Regulating the United States Railroads: 
The Effects of Sunk Costs and Asymmetric Risk*

JERRY HAUSMAN 
MIT 
Cambridge, MA 02139 
E-mail: jhausman@mit.edu 

STEWARD MYERS 
MIT 
Cambridge, MA 02139 

Abstract 
The Surface Transportation Board (STB) applies the theory of contestable markets to regulate dominant 
infrastructure markets. The STB bases its determination whether railroad revenues are excessive if 
they would be more than sufficient to support investment in a hypothetical stand-alone railroad designed 
to handle the at-issue traffic efficiently. The STB regulatory approach does not take account of the 
importance of sunk costs and irreversible investments in the railroad industry. We estimate how large the 
mistakes can be by applying a real options approach that takes into account the effect of sunk costs, 
irreversible investment, and asymmetric returns. 

1. Introduction 
Economic advice to regulators regarding the correct principles to set regulated prices has 
often been incorrect in that it does recognize the underlying technology of the industry. 
Economists recognized early on that in the situation of privately owned utilities in the 
United States that the first-best prescription of price set equal to marginal cost could not be 
used because of the substantial fixed (and common) costs that most regulated utilities 
needed to pay for. This realization typically accompanied the claim that the economies of 
scale of the regulated firm were so significant that competition could not take place

* Authors are MacDonald Professor of Economics and Billiard Professor of Finance, respectively. We thank 
Robert Willig for helpful comments and Paul Liu for helpful comments and research. Helpful comments 
are appreciated from conference participants at the IDEI conference on railroad regulation in Paris held in 
June 2000 and from the referees of this journal. We served as consultants for Union Pacific Railroad in a 
regulatory proceeding which considered the issues discussed in this paper.
1 See, for example, A. Kahn, The Economics of Regulation, vol. 1, New York, 1970.
because the regulated firm’s cost function was significantly below new entrants. Nevertheless, the most common advice from economists was that prices should be set similar to the outcome of a competitive process.

What the competitive process would be was typically never specified with any detail, which was to be expected since economic theory had no well-accepted model of competition with a technology exhibiting strong economies of scale, especially in the multi-product situation. In the United States regulators following legal principles adopted the position that the regulated firm should cover its costs. However, regulators also adopted prices for certain services to attempt to meet social goals for these given services. For other services, regulators used arbitrary means to set prices balancing competing claims from increasingly well organized groups of consumers, all of whom claimed they should receive low prices with other groups paying for the fixed and common costs.

This regulatory approach arguably did not do undue damage when no actual competition existed. So long as the regulated firm was (nearly) productively efficient, the losses were essentially second order social welfare losses. However, when actual competition appeared and was allowed to exist by the regulators, the economists’ advice of setting prices as if they were the outcome of a competitive process soon led to a regulatory morass. The outcome of a competitive process would need to take into account demand factors, cost factors, and competitive interaction (oligopoly) factors, with the first set of factors difficult to measure and the competitive interaction factors unlikely to be agreed upon. While regulators had some imperfect information about costs, they typically had little or no information about demand and no well-developed idea regarding the effects of competitive factors.

A particularly difficult problem arose when a regulated firm wanted to decrease its prices for services subject to entrant competition. Economists recognized that price set above incremental (marginal) cost should be permitted. New entrants wanted the previously set regulator set prices to be maintained. Furthermore, from a social welfare viewpoint the argument became first order since inefficient new firms could be productively inefficient causing a first order loss of social welfare.

Regulators found it difficult to permit the regulated firm to decrease its prices, especially since under cost of service regulation other prices would need to increase. Even when cost of service regulation was replaced by incentive (price-cap) regulation in the 1980s and 1990s, regulators still found it extremely difficult to allow price decreases since they believed in ‘‘regulated competition’’ (an oxymoron) where the regulators could better manage competition than the market. Nevertheless, the regulated companies were not harmed too badly since competition did not proceed at such a rapid pace to cause extreme economic damage.

However, in one regulated situation, the railroads, the advent of post World War II competition from trucking had left the regulated patient barely alive. The Interstate Commerce Commission (ICC) established in 1887 and one of the first federal regulatory agencies had been established to protect farmers from ‘‘unjust discrimination’’ by the railroads. By the early 1980s railroads were in a dire financial situation with bankruptcies common and railroads unable to make new investments. This outcome was a consequence of ICC established rates (prices) and its refusal to permit railroads to cease providing loss-services along low-volume routes. The United States Congress passed legislation, the Staggers Act, in 1980 that attempted to cure this dire situation of regulatory failure before the railroads ceased to exist. The Staggers Act allowed for private contracts between
shippers and railroads, allowed for railroads to stop providing service in certain situations, and most important allowed railroads to set their own rates unless they were dominant providers for a given origin-destination (OD) route.\(^2\) The definition of dominance was left for future regulatory proceedings, but a “safe harbor” was created by using the (arbitrary) rule that if the ratio of the rate to variable cost (R/VC) was less than 1.8, dominance would be assumed to be absent.

1.1. “Competitive Framework” Used by the ICC for Regulation
If dominance were determined, shippers could petition the ICC to set a regulation-determined rate.\(^3\) The same difficult questions then arose—since railroads are a joint production network-based technology with significant fixed and common costs, how should regulation set the prices on dominant OD routes? Here the ICC adopted a new development in economic theory, contestability theory, which claimed to be able to determine the competitive outcome in the multiproduct situation with significant economies of scale and fixed and common costs. Using the approach of Foulhaber (1975) who demonstrated that to avoid cross subsidy the upper bound for a service provided by a regulated firm should be its stand-alone costs (SAC), contestability theory demonstrated that under certain conditions competition would lead to prices that would satisfy the SAC outcome. The most important of the “certain conditions” was the absence of sunk costs, so that a new entrant could engage in “hit and run” entry. While contestability theory does not provide an accurate description of many industries, it certainly does not provide an accurate description of the technology of the railroad industry. In the railroad industry where sunk costs (along with joint and common costs) are the most important economic factors determining the technology, the regulators decided that an industry with free entry and exit (and no sunk costs) would provide the “competitive ideal.” Thus, in the situation where dominance is determined to exist, the ICC used the standard of SAC to determine rates since it would be the outcome of a competitive process in the absence of sunk costs. The Surface Transportation Board (STB) model applies the theory of contestable markets: railroad revenues are excessive if they would be more than sufficient to support investment in a hypothetical stand-alone railroad (SARR) designed to handle the at-issue traffic efficiently.

In this paper we demonstrate that choosing a competitive ideal, contestability theory


\(^3\) The ICC was abolished by the ICC Termination Act of 1995 (ICCTA). Its replacement the STB has continued the previous form of dominance regulation with only minor modifications.
that does not recognize the technological foundations of an industry can lead to incorrect results. We estimate how large the mistakes can be. We apply the results to a recent STB decision, *FMC Wyoming Corp. v Union Pacific Railroad.* Here the STB found in favor of the shipper that the railroad’s rates were excessive. If the economic effect of sunk costs had been recognized the outcome would have been reversed. Economic regulation must take account of the underlying technology, and assuming a competitive outcome inconsistent with the relevant technology can lead to large mistakes and disincentives for efficient investment and innovation.5

The STB has considered and rejected our suggested approach to take account of the effect of sunk costs in their regulation of the railroads. We examine the outcome of STB regulation in the final section of the paper. We observe that according to the STB’s own calculations that the railroads have not earned their cost of capital any time over the recent past, despite the unprecedented growth of the United States economy. We then explore the possible effects this outcome may have on U.S. railroads by comparing it to Great Britain, where the regulated railroad service provider recently was forced into bankruptcy when it was unable to raise sufficient capital to maintain and upgrade its network. The United Kingdom regulator used a similar regulatory approach to the United States in the sense that it ignored the effect of sunk costs. Improper regulation, as practiced in both the United States and United Kingdom can have real effects on railroad performance.

### 1.2. Regulatory Approach to Setting Price Where Dominance Exists

Suppose that the regulator, the STB determines that over a given OD route pair a railroad has market dominance, in the sense of having significant market power.6 The shipper then applies the “competitive economic theory” of the SARR approach, which the STB uses as a surrogate for the competitive constraint in applying price regulation to market-dominant traffic.7 The SARR is any railroad that the shipper can design that covers its SAC and carriers the subject, at-issue, traffic. The STB requires that the methodology recognize the assumed absence of entry and exit barriers (i.e., contestability theory) in an environment where sunk costs are a significant component of investment. Thus, the STB ignores the economic effect of sunk costs in an industry in which they are an important economic factor. Contestability theory is fundamentally inconsistent with the existence of sunk costs. As we explain later in this paper, the correct method to provide the correct economic incentives for investment is to apply modern options theory that takes account of the importance of sunk and irreversible investment. Failure to account properly for sunk

---

4 *FMC Wyoming Corp. v Union Pacific Railroad*, Doc. No. 42022 (May 12, 2000).


6 We will not analyze the economic factors that go into the dominance determination.

and irreversible investment will systematically understate the competitive rate standard against which benchmark the STB measures.

2. The Competitive Economic Theory of the SARR

2.1. Use of Contestability Theory for the Competitive Surrogate SARR

The ICC and now the STB have adopted the use of contestability theory (CT) as a competitive surrogate for railroad regulation when setting rates for railroads. CT explicitly assumes that no barriers to entry or exit exist. Railroad regulation by the ICC and the STB has used CT as the “competitive standard” to which sufficient competition among railroads and other forms of transport would lead. As the STB has stated: “The ultimate objective of the SAC constraint is to simulate a competitive rate standard for non-competitive rail movements by determining the rate that would be available to a shipper in a contestable market environment.”

To simulate a contestable market, a hypothetical SARR is designed with prices set to cover operating costs and a risk adjusted return on investment equal to the cost of capital for Class I (i.e., large) railroads. This SAC analysis is used to set the regulated prices.

But in simulating a contestable market, the ICC and the STB have acknowledged a critically important fact that faces any prospective railroad entrant: the presence of sunk costs. The STB has stated:

The objective of the SAC constraint is to simulate a competitive rate standard for non-competitive rail movements by determining the rate that would be available to shippers in a contestable market environment. A contestable market is one into which entry is absolutely free and exit absolutely costless, where the new entrant suffers no disadvantage relative to the incumbent. In simulating a contestable market in our SAC analysis, we do not eliminate sunk costs, rather, we assume that the costs that are sunk for actual rail carriers will also be sunk for the hypothetical efficient entrant.

Since the STB uses CT as its competitive benchmark, participants in rate cases must apply CT in a consistent manner to a hypothetical world in which no barriers to entry or exit exist. However, the STB also recognizes that a substantial proportion of investment for the hypothetical efficient entrant is sunk and irreversible, as a substantial portion of the actual investments made by railroads are sunk and irreversible. This framework requires the hypothetical SARR to maintain its rates at levels constrained by the threat of future and barrier-free entry by another SARR. Otherwise, inter-temporal cross subsidy will arise where later users of the SARR will cross subsidize earlier users of the railroad.

---

8 West Texas v. Burlington Northern (p. 27).
9 Nevada Power I (pp. 266–267) (footnotes omitted).
10 The ICC and STB have both stated that the primary use of SAC is to rule out cross subsidy as explained in the Faulhaber (1975) paper. See, for example, Coal Rate Guidelines, Nationwide, I 1.C.C.2d 520, 527 (1985): “One means of assuring that such cross-subsidization does not occur is the ‘stand-alone cost’ (SAC) test.”
Inter-temporal cross-subsidy will lead to regulated rates below the economically efficient level. From an economic analysis standpoint, the application of CT in a consistent manner to the hypothetical SARR leads to results similar to applying modern options theory to investment in railroads where the sunk and irreversible nature of much of the investment is accounted for.

2.2. Economic Effect of Potential New Entrants to Compete with the SARR

One way to model the lack of barriers to entry and exit ("hit and run entry") is to realize that the application of SARR theory must account for the threat that another SARR will enter if economic conditions permit. That is, the application of CT theory posits no barriers to entry or to exit in any period, so entry would occur whenever economic conditions would permit the second hypothetical SARR, say SARR2, to earn above its cost of capital. Otherwise, inter-temporal cross subsidy would occur. Thus, the application of CT establishes a limit on upside outcomes for the first SARR, say SARR1.

However, if the economic situation becomes worse than expected there will be no competitive entry. Instead, SARR1 will suffer losses due to non-retrievable investments that are the sunk costs that arise from the irreversible investment. Consider the example of an investment in a bridge or tunnel for a railroad in a contestable market. If economic conditions take a turn for the worse—say, there is an unexpected and significant decrease in demand in the future for rail service—SARR1 cannot sell the tunnel or use the tunnel elsewhere because of the sunk nature of the investment. So while CT assumes SARR1 (and any future entrant) could exit the industry costlessly, the railroad will bear the decreased economic value of its tunnel due to the unexpected decrease in demand for rail service because it cannot sell the tunnel for use elsewhere. Thus, while there exists a limit on the upside of the (probability) distribution of potential economic returns due to the absence of barriers to entry, no floor exists on downside outcomes due to the existence of sunk costs. The only limit on the downside is a complete loss of the investment in sunk assets. As previously stated, an actual railroad faces the same situation; if economic conditions worsen (e.g., an unexpected decrease in the demand for rail service), it cannot shift its investment in railroad facilities to other purposes because of the sunk nature of many railroad investments.

2.3. The Effects of Uncertainty from Future Revenue Projections under Contestability

Forecasts used in future revenue projections in simulations of the economic performance of the SARR are typically point estimates of the average of possible future outcomes. The further the time in the future, the greater the uncertainty and the greater and more diverse the possible future outcomes. Since no one can predict the future for certain, a band of uncertainty always surrounds a forecast. In a contestable market, worse-than-

---

11 For the purposes of the subsequent real options analysis, the SARR is assumed to “step into the shoes” of the first stand-alone railroad (SARR1) if the revenues are sufficiently high such that the appropriate cost of capital is achieved. The second stand-alone railroad (SARR2) is identical in every way to the first stand-alone railroad, with the same network, same real investment costs, and the same cost of capital.
expected outcomes will not attract competitive entry while better than expected outcomes are likely to attract competitive entry. Thus, the band of uncertainty will be asymmetric, and the actual expected value of future revenues for SARR1 will be less than the mean forecasts because of the asymmetric nature of market entry. The forecasts will be upward-biased for future revenues because of the failure to allow for the effects of the possible future competitive entry that can occur in a contestable market.

Thus, in considering the revenue forecast for the SARR, this fundamental mistake which arises by not taking into account the asymmetry in possible future results is an upward biased forecast of the future economic performance of the SARR. To be correct, the application of CT theory must account for a possible SARR2 and SARR3 and so on, all of which would find it economically worthwhile to enter when economic times were especially good. The concept of no barriers to entry leads to the truncation of the distribution of possible future returns. We incorporate the effect of the truncation of returns in the analysis of a SARR and demonstrate how significant the correction turns out to be.

It is important to note that one cannot use the observed cost of capital in the railroad industry, apply it to the SARR, and claim that the effects of truncation are already accounted for since actual railroads do not operate in a contestable world. Indeed, economic analysis of irreversible investment under uncertainty demonstrates that a given project typically must offer expected returns well above “breakeven”, including the cost of capital, before the project is undertaken because of the sunk and irreversible investment. In a contestable world with the availability of “hit and run” entry, a competitor will enter as soon as the breakeven point is reached because investments are not irreversible under contestability assumptions. Thus, the real world effect of irreversibility is assumed away when contestable markets are assumed. Contestable markets do not allow for sunk and irreversible investment.

Similarly, it is incorrect to claim that the outcome of the theory would be to cause the hypothetical SARR to have a higher cost of capital than an actual (real world) railroad. The asymmetric risk and symmetric returns, which arises from the concept of no barriers to entry and leads to the truncation of the distribution of possible future returns, must be compensated for investors to undertake the investment. The cost of capital for the SARR remains the same, to the extent that it has the same financial characteristics as a real world railroad, but the expected return for a project must increase for the project to be undertaken

---

12 The ICC’s logic in Nevada Power I (p. 267) means that SARR1 could not retaliate by lowering prices to reduce revenue for the SARR2 to make its entry more difficult.

13 The application of modern options theory to sunk investments demonstrates that companies require an options adjustment which is greater in early years to account for the irreversible nature of the investment, and therefore the inability of the company to redeploy the resources should circumstances change for the worse. The sunk nature of the investment creates the prospect that in future years, if the return is less than anticipated, no return on the investment will be earned, and therefore the options adjustment is required in the earlier years to offset that possibility.

14 See, for example, Dixit A. and R. Pindyck, Investment Under Uncertainty, Princeton, 1994. The higher return compensates for using up the “option to invest.” Thus, the analysis is often called “real options analysis.”
because of the truncation of the expected returns. Indeed, the investment rule with sunk and irreversible investments becomes in the situation of constant changes in prices and economic depreciation:

\[ V > \frac{\beta_1}{\beta_1 - 1} (\delta + \lambda) I, \]  

(1)

where \( V \) is the expected value of the investment project, \( I \) is the cost of the investment, economic depreciation is at a constant rate \( \delta, \lambda = r - \alpha \) the difference of the risk adjusted discount rate and the expected price change (net of variable costs) of the output of the project, and \( \beta_1 > 1 \), where the parameter \( \beta_1 \) depends on \( \lambda, \delta \), and the amount of uncertainty in the underlying stochastic process. It follows that the markup compared to the case with no sunk investment is \( m = \beta_1 / (\beta_1 - 1) > 1 \). The parameter \( \beta_1 \) takes into account the sunk cost nature of the investment coupled with inherent economic uncertainty. Parameter \( m \) is the markup factor required to account for the effect of uncertain economic factors on the cost of sunk and irreversible investments. The formula demonstrates that the risk adjusted discount rate \( r \) does not change, but that a markup is required to take account of the sunk and irreversible investment.

Rearranging equation (1) gives:

\[ \frac{V}{m} > (\lambda + \delta) I. \]  

(2)

For fixed, but not sunk costs, \( m = 1 \) so that the usual investment rule holds. However, for sunk costs the factor \( m \) lead to a “markdown” of future returns because of the sunk and irreversible nature of the investment. In poor economic times the fixed cost investments can be used in other projects, but the sunk investment cannot be reused in other projects. Thus, acceptable investment projects must have a higher expected return with sunk and irreversible investments, but the cost of capital to the firms making the investments does not necessarily increase.

3. Application of Real Options Theory to Investment under Uncertainty

3.1. Real Options Theory

At approximately the same time that the ICC adopted the current approach to regulating dominance in railroads in Coal Rate Guidelines, the development of real options theory

---

15 See, for example, Dixit and Pindyck, op. cit. for a further discussion of this point and J. Hausman, “Valuation and the Effect of Regulation on New Services in Telecommunications,” Brookings Papers on Economic Activity: Microeconomics, 1997.

16 In our application we do not assume constant rates of growth and price changes so that a closed-form expression will not exist. However, the basic economic parameters take the same form.
was beginning at MIT, applying to real assets the famous Black-Scholes approach to investment in financial assets. Economic theory states that with investment in a sunk and irreversible asset that you are giving a “hostage to fortune” because you cannot remove the investment—which is why the investment is sunk. Real options theory estimates the value of the option of waiting to invest to decrease the inherent uncertainty in a risky investment.

3.2. Value of an Option to Invest

When railroad networks are constructed (and as their components wear out and are replaced or added to or improved), they are in large part sunk investments. And, as the STB has determined, “we assume that the costs that are sunk for actual rail carriers will also be sunk for the hypothetical efficient entrant.” Therefore, when simulating a contestable market, the analysis must recognize that if the economic return for the SARR falls below competitive levels, the SARR cannot shift the investment in sunk assets to other uses because of their sunk and irreversible nature. Thus, the use of a perfectly contestable market standard must recognize the important feature of sunk and irreversible investments. Because of its failure to take into account the sunk and irreversible nature of investments, the STB approach fails to simulate a real world railroad market. An industry cannot be expected to behave in a manner that is fundamentally inconsistent with its underlying technological and economic characteristics, a fact that the STB has understood in its recognition of sunk costs.

3.3. Potential Solutions to the Problem

If all shippers signed contracts for the life of SARR1, the options theory would be satisfied because SARR1 would not bear any asymmetric risk on these contracts and would be guaranteed the returns on which prices calculated by the STB SARR analysis are based. Thus, the predicted revenues would equal the actual revenues for SARR1. Instead the shipper would bear the risk of its future demand uncertainty, and the assumptions built into the SARR analysis would hold true.

If lifetime contracts are not used, prices charged to the shippers in SARR1 must include payment for the value of the option in each period that shippers receive because they can decide to use or not to use the SARR1. Otherwise, shippers with shorter contracts or no contracts are given a “free option” by STB regulation through the fictional railroad SARR1. Thus, shorter contracts on SARR1 would have higher prices because of the value of the option not to use SARR1 if the shippers had better economic opportunities in the future. These conclusions are consistent with the real-world experience of railroads, where long-term contracts typically receive a lower price than similar contracts that have a shorter duration.

17 Nevada Power II (p. 267).
18 This feature of sunk and irreversible investment has been widely recognized by economic research in the past 10 years. See, for example, Dixit and Pindyck, op. cit.
19 The life-of-SARR contracts would have to be indexed properly to account for changes in input costs during the covered period.
4. Application of the Real Options Approach

There are two techniques for applying the real options approach. One solution to a given real options problem typically involves the solution to a partial differential equation with associated boundary conditions. Because of the 20-year horizon of the SARR as required by the STB, an analytical solution of the differential equation is impossible. When the solution of the partial differential equation is intractable, an alternative “Monte Carlo” approach estimates the solution to the partial differential equation. The Monte Carlo approach simulates many sample paths and to find the solution for each replication. We explain our Monte Carlo approach in detail in the Technical Appendix. The Monte Carlo simulation makes a draw from the underlying stochastic distribution and determines the economic outcome based on this random draw. This procedure is repeated many times and the random distribution of the outcome is given from the thousands of times the model is run, each time with a different random set of inputs that are drawn randomly from the same underlying distribution.

We use Monte Carlo simulations for a lognormal distribution where thousands of simulations (that use random outcomes that arise from the stochastic distribution used to model uncertainty) lead to the correct empirical distribution. Thus, the Monte Carlo approach in our SAC analysis accounts for the effects of uncertainty regarding the economic return to investment.

4.1. Methodology of the STB Model

The STB model uses the theory of contestable markets as explained above. Therefore, if an incumbent railroad has market dominance over a type or class of traffic, its rates and revenues are deemed excessive if a rational investor, earning its risk-adjusted cost of capital, would build a SARR designed to enter and take over the incumbent’s traffic. If this is the case, the STB then considers whether rates for the at-issue traffic should be constrained.

The STB model is detailed discounted cash flow analysis. The following points highlight some major assumptions.

1. The SARR is a stand-alone corporation, partly debt-financed.
2. The analysis is based on after-tax cash flows to debt and equity investors.
3. The discount rate equals the weighted average cost of capital for Class I railroads.
4. Capital investment is itemized by category. For each category, an inflation rate for future reproduction cost is projected. The model calculates an overall growth rate of reproduction cost for each year as a weighted average of the inflation rates category by category.22

20 This approach leads to the closed form expression that we analyzed above.
22 The average growth rate increases each period, because the categories with more inflation receive more and more weight. In other words, the growth rate accelerates over the life of the SARR.
5. The model assumes that the SARR’s assets are replaced at the end of their useful lives, and that replacement cycles continue in perpetuity.
6. The model calculates the SARR’s capital charges to a 20-year horizon. Capital charges are cash flows before interest, taxes and depreciation. The time pattern of capital charges depends on the overall inflation rate of reproduction cost and also on the projected growth in volume shipped. After-tax cash flows to investors are also calculated.
7. A terminal value is included at year 20 to capture the value of the SARR at that horizon. The terminal value is calculated by assuming that the capital charges in the horizon period will continue to grow in perpetuity at the ending overall growth rate in reproduction cost.

Figure 1 illustrates how the SARR’s capital charges are assumed to grow over time. The terminal value is determined by the level and growth rate of capital charges at the 20-year horizon.

The STB model solves for the initial level of capital charges that makes the present value of after-tax cash flows to investors equal to the required investment.23 Thus, the STB model calculates the capital charges just sufficient to support investment in the SARR. After capital charges are calculated, operating expenses are added to obtain total SAC—i.e., the future revenues necessary to justify investment in the SARR. This completes the first stage of the STB model.

23 The model actually solves for the capital charge per ton, which is multiplied by initial volume.
In the second stage, these SAC are compared to actual and projected revenues for the SARR’s traffic. The differences between revenues and SAC are projected for 20 years and the present value of the differences is calculated. If the present value is greater than zero—that is, if there are excess revenues, the STB moves on to consider whether rates should be reduced. At this stage, certain other tests apply. For example, rates are not reduced unless revenues exceed a jurisdictional threshold of 180% of variable cost, which is the dominance standard discussed above.

4.2. Uncertainty and Contestable Markets: Asymmetric Risk and Returns

The STB model does not explicitly recognize uncertainty. When revenues are risky and investment costs are sunk, a rational entrant would require higher rates and revenues than the STB model produces. Explicit analysis of uncertainty and risk has not previously been used in STB rate proceedings. However, uncertainty matters when investment is sunk and irreversible. It matters because the effects of uncertainty in the railroad industry are asymmetric as we discussed above. A SARR would face this kind of asymmetric risk and asymmetric returns. Asymmetric risk and returns increases the required capital charges and SAC for a SARR. If the capital charges are not adjusted for asymmetric risk and returns, they will be understated by a significant amount as our simulations demonstrate.

4.3. Implications for Railroad Investment

Much of the investment in a railroad is sunk and irreversible. The time it takes to build a railroad is much less than the life of the railroad’s assets once they are in place. A railroad facing competition, either from other railroads or other modes of transportation, faces upside limits. But there is no offsetting downside protection. Thus the SARR, or any similar railroad investment, would face asymmetric risk, as shown in figure 2b, which differs from the no competition outcome of figure 2a where no upside limits exist.

An investor considering construction of a SARR would recognize this asymmetry, and would not go ahead unless revenues and profits were sufficiently high to compensate. In other words, the investor would not go ahead just because a routine extrapolation of today’s traffic and revenues “looks good.” The investor would not go ahead just because the discounted value of the most likely future cash flows exceeds the required investment. Prospects have to be better than that. The greater the uncertainty, the more cautious the investor would be.

Most of the traffic on an SARR would be subject to competition. Competition would limit its profits on this traffic, just as shown in figure 2b. What if the actual incumbent railroad is protected from competition and is market-dominant? Absent regulation, a market-dominant railroad might have a less-constrained upside as in figure 2a. However, regulation constrains the pricing power that a “real world” incumbent can exercise.24

---

24 Thus, an alternative motivation for the asymmetric risk approach that we use is the effect of STB regulation. Shippers will only approach the STB when economic outcomes are “good,” and they believe they can achieve a rate reduction. However, the railroads are unable to increase their revenues when economic outcomes are “bad.” Regulation will truncate the distribution of outcomes and create asymmetric risk.
Figure 2a. SARR potential range of outcomes.

Figure 2b. SARR operating in a competitive market.

Figure 2c. SARR operating in contestable markets.
The contestable market standard imposes an upper bound on the revenues and profits that the incumbent or the SARR could earn in the future. Figure 2c shows how entry by a future SARR would put the best outcomes out of reach. The range of outcomes is truncated by the future operation of contestable markets. If a SARR were built today, and later profits turned out to exceed today’s projections, then another SARR could enter, or threaten to enter. This constrains the upside. But there is no constraint on the downside, because much of the investment in the SARR would be sunk and irreversible.

This asymmetry of payoffs is not a barrier to entry or exit. It is a cost of doing business when investment is sunk and irreversible. For example, in contestable markets assets can always be sold at their full economic value. But that value is uncertain, and it is limited on the upside by entry, or by the threat of entry. There is no similar protection on the downside. Therefore uncertainty means that the economic value of the SARR’s assets can fall further than it can rise. Asymmetric risk has nothing to do with price wars, or with one-on-one rivalries between two competitors for the same business. In contestable markets, the rival replaces the incumbent. Thus an existing SARR’s future profits are limited by the possibility that it will be replaced. But there is no replacement, or threat of replacement, if future profits turn out much lower than originally projected.

Asymmetric risk and returns increase the capital charges required by a potential entrant in contestable markets. Capital charges are not a barrier to entry. They are a condition for entry, and therefore a constraint on the revenues and profits of the incumbent railroad. The key is to get the capital charges right. That requires explicit analysis of uncertainty, which we now proceed to do.

4.4. How Risky is the Railroad Business?

The railroad business is not a “safe business”—as is readily confirmed by examining the risk of railroad common stocks. Railroad stocks respond to all of the risks that may affect the value of railroad assets and operations. In financial markets, risk is measured by the standard deviation of the stocks’ rates of return to investors.\(^{25}\) Table 1 shows railroad standard deviations for a recent five-year period, 1993–1997.\(^{26}\)

Figure 3 shows that the normal standard deviation for large railroads is between 15% and 20% per year. In our application we use a range for the standard deviation of 15%–25% per year.

---

25 Standard deviation is the correct measure for analysis of asymmetric risk. Standard deviation would not be used for estimating the cost of capital, because much of the volatility of individual stocks can be eliminated in a diversified portfolio of stocks. But the effects of asymmetric risk on asset values and required capital charges cannot be eliminated by diversification. When investments are sunk and irreversible, asymmetric risk reduces average operating profits and cash flows.

26 Some of the volatility of railroad stocks is due to financial leverage, i.e., to borrowing. Borrowing amplifies the risks borne by stockholders, because debt is a prior claim on assets and operating income. The effects of leverage on volatility have been removed in the calculations.
Table 1. Standard Deviation of Returns to Investors Individual Railroads

<table>
<thead>
<tr>
<th>Large Railroads</th>
<th>Standard Deviation (Annual)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Burlington Northern</td>
<td>14.9%</td>
</tr>
<tr>
<td>CSX Corp</td>
<td>15.7%</td>
</tr>
<tr>
<td>Kansas City Southern Inds Inc</td>
<td>25.1%</td>
</tr>
<tr>
<td>Norfolk Southern Corp</td>
<td>12.8%</td>
</tr>
<tr>
<td>Union Pacific Corp</td>
<td>14.0%</td>
</tr>
<tr>
<td>Canadian Pacific</td>
<td>15.0%</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td><strong>16.2%</strong></td>
</tr>
</tbody>
</table>

Note: The standard deviation of returns on the S&P 500 index is for comparison.
The S&P returns are not adjusted for financial leverage.

Figure 3. Standard deviation of returns to investors.

4.5. Calculating the Effects of Uncertainty on a SARR

We now apply the analysis of uncertainty and asymmetric risk to a SARR, used in a recent regulatory proceeding. Figure 4 summarizes the implications of asymmetric risk for the SARR. The lower dashed line represents the capital charges that would be calculated from the standard STB model—in other words, the projection of pre-tax operating cash flows that would be just (barely) sufficient (ignoring asymmetric risk) to

---

27 *FMC Wyoming Corp. v Union Pacific Railroad*, Doc. No 42022 (May 12, 2000). We have altered the exact numbers to preserve the confidentiality of the data. However, the order of magnitude of the numbers remains the same.
justify construction and operation of the SARR. The line slopes up because of the inflation rate of reproduction cost. These capital charges are calculated by discounting a single most likely projection. Because of competition and contestable markets, the SARR’s chances of performing better than this projection are constrained. The chances of performing worse are not because the SARR’s investment would be sunk and irreversible. Therefore, no rational investor would commit to build a SARR based on this projection of capital charges shown by the dashed line in figure 4. In other words, the lower line cannot define the entry point for a SARR in contestable markets.

Investment in the SARR would make sense only if projected capital charges were on the higher, solid line in figure 4. The problem is to calculate the numerical difference between these two lines. Since the lower line can be calculated from the usual application of the STB model, we take it as a fixed starting point. In other words, we follow the STB model in every respect except for the adjustment for asymmetric risk.

According to the initial model, the initial capital charges required for the SARR are approximately $500 million. This amount would be the starting point for the dashed line in figure 4. Suppose the SARR enters at this point with its upside constrained by the potential entry of SARR2. The constraint is represented by the solid line in figure 4. Suppose as an example that the solid line is 22.6% higher than above the capital charges

---

28 The line for the SARR slopes up because of growth in volume.
calculated by the STB model says would be adequate that is, 22.6% above the dashed line. With a 20% annual standard deviation of the stochastic process, the SARR has a negative NPV of $788 million. Therefore no rational investor would build and operate the SARR at revenues which gave an initial capital charge of only $500 million. Entry would not occur unless revenues and initial capital charges were significantly higher.

In fact they would have to be 40% higher. In other words, a SARR would not enter unless it could start life on the upper line in figure 4. We calculate the height of the upper line by solving for the initial capital charges that would generate a NPV of zero with asymmetric risk and returns taken into account. The SARR is economically justified only if actual capital charges begin at $702 million.

4.6. More Numerical Results

Table 2 reports the implications of asymmetric risk for the SARR. The second line of the table summarizes the example just given. At a standard deviation of 20%, an initial capital charges of $702 million is required to justify investment in the SARR, an increase of about 40% over the capital charges of $500 million that the STB models says would be required.

These results were obtained by simulating the SARR’s capital charges, taking account of the range of possible outcomes, given the standard deviation chosen. The constraint caused by the potential of future entry takes the required sunk cost investment of that future entry into account. Thus, the potential new entrant that arises from contestability theory is in the same position as the SARR1, but this threat of entry constrains the SARR1. If SARR1 attempts to charge too high prices, under contestability theory SARR2 will enter and take away the business from SARR1.

Table 2 presents results for standard deviations from 15% to 25%. Notice that moving

<table>
<thead>
<tr>
<th>Table 2. Impact of Asymmetric Risk on Required Capital Charges</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard Deviation</td>
</tr>
<tr>
<td>--------------------</td>
</tr>
<tr>
<td>Absorbing barrier</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Reflecting barrier</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
from a lower to a higher standard deviation increases the required initial capital recovery cost. The greater the risk, the more asymmetric risk matters.

There are two blocks of numbers in table 2. The top block follows the theory of contestable markets exactly. It assumes that the SARR’s upside is cut off by the threat of entry by a future SARR. Refer again to figure 4, and assume a scenario in which the SARR could potentially do exceptionally well. The calculations for contestable markets recognize this potential, but allow the SARR to earn no more than the upper (solid) line. If the potential earnings are above that line for several periods, then the SARR’s actual earnings will stick on the solid line for each of those periods.29 This “absorbing barrier” states that regulation places an upper limit on the earnings of the railroad. If earnings ever exceed this amount a shipper will use STB regulation to decrease rates to the absorbing barrier. Thus, STB regulation is a form of “maximum rate regulation” that truncates the distribution of return for the real world railroad. However, the truncation point does not take account of the sunk cost nature of most railroad investment. Instead, it assumes that the world is contestable so that all investment is fixed but not sunk.30 Thus, the truncation point is downward biased compared to what it should be according to a real options calculation. Modern real options theory demonstrates that the expected return for a sunk investment must exceed that of fixed investment by a factor typically around 2.0 or higher.31

We then consider another possible competitive outcome. The bottom block of numbers assumes that the constraint on profits comes not from a single potential entrant, but from many competitors. Again consider a scenario in which the SARR does very well. As its earnings hit the top line, competitive entry by other firms tends to push earnings back down again.32 After competitive entry has occurred prices will tend to decrease below the truncation point as the multiple entrants compete with each other. The odds that the SARR’s earnings can stick at the upper limit for several periods in a row are therefore reduced. Thus, the truncation point creates a “reflecting barrier,” which provides an upper bound to the return, but the competitive entry that occurs typically causes return to decrease below this point.33 Of course this makes average earnings lower, relative to starting capital charges. The required initial capital charges (the initial operating cash flow just sufficient to justify construction of the SARR) are correspondingly higher. This is reflected in the bottom block of results in table 2. With a 20% standard deviation, the required initial capital recovery would increase by 66.1% from about $500 million to $830 million.

29 Thus, we are applying an “absorbing barrier” solution to the problem. See, for example, D. Cox and H. Miller, The Theory of Stochastic Processes, London, 1965 for a discussion of absorbing barriers.
30 Fixed investment can be redeployed immediately in other uses if returns are below the competitive level while sunk investment cannot be redeployed.
32 Here we are applying a “reflecting barrier” solution because competition will decrease profits below the maximum. See, for example, Cox and Miller op. cit. for a discussion of reflecting barriers.
33 For a further discussion of reflecting barriers see Dixit and Pindyck, op. cit. p. 254ff, where they discuss dynamic equilibrium in a competitive industry which causes this type of behavior. Rational expectation by competing firms on their entry decision leads to this type of outcome.
Thus, the most straightforward way to allow for uncertainty and asymmetric risk, by assuming that profits never exceed the amount to cause new entry, demonstrates that the STB’s failure to account for these factors leads to an understatement of the required initial capital recover of approximately 30%-50%. Using a more realistic model of competitive entry which will force profits below the maximum demonstrates that the STB’s failure to account for the factors of uncertainty and asymmetric risk leads to an understatement of the required initial capital recover of approximately 50%-84%. These amounts are significant economic factors in choosing an economically correct approach to setting regulated prices for railroads.

5. Effects of STB Regulation

The STB in a recent decision has rejected the necessity of taking sunk costs into account, along with the associated asymmetric risk, in terms of correct regulation. The STB stated:

To the extent UP may face some (more limited) asymmetric earnings risk itself, as its counsel suggests, UP has not shown why that risk is not already reflected in its cost of capital. We presume efficient capital markets recognize and reflect all of the risks faced by railroads... (p. 53, footnote omitted)

The STB has failed to recognize that cost of capital that a railroad faces is composed of many investments that regulation cannot affect because the R/VC ratio is below the statutory limit and also investments that regulation does affect as we have discussed above. Since STB regulation truncates the distribution of returns by failing to account for the asymmetric risk it creates, the distribution of returns differs for regulated and unregulated investments ad demonstrated in figures 2a–2c. Modern finance theory recognizes that different projects have different potential risk and return characteristics and insists that, to the extent possible, each individual project should be analyzed in terms of its specific characteristics. The STB has failed to consider this important distinction.

We now consider whether this incorrect regulation matters. In establishing its Standards for Railroad Revenue Adequacy in 1986, the ICC defined and explained the concept of revenue adequacy:

...[T]he concept [of revenue adequacy] itself is forward looking, the focus being the railroads’ ongoing revenue requirement to sustain a viable plant investment over the long run.

This approach is consistent with the economic approach to regulation that concludes that regulated firms must have the opportunity to achieve their cost of capital. If regulated firms

34 STB Docket 42051, September 13, 2001, Wisconsin Power and Light v. Union Pacific Railroad.
35 Ex Parte No. 393 (Sub-No.1), decided December 16, 1986, p. 297.
are unable to achieve their cost of capital over reasonable time periods, investors will not provide the funds for new investment and for improvements and replacement of obsolete capital. The result will be deterioration in service and economic inefficiency. However, despite the passage of 15 years and unprecedented economic growth and prosperity in the United States economy over approximately the past six years, the railroads have not become revenue adequate.

We now consider the STB’s revenue adequacy determination for the past 11 years for the three major Class I railroads.36

The STB’s own determination over the 12-year period 1989–2000 has found that the BNSF and UP were revenue adequate in only one year each: the BNSF in 1989 and the UP in 1995. The CSX has never been revenue adequate. The STB estimated industry cost of capital has been quite steady over the period. Thus, looking at the average is a good summary statistic. Over the 12-year period the BNSF has reached 76.5% of average revenue adequacy, the CSX has achieved 48.7% of average revenue adequacy, and the UP has achieved 70.4% of average revenue adequacy. Most importantly, the gap between the cost of capital and the individual railroad’s return to capital has been increasing over the past 4 years, even when merger related difficulties are considered. Thus, the “revenue adequacy gap” has been growing over time. Even in the best sustained period of U.S. growth in 30 years, over the 1997–2000 period, none of the three major U.S. railroads was revenue adequate, according to the STB’s own calculations. This continuing revenue inadequacy is consistent with the STB’s form of regulation that fails to recognize the importance of sunk costs and asymmetric risk, which is a fundamental economic characteristic of railroad investment.

36 In earlier years I used the predecessor railroad before merger, e.g., I used the BN where in later years I use the BNSF. This procedure actually overstates return to capital because typically the railroads that were merged into the remaining railroads had inferior economic performance. Also, where the return was negative in two instances, I used a zero return.
When regulation does not provide correct regulatory incentives for investment, problems occur. In the United Kingdom the British government vertically dis-integrated British Rail into two components. Railtrack, a regulated monopoly provider of trackage service, was formed, as were a number of separate passenger and freight operator companies that would buy access service from Railtrack.\textsuperscript{37} A regulatory authority established cost-based prices for Railtrack in a manner similar to the STB that failed to recognized the sunk cost component of railroad investment.\textsuperscript{38}

Initially, Railtrack performed well with a successful IPO. Up until early 2001, its stock had traded in a range of £10–£12, well above its initial post-privatization price of £3.90. However, capital markets came to realize that Railtrack was not receiving adequate compensation from its regulated rates. Railtrack became unable to raise the necessary funds to maintain and modernize its network. The head of Britain’s Office of the Rail Regulator, called the Regulator, belatedly recognized the absence of economic incentives for Railtrack to invest and has stated that ‘‘[e]veryone accepts that the railway industry has been starved of long-term investment.’’\textsuperscript{39} The Office of the Rail Regulator identified two specific concerns about track quality: a downward trend in the quality of track, and a significant increase in the incidence of broken rails over the previous two years.\textsuperscript{40} However, Railtrack was unable to raise the necessary capital to do the required investment. Railtrack was put into administration (forced into bankruptcy) by the United Kingdom government in October 2001.\textsuperscript{41} Incorrect regulation did not permit Railtrack to modernize adequately, even though the need for investment received widespread agreement from Railtrack management, the Rail Regulator, and most of the users of Railtrack.

6. Conclusion

Application of real options analysis in the SAC approached used by railroad regulation in the United States is required to retain coherence and consistency with the real world of railroad investment and the application of contestability theory to railroad regulation. Ignoring the modern economic theory of investment under uncertainty, as the STB model does, would assume no barriers to entry and exit in the current period, but would assume absolute barriers to entry and exit for subsequent SARRs such as SARR2, SARR3, and so

\begin{itemize}
\item[37] Thus, Railtrack had a higher percentage of sunk cost investments than U.S. railroads because Railtrack did not operate rolling stock.
\item[38] An incentive component was also included, but the cost component did not recognize the existence of sunk costs. An important difference is that U.K. regulation set all of Railtrack’s tariffs by incorrect regulation, while U.S. STB regulation only affects U.S. railroad tariffs above the statutory R/VC ratio.\textsuperscript{39}
\item[39] Tom Windsor, ‘‘Creating an Investment-Friendly Environment,’’ July 11, 2001. Mr. Windsor was the government regulator of Railtrack.
\item[40] See Railtrack’s Stewardship of the Network, §4.1 (Track quality), 1999. The report stated that it ‘‘would appear to indicate that Railtrack’s maintenance and renewal work is not keeping pace with the aging of the network and the growing volumes of traffic which have led to increased wear and tear.’’
\item[41] See, for example, Juliette Jowit and Nicholas Timmins, ‘‘Railtrack collapses as Byers pulls plug’’, Financial Times, October 8, 2001. See www.ft.com for an extended account of the demise of Railtrack.
\end{itemize}
on. The result would be an inconsistency in the application of contestability theory to railroads. In addition, the market realities of uncertainty would also be ignored.

As we demonstrate, the truncation effect that creates the fundamental asymmetry in potential future outcomes in a contestable market leads to lower expected returns for a SARR. This requires higher rates for SARR1 so that it can earn the appropriate economic return to its sunk investments under conditions of uncertainty. The STB’s simulated competitive rate benchmark is too low because it fails to account for the sunk and irreversible nature of much investment in railroads. We find that failure to take account of uncertainty and asymmetric risk in railroad regulation has a quantitatively important effect. We find that the required returns calculated from the STB model that ignores these factors is too low by between 30% and 84.4%. Thus regulation by truncating the return of investment by the railroads will force investment below economically efficient levels, so that too little investment will be made in situations where regulation of railroad prices can occur. Overall, our results demonstrate the importance of taking account of sunk costs in designing an appropriate regulatory framework and how the application of the real options framework allows for regulation with sunk costs present.

We lastly note that according to the STB’s own calculations almost no large railroads have earned their cost of capital over the past 12 years. This outcome is consistent with STB regulation that fails to account for sunk cost investment in setting regulated rates. A similar situation in the United Kingdom has led recently to the regulated railroad trackage service provider going into bankruptcy.

Technical Appendix

The STB model as used to determine rate reasonableness has two components: (1) the STB model solves for the initial level of capital charges that makes the present value of after-tax cash flows to investors equal to the required investment. Thus the NPV of the SARR from this stages equals zero. (2) the present value of the traffic revenues is compared to the present value of SAC. SAC is calculated by adding operating expenses to the capital charges. If revenues exceed SAC then the rates are deemed too high and vice-versa.

The real options approach uses the first component of the STB model. Tax calculations and the terminal value calculations are the same. The real options model (ROM) adds three additional components: (1) uncertainty is introduced into the capital charge (2) the uncertain capital charges are capped by the threat of entry (3) a multiplier to the capital charge is introduced. The multiplier is determined by setting the expected net present value to zero. The adjusted capital charge is then applied in the second stage of the STB model, as discussed above.

1. Introducing Uncertainty

The options pricing model requires that all quantities by put into a “risk-neutral” equivalent basis. Let the cost of capital \( r \) be the appropriate discount rate and the risk free rate be \( r_f \). The conversion factor is:
\[ a_t = \left( \frac{1 + r_t}{1 + r} \right)^t \]  

(A1)

Capital charges, interest tax shields, and depreciation are all multiplied by this factor to find the risk adjusted quantities. A random factor in then introduced using a geometric random walk. Each period has its own independent random draw \( \eta_t \) that is distributed as log normal with mean one and standard deviation \( \sigma \). Thus in each period the random variable takes the form:

\[ \tilde{\epsilon}_t = \tilde{\epsilon}_{t-1} \eta_t, \]

(A2)

so that capital charges have the realization:

\[ \tilde{C}_t = a_t C_t \tilde{\epsilon}_t, \]

(A3)

where \( C_t \) is the (non-stochastic) capital charge in the STB model.

2. The Cap Introduced by SARR2

Two different barriers exist in the ROM, the absorbing barrier and the reflecting barrier. The limit or cap \( X_t \) is determined by the process:

\[ X_0 = C_0 \ldots X_t = X_{t-1} (1 + i_t), \]

(A4)

where \( i_t \) is the asset inflation rate. The absorbing barrier takes the form

\[ \tilde{C}_t = \min(X_t, a_t \tilde{C}_t), \]

(A5)

where \( \tilde{C}_t \) is the capped realization of the random variable. For the reflecting barrier the random factor is capped:

\[ \tilde{\epsilon}_t = \min \left( \tilde{\epsilon}_{t-1} \eta_{t-1} \frac{X_t}{a_t \tilde{C}_t} \right), \]

\[ \tilde{C}_t = C_t \tilde{\epsilon}_t. \]

(A6)

3. Solving for the Multiplier

Introducing the cap will cause the NPV to be negative using the (non-stochastic) STB capital stream of \( C_t \). We calculate a multiplier, \( m > 1 \), used in the tables, to return the NPV to zero. We solve for the multiplier \( m \) that sets the NPV to zero. For instance, for the absorbing barrier we find:

---

42 Our approach allows for other stochastic processes, e.g., mean reversion. The random draws are changed accordingly.
\[ \hat{C}_t = mC_t \tilde{z}_t, \]
\[ \hat{C}_t = \min(X_t, a_t \tilde{C}_t). \]  \hspace{1cm} (A7)

For the reflecting barrier the analogous equation is
\[ \hat{C}_t = ma_t C_t \tilde{z}_t, \]  \hspace{1cm} (A8)

where equation (A6) is used to determine \( \tilde{z}_t \). A large number, 10,000, of random factor sequences \( \eta_t \) are drawn using a Monte Carlo technique.\(^{43}\) For each sequence the random factor \( \xi \) is applied to capital charges and the resulting NPV is calculated. The multiplier \( m > 1 \) is calculated that sets the average NPV to zero for the 10,000 Monte Carlo outcomes. Thus, the expected outcome, not any individual outcome, determines the value of the multiplier \( m \). The estimated multiplier is then applied to the stream of capital charges to generate a new stream of adjusted capital charges and the SAC is calculated. As with the STB model, the revenues of the SARR are then compared to the SAC to determine whether current rates are reasonable.

References


\(^{43}\) See, for example, Press (1992), pp. 305–306 for the Monte Carlo technique used.