stand-alone costs.

[16] Note that if payments of one group of shippers help make up for smaller margins in payments by another, the first group might well believe it is being forced to cross-subsidize the second. Yet mere payment of a relatively higher rate is not evidence of a cross-subsidy where fixed and common costs must be covered. Rather, a cross-subsidy in an economic sense can occur only if a shipper (or a group of shippers) pays more than the total cost of serving it alone. If no shipper pays more than that amount, differences in their rates simply reflect differing contributions to the common costs of the system, not cross-subsidies.

[17] Faulhaber (1975a), (1975b).

replacement and monetizing depreciation of its sunk capital assets. Long run economic welfare would be harmed by this outcome if the shipper's business could cover its total social costs of long run operation, including the costs of provision of its needed rail services, but it is facing the commercial exigency of paying more for those services than their long-run cost to the carrier. Thus, the regulatory cap of stand-alone costs on the captive shipper's obligations to pay for its market dominant rail services protects the shipper and long-run social welfare from excessive rail rates impelling the contraction of beneficial economic activity.

The third property of well-designed regulatory fall-backs is satisfied since the stand-alone cost is the cost of service by an efficient hypothetical entrant that offers alternatives to the shippers at issue, rather than determined by any of the costs actually incurred by the regulated railroad. Consequently, under the system of stand-alone cost rate ceilings, a railroad has no incentive to pad or otherwise increase its expenditures for the purpose of relaxing a regulatory constraint. Further, the ceilings and their calculation apply only to services over which the railroad has monopoly power, so they do not diminish the railroad's incentives to pursue aggressively additional traffic and other new business opportunities.

That the stand-alone cost test satisfies the fourth aforementioned property of a well-designed regulatory fall-back follows from its intrinsic connection to the economic theory of contestable markets. By their theoretical definition, contestable markets are driven by the competitive forces from the threats of diversion of incumbents' business to potential entrants. In a perfectly contestable market, firms not already active are capable of entering into the production and sale of any quantities of any of the market's products or services by making use of the generally available technology and inputs, and thus incurring efficient economic total costs of that production. Such potential entrants evaluate the profitability of their entry into the market by comparing these costs with their potential revenues from diverting sales of the incumbent firm or firms at prices just below those charged to the incumbent customers. The potential entrants into a perfectly contestable market are assumed to face no entry barriers (unlike rail carriers facing substantial economic entry barriers), and this is the primary reason

why contestability is a benchmark model of workable competition rather than a generally accurate description of real markets.[18]

In contestable markets, no customer or group of customers would agree to pay more to a supplier for their services than it would cost to produce them efficiently in the long run on their own, or than it would cost a competing entrant to implement their supply. In other words, the competition from the potential entrants characteristic of the contestable market benchmark would drive reductions in incumbents' prices by the threat of stand-alone supply at stand-alone costs of each of the services and combinations of services. Thus, equilibrium outcomes in perfectly contestable markets necessarily pass the stand-alone cost test.[19] The stand-alone cost test imposes the same constraints, discipline and ceilings on customers' prices that the market would impose if the incumbent supplier were subject to robust competition from potential entrants with efficient technology and inputs and facing no barriers to entry.[20] In short, the stand-alone cost test provides the same protections to shippers that they would be afforded by effective competition in contestable markets.

The relationship between contestable markets and stand-alone costs underlies the consistency between the stand-alone cost test and Ramsey pricing that constitutes the aforementioned fourth economic property of a well-designed regulatory fall-back for reasonable rates. The so-called "Weak Invisible Hand Theorem" demonstrates that the Ramsey prices for a multi-product natural monopolist are consistent with equilibrium in perfectly contestable markets.[21] This implies that Ramsey prices pass the stand-alone cost test, since this test is one of

[18] For the most complete original treatment, see Baumol, Panzar and Willig (1982; and for discussion of the role of contestable markets as a benchmark for regulation, see Baumol and Willig (1986).
[19] This result is most thoroughly and rigorously documented in Baumol, Panzar and Willig (1982).
[20] It is economically efficient that the workable competition of contestable markets constrains incumbents' prices with levels of the forward-looking long-run economic costs that comprise stand-alone costs rather than any backward-looking historic costs or accounting conventions that may have their appropriate uses but that do not play a role in shaping the outcomes of competition.
[21] See Baumol, Bailey and Willig (1977). The assumptions needed to establish this result include "normal" demands and cost conditions that in essence result from a balance between economies of scale and economies of scope.

the necessary conditions for equilibrium in perfectly contestable markets. Accordingly, the stand-alone cost test does not confine prices away from their Ramsey optimal levels.

In summary then, we have shown that the stand-alone cost test can serve as the well-designed standard for a regulatory fall-back mechanism to protect complaining captive shippers against unreasonable rates by a market dominant rail carrier. It is well-designed in that it satisfies the economic properties delineated above: protecting shippers against providing cross-subsidies; protecting long-run efficient shippers from rates that would drive them from the market; promoting incentives for rail carriers to be cost efficient and to market aggressively; and maintaining consistency with optimal differential (Ramsey) pricing.

As a simple illustrative example of the application of stand-alone costs to pricing analysis, consider an incumbent whose services to customers A and B make use of common facilities with long-run recurring costs of 10, specialized facilities used only to serve A with recurring cost of 6 and specialized facilities used only to serve B with recurring cost of 9. The stand-alone cost of service to A is 16 (10+6) because the common facilities would be needed for that production even in the absence of service to B, along with the specialized facilities that are employed only to serve A.[22] Similarly, the stand-alone cost of service to B is 19 (10+9) because the common facilities would be needed for that production even in the absence of service to A, along with the specialized facilities that are employed only to serve B. The total recurring cost is 25 (10+6+9), which is significantly less than the sum of the stand-alone costs of A and B. This is a reflection of the firm's economies of scope, which means that joint production of the services is more efficient than the total cost of producing them all separately.[23]

Now suppose the incumbent charged prices that generated revenues of 18 from the services provided to A and 7 from the services provided to B. These prices could not hold sustainably in a contestable market since an entrant could produce service to A for its stand-

[22] Of course, those common costs might be efficiently reduced if they were to serve only A without B, and this would somewhat complicate the example without changing its lessons.
[23] Panzar and Willig (1981).

alone cost of 16 and divert that business from the incumbent who was charging 18 for it.[24] Alternatively, suppose the incumbent charged prices that generated revenues of 14 from the services provided to A and 16 from the services provided to B. Then these levels of revenues are below the stand-alone costs of 16 for just A and the stand-alone costs of 19 for just B, so that in contestable markets an entrant would not be able to profitably divert the incumbent's business of just A nor that of just B. However, an entrant would be motivated to divert the business of both A and B since the incumbent's revenues from that group of services are 30 (14+16) while the stand-alone costs for the services of A and B together are just 25.

This same example provides an illustration of the benefits of demand-based pricing (as per Ramsey Pricing), the economic dangers of pricing based on fully allocated costs, irrespective of customers' demands, and the connections to stand-alone cost limits on pricing. Suppose the value of the services to A is a flow of benefits of 9 and the flow of benefits to B is 18. Since the incremental costs of service to A are 6 and the incremental costs of service to B are 9,[25] it might seem under some accounting principles that the equitable prices to A and B would raise revenues of 10 from A and 15 from B. These prices are both the same percentage above incremental costs (67%) that is needed to boost the total of the incremental costs (15=6+9) to the level of the total cost of 25. Another common accounting principle might exercise "fully allocated costs," or "fully distributed costs," by dividing up the common costs of 10 in accordance with relative use, for example equally, to derive revenue targets of 11 (6+1/2 of 10) and 14 (9+1/2 of 10) for A and B respectively.[26] The pricing methods based on seemingly equitable accounting principles that neglect customers' individual values of service lead to dramatic economic harms for this market. Customer A does not experience sufficient value from its services to warrant paying the revenues of either 10 or 11 derived from the accounting principles. Then, with customer A out of the market, and making no contribution to the needed

[24] Another reflection of this unsustainability is that the revenue from services to B of 7 is less than the incremental cost of 9 of producing those services for B.

[25] These are the costs of the specialized facilities needed only to support the services of one of the customers.

[26] See R. Braeutigam, op. cit., for explanation of the general economic harms from such methods for pricing the outputs of a firm with pervasive economies of scale and scope.

costs, customer B with the much higher level of value of service of 18 must face its own stand-alone cost of 19, which exceeds its willingness-to-pay. The result is no service at all, and an incumbent impelled to price this way will soon be out of business.

In sharp contrast, Ramsey prices generating revenues of 8 from A and 17 from B are socially optimal here, because they leave both customers with benefits from their services that are greater than their payments, and they cover the total costs of service effected jointly with the available economies of scope. This is an illustration of how demand-based pricing, or more formal Ramsey Prices, maximize the use of the common facilities and the net benefits to consumers, while still providing financial sustainability to the efficient enterprise doing the production. And, here, the optimal prices for social welfare are fully consistent with the workable competition provided by perfectly contestable markets. The revenues from the optimal prices do not exceed the stand-alone costs of A, of B and of A and B taken together.

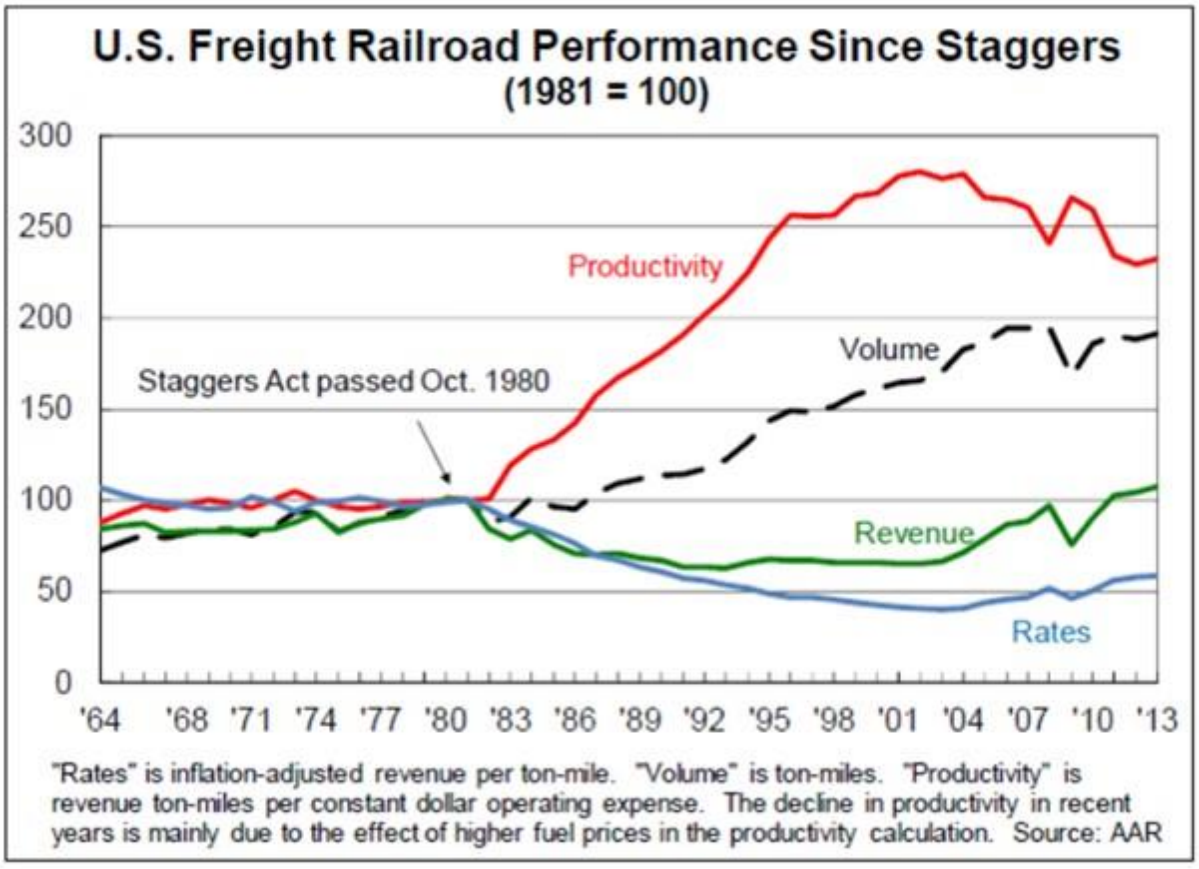
IV. The Suitability of the Economic Theory of CMP Going Forward

The lessons from economic theory for regulation of freight railroads that have been recounted above underlaid the formulation and acceptance by the ICC in the early 1980s of the regulatory system termed "Constrained Market Pricing (CMP)."[27] In particular, CMP embraces individualistic contracting between shippers and rail carriers, the railroads' ability to utilize differential pricing across customers (consistent with Ramsey pricing), and the rights of shippers to complain that they are captive to a market dominant rail carrier and thereby obtain protection against excessive pricing through the regulatory application of a stand-alone cost (SAC) test. The history of the U.S. freight railroad industry shows a dramatically successful industry turn-around resulting from the adoption of the CMP regulatory system as replacement for a previous disastrously counterproductive form of regulation, under the stimulus of the

[27] Indeed, the ICC specifically identified two economic concepts as foundational for CMP: differential pricing and contestable markets. See See *Coal Rate Guidelines, Nationwide*, Interstate Commerce Commission, August 8, 1985.

passage of the Staggers Act in 1980.[28] The complexity of that turn-around is concisely summarized by the data on the large U.S. freight railroads displayed in Figure 1. The figure shows that shortly after 1980, productivity began a new period of steep elevation, average prices began a long period of decline, revenues declined by far less than price and costs, and volume of service began an increasing trend.

FIGURE 1



Nonetheless, in recent years criticisms of CMP have emerged. They alternatively include charges that: (1) while the economic theory behind CMP is relevant, as a practical

[28] Indeed, the Staggers Act was so successful that the Senate Commerce Committee declared that it "is considered the most successful rail transportation legislation ever produced, resulting in the restortation of financial health to the rail industry." Senate Report No. 104-176 (1995). Economic studies have similarly found that the policy environment post-Staggers precipitated numerous positive economic improvements in the industry. See, e.g., McFarland (1989), Barnekov and Kleit (1990), Burton (1993), Ellig (2002), Winston (2005), Gaskins (2008), Caves, Christenen and Swanson (2010), McCollough and Thompson (2013), and Macher, Mayo and Pinkowitz (2014).

matter SAC is too cumbersome and complex to serve as a practical test of rate reasonableness; and (2) the economic theory underpinning the determination of the upper bound of reasonable rates is itself flawed.

We focus here on the second of these claims, noting only on the former that: (a) the 35 years of experience with rate cases utilizing SAC should permit the regulators to establish precedents in the application of practical considerations that can very much streamline the set of relevant issues in actual cases; (b) in recognition of the complexity of the original specification of the SAC test, the STB has developed the Simplified SAC method for adjudicating cases in practice;[29] and, (3) it is unsurprising and not unusual that important regulatory decisions are complex and require extensive procedural and substantive records to establish the congruence of public policy decisions with economic theory.

At the more conceptual level, a criticism has arisen that alleges that the SAC test is unyoked from the concept of fairness. That is, it is alleged that the SAC test does not provide a test of the unfair treatment of a customer, and therefore is not a sound measure of "reasonable" rates as required by statute. This charge, however, is inconsistent with at least one coherent definition of fairness. The fundamental linkage between the stand-alone cost test and fairness was laid out clearly by Professor Baumol in his book *Superfairness*.[30] He pointed out that the conventional equity consideration in price regulation historically centered on the concept of compensatory pricing. In particular, fairness would seem to require that each customer and customer group should pay at least the incremental cost of serving them; i.e., customers' prices should be compensatory. Rather than focusing on the inequity brought about by the recipients of uncompensatory prices, however, Professor Baumol enunciated the complementary thought that, "the object of an intercustomer fairness calculation is presumably to protect the interests

[29] A key simplifying assumption of the Simplified SAC is its reliance on existing infrastructure along the route used for transporting the traffic that is subject to complaint. In contrast, a full SAC case allows the complainant to design an operating plan that shows how an efficient railroad would serve the focal traffic, and determine the optimal network configuration. See *Rate Guidelines – Non-Coal Proceedings*, Surface Transportation Board, December 27, 1996.

[30] Baumol (1986, p. 120).

of …the potential *payers* who would be the victims of unfairness, rather than to inhibit the flow of benefits to the potential recipients of any subsidy."[31]

This reframing focuses on the fundamental unfairness of firm pricing that would extract more than reasonably required from a customer or customer group. This is precisely the stand-alone cost test. When prices violate the stand-alone cost test for a customer, that customer is inequitably and unreasonably required to provide to the railroad more revenues than would be necessary for the customer to efficiently self-supply the services. In this context, the customer (group) is required to unfairly cross-subsidize other customers.[32] Conversely, when prices satisfy the stand-alone cost test, the focal set of services demonstrably share in the economies of scale and scope afforded by a continuing relationship with the railroad at existing prices.

Another criticism has arisen from the observation that both Ramsey pricing and the theory of contestable markets are developed in a context where the firm earns zero economic profit, while freight railroads in the U.S. may earn more than that. However, it has been a reasonable judgement that the vast majority of rail shipments in the U.S. are in fact subject to the presence of competitive alternatives,[33] so that the working assumption of zero economic profit is generally tenable. Moreover, as a practical matter, detailed empirical analyses of the profitability of railroads in the United States reveal virtually no evidence of extra-ordinary profitability, indicating that market conditions have to a rough approximation produced results consistent with the underlying theories.[34]

Yet another criticism is also centered on the salient economic theory's seeming reliance on zero economic profits. Specifically, when positive economic profits exist, the economic equivalence of a stand-alone cost tests and incremental cost tests is severed. It is argued that this

[31] Id., (emphasis added).
[32] Braeutigam (1989) describes fair rates as those that are free from cross subsidies. He further observes that "The idea behind the (stand-alone cost) test is that if the revenues generated by [the focal] services exceed the cost of providing those services alone, then users of the services …are subsidizing users of other services." (p. 1339)
[33] For example, as noted in the introductory language of the Staggers Act "most transportation within the United States is competitive". Public Law 96-448, Findings, Sec. 2 (3).
[34] See Macher, Mayo and Pinkowitz (2019) in this volume.

result renders the stand-alone cost test inapplicable.[35] But while the uncoupling of the equivalence of the stand-alone cost test and an incremental cost test does occur with positive economic profits, it is easy to see that, if anything, the presence of such profits makes the case for the stand–alone cost test even more compelling.

To see this, initially suppose that profit (π) is zero. A violation of the stand-alone cost test for a subset of focal services *T* offered by a railroad providing the services *N* requires that:

$$\sum_{i \in T} p_i y_i > C(y_T), \tag{1}$$

where p_i is the price of service i, y_i is the output of service i and $C(y_T)$ is the cost of providing the relevant set ofservices in *T*. With zero economic profit, we know that:

$$\sum_{i \in N} p_i y_i = C(y_N), \tag{2}$$

where *N* is the set of all services provided by the railroad.

Subtracting (1) from (2) yields:

$$\sum_{i \in N-T} p_i y_i < C(y_N) - C(y_T) \equiv IC_{N-T}. \tag{3}$$

Note that $C(y_N) - C(y_T)$ is the incremental cost (IC_{N-T}) of providing the *N*-*T* services. Thus, (3) reveals that when the zero profits assumption holds the stand-alone cost test is equivalent to an incremental cost test. That is, violations of the stand-alone cost test for a focal set of services necessarily indicates the presence of non-focal services that fail to cover their incremental costs (i.e., they are non-compensatory). From a policy perspective, this would seem to make the application of the stand-alone cost test especially attractive because violations of stand-alone cost indicate not only that a group of customers is harmed by having revenues extracted from them that are greater than the costs of providing the service themselves, but also that these

[35] See, e.g., Faulhaber (2005) asserting that when positive economic profits exist SAC tests are "not helpful." (p. 446)

funds are used to cross-subsidize other customers whose rates are set below the incremental costs of supplying them.

But now consider the application of the stand-alone cost test when economic profits are positive. We can re-write (2) to include a positive profit $\pi > 0$:

$$\sum_{i \in N} p_i y_i = C(y_N) + \pi. \tag{4}$$

Now, subtracting (1) from (4) yields:

$$\sum_{i \in N-T} p_i y_i < C(y_N) - C(y_T) + \pi. \tag{5}$$

Here, while (1) still implies that more revenue is extracted from the focal services than the competitive benchmark provided by their associated stand-alone costs, (5) shows that the revenue from non-focal services (*N-T*) may or may not fail an incremental cost test.

Rearranging (5) yields:

$$\sum_{i \in N-T} p_i y_i - [C(y_N) - C(y_T)] < \pi. \tag{6}$$

The left-hand side of (6) is the excess of incremental revenues from the non-focal services above the associated incremental costs of providing those services. In the event that the pricing of the non-focal services is compensatory (i.e., the left-hand side of (6) is positive), then we see that the contribution of these services is less than the total level of monopoly profit earned by the supplier. In such cases, violations of the stand-alone cost test with positive total profits indicate that focal services are contributing to those monopoly profits. This condition would appear to strengthen the case for the application of the stand-alone cost test as in this event the pricing of focal services is sufficiently high to not only violate the stand-alone cost test, but additionally are a source of monopoly profits.

Alternatively, suppose that the left hand side of (6) is negative so that the non-focal services fail to cover their incremental costs. This condition indicates the presence of intra-customer cross-subsidies, with non-focal services benefitting from these transfers while

customers of focal services are payers of the cross-subsidy flows, and in addition fund the firm's total positive economic profits. In this situation, the logic is only enhanced for interpreting the failure of the stand-alone cost test as indicating unreasonably excessive rates. In sum, the merits of the stand-alone cost test are robust to the presence of positive profits, were that situation to arise. Another criticism of the application of the stand-alone cost test has been that it is developed within the context of contestable markets theory even though the provision of freight rail services does not comport with the underlying free-entry assumptions of a perfectly contestable market. Specifically, there is no doubt that there are high barriers impeding the entry of new rail carriers due to their need to sink substantial capital costs for rail infrastructure as well as their need to somehow assemble needed rights of way. However, this seeming incongruity is in no way a reason to doubt the validity of applying the lessons from contestable market theory to the formulation of applicable regulatory standards. In fact, if the industry were contestable, there would be no rationale for the application of regulation to it since the forces of competition from rivals and potential entrants would enforce reasonable and efficient pricing.

But, here, the theory is not intended to describe the industry actually as perfectly contestable. Rather it is used to evoke the counterfactual: what would be the properties of prices in the rail industry were the industry subject to the competitive forces present in perfectly contestable markets? The theory reveals that in such markets competition would yield prices consistent with stand-alone cost tests, and thereby protect shippers from excessive pricing. Thus, the stand-alone cost test emerges not as a consequence of the actual presence of contestable market conditions in the rail industry but rather as the standard that would guide regulation to offer captive shippers the same protection from excessive pricing that they would experience from workable competition in actually contestable markets.

While the CMP approach described herein and employed by rail regulators has largely adhered to sound principles of economics, several alternative approaches to separating "reasonable" from "unreasonable" rates have been suggested in recent years. One of the most threatening of these alternatives is that standard rate-of-return regulation will be applied to rail

carriers, perhaps due to interpretations of the statutory language that requires rail revenues to be "adequate."[36] Determination of the reasonableness of rail rates based on the measured profitability of the carrier, in terms of its expended costs and its book of capital assets, systematically generates market distortions and misincentives for the firm's decision-making. Indeed, a comprehensive literature review of economic research papers studying regulation indicates that the effects of such profit-based regulation generally included: "(1) limited incentives for innovation and cost reduction; (2) over-capitalisation; (3) high costs of regulation; (4) excessive risks imposed on customers; (5) cost shifting; (6) inapproapriate levels of diversification and innovation; (7) inefficient choice of operating technology; and (8) insufficient pricing flexibility in the presence of competitive pressures."[37] More specifically with respect to the rail industry, it has been shown that comprehensively regulating the earnings of railroads would introduce substantial disincentives to innovation.[38] Moroever, attempts to target earnings regulation to only a subset of rail services, say those that are thought to be contributing to positive economic profits, are also frought with economic distortions.[39]

However, in this regard, it is significant to note that the stand-alone cost test can validly be applied to the traffic of large groups of shippers that could, in concept, approach the totality of a carrier's operations, or its operations within one of its regions. Such a test would compare the shippers' total expenditures for their services, which are essentially the total revenues of the carrier, to the total costs of an efficient hypothetical entrant providing those services. This is a test of the profitability of that entrant, and not a test of the profitability of the actual carrier based on its expended costs and booked capital stock. Here, the notion of a regulatory constraint based on "revenue adequacy" is properly interpreted as whether actual prices generate more than adequate revenues to cover the stand-alone costs of the analyzed services, no matter how extensive they are. Since the stand-alone costs are long-run replacement costs, inclusive of competitive costs of capital, and independent of the actual expenditures of the

[36] For an extended discussion, see Macher, Mayo and Pinkowitz (2014).
[37] Sappington (2002, p. 240)
[38] Mayo and Sappington (2016), reprinted in this volume.
[39] Id.

carrier, such appropriate considerations of adequate revenues avoid most of the distortionary and repressive effects of standard rate-of-return regulation.

Another suggestion for determining rate reasonableness has been to simply establish a cap on the ratio of revenues to variable costs.[40] Such a cap is devoid of any economically grounded benchmark. Consequently any such cap would necessarily be arbitrarily chosen with the result that "reasonable" rates would be distinguished from "unreasonable" rates by the arbitrary judgment of the regulatory body at the time that a set of challenged rates were adjudicated. Most important, in the absence of an economically grounded benchmark (such as stand-alone costs), this proposal would satisfy none of the properties delineated above for a well-designed regulatory fall-back: It would not protect against cross-subsidization, it would not promote long-run efficiency in shippers' operations, it would engender misincentives inasmuch as carriers' conduct would influence the allowed ratio; and it would diminish overall consumer welfare by impeding Ramsey pricing.

Most recently a proposal has arisen to replace the existing rate reasonableness methodology utilized by the STB with an arbitration procedure.[41] This approach would rely on independent arbitrators to adjudicate the reasonableness of rates based on a final-offer arbitration method, as is used in the Canadian rail industry. The principal motivation for and appeal of the proposal is that it would substantially simplify the rate reasonableness adjudication process. "Reasonable" rates would be chosen by the arbitrator after "a brief hearing" with restricted opportunities to appeal arbitration rulings.[42] The Canadian model of arbitration is, however, completely confidential to the parties and therefore provides no basis for knowing what standards are used by the arbitrator to decide which among the alternative rates proffered are "reasonable." The absence of transparent standards that are demonstrably consistent with economic principles creates the distinct possibility of arbitrary arbitration outcomes. Of course, there can be no confidence that the arbitration decisions would be

[40] Pittman (2010).
[41] Transportation Research Board (2015), pp. 107-119.
[42] Id., p. 212.

consistent with the principles of well-designed regulatory fall-backs. The resulting uncertainty and pricing consequences are likely to be detrimental to both shippers and railroads.

Thus, while the goal of regulatory expedience is laudable, it is doubtful that the gains from implementing an expedited arbitration mechanism for determinining the reasonableness of rates is worth the corollary sacrifices that would occur to transparency, economic standards and expert adjudication embedded in the current system. But that is not to deny the possibility of other procedural innovations that could accomplish far greater speed and access to smaller shippers than today's regulatory procedures, without substantial sacrifice of the reliability of principled outcomes.

V. Conclusions

No regulatory mechanism is perfect either in theory or in practice. Consequently, periodic reviews of prevailing regulatory systems are warranted. In this chapter we have reviewed the economic underpinnings of the core elements of regulation used by freight rail regulators in the United States since the mid-1980s. We find that permitting individualized contracting, differential pricing by rail carriers (with the attendant Ramsey-like features), and availability of a regulatory fall-back based on stand-alone cost tests for captive complaining shippers constitutes an economically sensible and attractive basis upon which to regulate freight rail prices in the 21st century. This system has both powerful efficiency and equity features. Moreover, during the post-Staggers' period in which this light-touch regulatory structure has been in place numerous economic characteristics of the freight rail industry have much improved, including prices, output, investment, safety and innovation. It may be possible to further improve the performance of this regulatory system with attention to procedural innovations to speed, simplify and lower the costs of the needed ajudication, but the impulse for such worthwhile ends to undermine the principled operation of CMP regulation must be resisted.

REFERENCES

Barnekov, Christopher C. and Andrew N. Kleit. 1990. "The Efficiency Effects of Railroad Deregulation in the United States." *International Journal of Transport Economics* 17 (1): 21-36.

Baumol, William J. 1986. *Superfairness: Applications and Theory*. Cambridge, Mass: MIT Press.

Baumol, William J. and David F. Bradford. 1970. "Optimal Departures from Marginal Cost Pricing." *American Economic Review* 60: 265-283.

Baumol, William J. and Robert D. Willig. (1986) "Contestability: Developments since the Book," *Oxford Economic Papers*, New Series, Vol. 38, Supplement: Strategic Behaviour and Industrial Competition: 9-36.

Baumol, William J. and Robert D. Willig. 1983. "Pricing Issues in the Deregulation of Railroad Rates" in *Economic Analysis of Regulated Markets*, Jörg Fingsinger, Ed. Macmillan Press.

Baumol, William J., Elizabeth E. Bailey and Robert D. Willig. 1977. "Weak Invisible Hand Theorems on the Sustainability of Multiproduct Natural Monopoly." *American Economic Review* 67 (3): 350-365.

Baumol, William J., John C. Panzar and Robert D. Willig. 1982. *Contestable Markets and the Theory of Industry Structure*. Harcourt Brace Jovanovich, New York.

Baumol, William J., Michael F. Koehn, and Robert D. Willig. 1987. "How Arbitrary is 'Arbitrary'? -- Or, Toward the Deserved Demise of Full Cost Allocation." *Public Utilities Fortnightly* 120 (5): 16-21.

Braeutigam, Ronald R. "Optimal Policies for Natural Monopolies." In *Handbook of Industrial Organization*, Volume II. Richard Schmalensee and Robert D. Willig, Eds. North Holland, Amsterdam.

Braeutigam, Ronald R. 1980. "An Analysis of Fully Distributed Cost Pricing in Regulated Industries." *The Bell Journal of Economics* 11 (1): 182-196. https://doi.org/10.2307/3003407.

Brennan, Timothy J. 1990. "Cross-Subsidization and Cost Misallocation by Regulated Monopolists." *Journal of Regulatory Economics* 2 (1): 37–51. https://doi.org/10.1007/BF00139361.

Burton, Mark L. 1993. "Railroad Deregulation, Carrier Behavior, and Shipper Response: A Disaggregated Analysis." *Journal of Regulatory Economics* 5 (4): 417-434.

Caves, Douglas W., Laurits R. Christensen and Joseph A. Swanson. 2010. "The Staggers Act, 30 Years Later." *Regulation* 33 (4): 28-31.

Davidoff, Thomas and Hermalin, Ben E. 2004. "The Parable of Red Pens and Blues. " Lecture Notes for 201a, University of California, Berkeley.

Ellig, Jerry. 2002. "Railroad Deregulation and Consumer Welfare." *Journal of Regulatory Economics* 21 (2): 143-167.

Faulhaber, Gerald R. 1975a. "Cross-Subsidization: Pricing in Public Enterprises." *American Economic Review* 65 (5): 966-977.

Faulhaber, Gerald R. 1975b. "Cross-Subsidy Analysis with More than Two Services." *Journal of Competition Law and Economics* 1 (3): 441-448.

Gaskins, Darius W. 2008. "Regulation of Freight Railroads in the Modern Era: 1970-2010." *Review of Network Economics* 7 (4): 561-572.

Kaserman, David L. and John W. Mayo. 1995. *Government and Business: The Economics of Antitrust and Regulation*. Dryden Press, Ft. Worth, TX.

Macher, Jeffrey T. , John W. Mayo and Lee F. Pinkowitz. 2019. "The Law and Economics of Revenue Adequacy." in *U.S. Freight Rail Economics and Policy: Are we on the Right Track,* edited by Jeffrey T. Macher and John W. Mayo. New York, Routledge.

Macher, Jeffrey T. and John W. Mayo. 2018. "The Evolution of Contracting." Working paper. Georgetown University.

Macher, Jeffrey T., John W. Mayo, and Lee F. Pinkowitz. 2014. "Revenue Adequacy: The Good, the Bad and the Ugly." *Transportation Law Journal* 41 (2): 85–127. http://search.ebscohost.com/login.aspx?direct=true&AuthType=ip,uid&db=ofm&AN=102647998&site=ehost-live&scope=site.

Mayo, John W., and David E. M. Sappington. 2016. "Regulation in a 'Deregulated' Industry: Railroads in the Post-Staggers Era." *Review of Industrial Organization* 49 (2): 203–227. https://doi.org/10.1007/s11151-016-9524-1.

Mayo, John W., Statement before the House Subcommitte on Railroads, Pipelines and Hazardous Materials. Hearings on "The 35th Anniversary of the Staggers Rail Act: Railroad Deregulation Past, Present and Future," May 13, 2015. https://cbpp.georgetown.edu/sites/default/files/Anniversary_Staggers_House_Railroad_John_MayoJohn-20150513.pdf

McCullough, Gerard J. and Louis S. Thompson. 2013. "A Further Look at the Staggers Act: Mining the Available Data." *Research in Transportation Business and Management* 6: 3-10.

McFarland, Henry. 1989. "The Effects of United States Railroad Deregulation on Shippers, Labor, and Capital." *Journal of Regulatory Economics* 1 (3): 259–70.

Panzar, John C. and Robert D. Willig. 1981. "Economies of Scope," *American Economic Review*, 72 (2): 268-272.

Pittman, Russell. 2010. "Against the Stand-Alone-Cost Test in U.S. Freight Rail Regulation." *Journal of Regulatory Economics* 38 (3): 313–26. https://doi.org/10.1007/s11149-010-9130-3.

Ramsey, Frank P. 1927. "Contribution to the Theory of Taxation." *Economic Journal* 37 (1): 47-61.

Sappington, D.E.M. 2002. "Price Regulation." In *Handbook of Telecommunications Economics*, Volume 1. Martin E. Cave, Sumit K. Majumdar and Ingo Vogelsang, Eds. North Holland, Amsterdam.

Sherman, Roger. 1989. *The Regulation of Natural Monopoly*. Cambridge University Press. Cambridge.

Sweeney, George. 1982. "Welfare Implications of Fully Distributed Cost Pricing Applied to Partially Regulated Firms." *American Economic Review* 71 (3): 437-447. https://doi.org/10.2307/3003471.

Train, Kenneth E. 1991. *Optimal Regulation: The Economic Theory of Natural Monopoly*. MIT Press. Cambridge, MA.

Transportation Research Board. 2015. *Modernizing Freight Rail Regulation*. Special Report 318. The National Academies of Sciences, Engineering, Medicine.

Willig, Robert D. 1978. "Pareto Superior Nonlinear Outlay Schedules," *Bell Journal of Economics*, 9 (1): 56-69.

Willig, Robert D. 1979. "Multiproduct Technology and Market Structure." *American Economic Review* 69 (2): 346-351.

Winston, Clifford. 2005. "The Success of the Staggers Rail Act of 1980." *AEI-Brookings Joint Center for Regulatory Studies.*