



New and Larger Costs of Monopoly and Tariffs

Careful examination of industries in transition finds that both monopoly and tariffs generate significant costs

James A. Schmitz, Jr.*

Federal Reserve Bank of Minneapolis

Introduction

In standard economic theory, monopoly leads to a welfare loss. This loss stems from a misallocation of resources *across* industries: Too few goods are produced by the monopolist; too many in other industries. Economic theory had long suggested that this welfare loss exacted high costs from the economy. But modern understanding took a turn when, in a landmark 1954 paper,¹ Arnold Harberger analyzed the quantitative significance of monopoly costs in the United States. Were these costs as high as conventional economic theory suggested? The clear but surprising answer that Harberger provided was no.

Harberger estimated that, contrary to his expectation and to standard theory, the costs of monopoly were quite trivial. “We come to the conclusion that monopoly misallocations entail a welfare loss of no more than a thirteenth of a per cent of the national income. Or, in present values, no more than about \$1.40 per capita,” he wrote. “I must confess that I was amazed at this result. ... Monopoly does not seem to affect aggregate welfare very seriously through its effect on resource allocation” (Harberger 1954, pp. 85, 86, 87).

Other economists extended Harberger’s work to estimate costs associated with tariffs, and here, too, the costs were trivial. A consensus quickly developed that Harberger’s conclusion was indeed valid.

Recently, a new literature has taken a different approach to understanding the costs of monopoly.

* The author thanks Tasso Adamopoulos, John Asker, Doug Clement, Tom Holmes, Pete Klenow, Sam Kortum, David Lagakos, Justin Pierce, Todd Schoellman and Arilton Teixeira. A more complete version of this policy paper is available online at minneapolisfed.org.

Economic Policy Papers are based on policy-oriented research by Minneapolis Fed staff and consultants. The papers are an occasional series for a general audience. Views expressed are those of the authors, not necessarily of others in the Federal Reserve System.

ABSTRACT

Fifty-eight years ago, Arnold Harberger estimated that the costs of monopoly, which resulted from misallocation of resources across industries, were trivial. Others showed that the same was true for tariffs. This research soon led to the consensus that monopoly costs are of little significance—a consensus that persists to this day.

This paper reports on a new literature that takes a different approach to the costs of monopoly. It examines the costs of monopoly and tariffs within industries. In particular, it examines the histories of industries in which a monopoly is destroyed (or tariffs greatly reduced) and the industry transitions quickly from monopoly to competition. If there are costs of monopoly and high tariffs within industries, it should be possible to see those costs whittled away as the monopoly is destroyed.

In contrast to the prevailing consensus, this new research has identified significant costs of monopoly. Monopoly (and high tariffs) is shown to significantly lower productivity within establishments. It also leads to misallocation within industries: Resources are transferred from high- to low-productivity establishments.

From these histories, a common theme (or theory) emerges as to why monopoly is costly. When a monopoly is created, “rents” are created. Conflict emerges among shareholders, managers and employees of the monopoly as they negotiate how to divide these rents. Mechanisms are set up to split the rents. These mechanisms are often means to reduce competition among members of the monopoly. Although the mechanisms divide rents, they also destroy them (by leading to low productivity and misallocation).



Looking *within* industries, it examines the histories of industries in which a monopoly is destroyed and the industry transitions quickly from monopoly to competition, as well as the histories of industries that rapidly moved the opposite way, from competition to monopoly. If there are costs of monopoly, those costs should be whittled away as the monopoly is destroyed. Likewise, if an industry is monopolized, costs should be created. In both cases, costs should be apparent when comparing the industry before and after monopolization.

Several industries have been studied with this method, including transportation in the United States and U.S. manufacturing of sugar, iron ore and cement. The historical records of these disparate industries show that there are costs of monopoly and tariffs within industries. In these industries, this new literature has shown that monopoly led to, among other costs, the following:

1. *Low productivity* at each factory. That is, for any given amount of inputs, monopoly meant that less output was produced than under competition.

2. *Misallocation of resources between high- and low-productivity factories.* That is, monopoly led to resources (capital, labor, etc.) being transferred from productive factories to unproductive factories. Again, this misallocation occurs *within* an industry and is different from the misallocation that Harberger studied.

In sharp contrast to Harberger's finding, these studies show that the welfare costs associated with monopoly and tariffs are not small. The consequence of cases (1) and (2) above is that industry output could have been produced with fewer inputs. One way to measure the loss, then, is to calculate the value of the "wasted" inputs. The histories of these industries show that as monopoly was destroyed in each, productivity at each factory soared. Doubling of productivities in a few years was common. The value of the wasted inputs was as much as 20 percent to 30 percent of industry value added.

A common theme (or theory) emerges from the histories as to why monopoly led to these costs. When a monopoly is created, "rents" are created.

(In this usage, "rent" is the difference between what a factor of production is *actually* paid and what it would *need* to be paid to remain in use; as such, it is a measure of that factor's monopoly power.) Conflict emerges among shareholders, managers and employees of the monopoly as they negotiate how to divide these rents among themselves—or, more colloquially, how to "split the spoils." Mechanisms are set up to split the rents. Although they divide rents, they also destroy them (by leading to low productivity and to misallocation).

As used in this paper, the term "monopoly" means more than the strict definition: an industry with a single producer. One industry mentioned later in the paper was a cartel for 40 years. Conflict over rents emerged between groups in the cartel, firms, workers and managers. In some industries, there were high tariffs (and other forms of protection). This high protection led to strong incentives among groups in the domestic industry to form monopolies. Firms attempted to collude, and workers formed industrywide unions (i.e., monopolies). So, the statement that "tariffs led to large welfare losses" means that tariffs led to incentives to form monopolies and then to actual monopolies, and these monopolies then led to large welfare losses.

A body of literature in the 1960s and 1970s argued that the costs of monopoly and tariffs were not trivial, saying (in essence) that there were costs within industries. This theoretical literature, and why it did little to dent the "Harberger consensus," is briefly reviewed in Minneapolis Fed Staff Report 468 (online at minneapolisfed.org), on which this policy paper is based. In this policy paper, I discuss historical studies that look at the collapse of monopoly. I describe how the monopolies emerged and how they were destroyed. Then I discuss the mechanisms that were used to split rents and why these mechanisms led to welfare losses.

Monopoly: Its Creation and Destruction

When a monopoly is created, the government often has a hand in the process. This is the case in most of the industries studied, to greater or lesser degrees. In U.S. sugar manufacturing, the government played a central role in creating monopoly. During the Great Depression, sugar manufacturers were permitted, indeed encouraged, by U.S. law to form a cartel.



Jim Schmitz

PHOTOGRAPH BY MARC NORBERG

Many U.S. cartels were created during the Depression (as part of the New Deal), but the New Deal sugar cartel survived much longer than most. For 40 years, from 1934 to 1974, the industry was repeatedly able to renew the U.S. laws that enabled it to operate as a cartel. Soaring world sugar prices in 1974 resulted in the cartel losing political support, and the laws permitting it to operate as a cartel were not renewed.

To describe the government's role in creating monopoly in the other industries, a useful approach is to first sketch a very simple model. Consider an industry where transportation costs are large relative to production costs. If the domestic price is initially set equal to the cost of domestic production, then domestic producers will have a strong incentive to push their price up to the sum of foreign production cost plus the cost of transportation (or tariff) involved in bringing the foreign product to domestic markets.

The incentive to do so is great in this "industry" because, by assumption, transport costs are large relative to production costs. A very large tariff will be an incentive to increase prices, just as a large transportation cost would.

If the transportation or tariff cost is large, assume that groups will make investments to form monop-

olies. Firms will attempt to collude, and workers to form strong unions. Some groups may succeed. If later on protection is cut, the incentives to make these investments will fall, and the monopolies will weaken (or disappear). This same logic applies if, rather than a transportation cost advantage, local firms have a production cost advantage.

This simple abstraction is a good representation of both the iron ore and cement manufacturing industries. In the early 1950s, U.S. producers had production cost advantages over foreign producers, and the industries received significant protection. Groups invested in creating monopolies. At various times, firms in these industries were charged with trying to collude. The U.S. government investigated the industries for antitrust violations. It is unnecessary to enter into the argument as to how effective collusion was; there is little doubt that very strong, industrywide unions emerged in these industries. Although antitrust laws in the United States made firm collusion difficult, building monopoly unions was easier. Collective bargaining laws enacted by the U.S. government allowed unions to organize all workers in an industry and not be bound by antitrust laws.

The monopolies in these industries—in particular, the strong monopoly unions—lasted for many decades in the post-World War II period. The monopoly unions were able to provide very high wages. For example, by the 1970s, cement workers were paid as much as U.S. autoworkers (who were the highest-paid manufacturing workers). The unions also had very stringent work rules (as described later on).

In the 1980s, the monopolies in these industries weakened or were dissolved. The union in the cement industry dissolved. In the iron ore industry, the union did not disappear, but lost much of its clout. For example, work rules became much less stringent, and plant managers had more control over how to structure plant operations.

Why the weakening of the monopolies in the 1980s? Foreign producers were now threatening to enter local markets. Brazil offered to sell iron ore in Chicago and Cleveland, the heart of the U.S. market, at half the local price. Firms around the world offered to sell cement on the West Coast and Gulf of Mexico at half the U.S. prices.

How could foreign firms offer to sell at such discounts? There are two proximate reasons. First, transportation costs greatly decreased (relative to production costs) in the postwar period. This, by itself, would have meant a weakened incentive for continued investment in keeping monopoly. But, second, the production cost advantage of U.S. producers decreased. This development was, of course, to be expected, as the whole purpose of creating strong unions was to increase wages (and hence costs). The monopolies also led to lower productivity, increasing costs further. But what was striking is that U.S. producers were at a production cost disadvantage.

An obvious question is: Why did the unions (and other groups discussed later on) push wages so high and lower productivity to the point where foreign producers could offer such steep discounts? At least three possibilities come to mind. First, the groups realized that wage demands and work rules would lead to the demise of monopoly, but that this strategy was the best. Second, the groups realized that wage demands and work rules would lead to the possibility of foreign entry, but they expected more government protection than they were able to receive. Many calls for protection were made, and some protection was given, but it was not enough. Third, perhaps the outcome (foreign entry) was not expected. Although I do not know which story best describes the events, the story itself is not important for the issue at hand. The main point is that there are significant costs of monopoly and tariffs.

Monopoly: Splitting the Spoils (and Destroying Them as Well)

During the period when monopolies in these industries were strong, groups set up mechanisms to split rents. Here I discuss some of the mechanisms used and how they led to the destruction of rents—in particular, to low productivity and misallocation. When monopoly was weakened in these industries, the mechanisms were abandoned, leading to large productivity gains in establishments and to resources being reallocated from low- to high-productivity producers.

One mechanism used to split rents was competition-reducing rules. Here I discuss two types that were used: *quotas* and *work rules*.

Quotas

In the U.S. sugar industry, the New Deal cartel included factory owners, factory workers, farmers, farm workers and others. As the cartel was established, each of these groups sought to secure (for themselves) as large a share of rents as possible. A major mechanism to split rents was quotas. In the cartel, firms were given quotas—the right to sell a certain amount of sugar each year.

Incumbent farmers also sought, and were successful in acquiring, quotas—the right to grow sugar beet crops on a given number of acres each year. Without these quotas for incumbent farmers, nothing stopped firms from moving the locations of their factories or even using different farmers in the same location. Just as firms in the cartel used firm quotas to limit competition, incumbent farmers wanted quotas to limit competition among themselves (and from other farmers). Without these quotas, there was no way to ensure that incumbent farmers would receive a share of the monopoly profits.

As is often the case, these quota rights (both those of the firms and those of the farmers) could not be sold. Although the allocation of quotas for acres in 1934 was “efficient,” over time there was a change in the comparative advantage of locations in manufacturing sugar. Hence, there emerged a significant misallocation of resources between factories, with low-productivity factories producing too much sugar and high-productivity factories too little.

I can estimate the magnitude of the welfare loss due to these mechanisms to split rents (the quotas), that is, from the misallocation of resources within the industry. Recall the introduction to this paper, which mentioned that one way to measure welfare loss is to calculate the value of wasted inputs in producing industry output. The estimates indicate that the losses were roughly 20 percent to 30 percent of industry profits.

Work Rules

In the iron ore and cement industries, those who were in a position to gain from the large transportation costs into local markets, and the protection offered by tariffs, were the factory owners, factory workers and even the local governments (e.g., townships) where factories were located. What mechanisms were used to acquire rents? Local

townships placed significant taxes on the production of iron ore and cement. Workers formed very strong unions. Although claims of collusion within both industries have been made, these claims are harder to document than the taxes and union contracts that emerged in these industries.

A major mechanism to split rents was the work rules in union contracts. Among other things, work rules were a way to limit competition among workers. They were structured so that managers could not play workers off each other.

Union contracts split the tasks in plants into groups or categories. Workers were then assigned to one of these groups or categories, that is, given the right to complete tasks in that category. Only the workers in this group could complete the tasks assigned to the group. Very often these distinctions among workers were arbitrary in that a worker in a particular category was able, but not allowed, to complete tasks in many other categories.

These types of work rules dividing work among members of the union are most often called *job classification systems*. They are similar to the quotas discussed earlier. In particular, work rules are a way to limit competition between workers, just as quotas limited competition between farmers. They ensure that groups of workers receive a share of the monopoly profits. But they also destroy profit.

What is the quantitative significance of work rules? In the 1980s, when the work rules in the iron ore and cement industries were made much less stringent, labor productivity doubled in a few years. Other productivities increased as well. If these increases in productivity can be tied in large part to the relaxing of work rules, then obviously these are big welfare gains.

We can estimate the magnitude of this welfare loss due to these mechanisms to split rents (the work rules), that is, from the low productivity in establishments. Again, one way to measure the welfare loss is by the value of the wasted inputs. With these work rules, machines were down longer than necessary. The energy that was being consumed elsewhere in the plant when output was not produced was a wasted input. Capital was also wasted, as work rules meant that disabled machinery took longer to repair than was necessary.

Labor input was wasted as well. For example, a

machine operator could not hold a tool for a repair person (who would need to bring in another repair person for such tasks). The value of this wasted input was the opportunity cost of the machine worker's time multiplied by the amount of time involved. A rough estimate suggests a dead-weight-loss-to-industry-value-added ratio of 16 percent to 17 percent. (See Staff Report 468, pp. 14-15, for details of this calculation.)

Using a dead-weight loss for the wasted capital and energy of a few percentage points (possibly more) of value added, together with the wasted labor estimate of 16 percent to 17 percent of value added, gives an estimate of over 20 percent in total.

Splitting the Spoils: Other Industries, Other Countries and a U.S. Cost Estimate

In other industries as well those just discussed, and in other countries, work rules have likely led to the same type of misallocation—with low-productivity plants producing too much output and high-productivity plants too little. However, I cannot be sure of their quantitative significance because no studies like those described in the preceding section have been completed for these industries.

Many U.S. industries had significant market power after World War II, first by virtue of the devastation that many countries faced as a result of the war and later because of government protection of U.S. manufacturing. Monopolies emerged; in particular, the postwar years saw the emergence of industrywide unions in the auto, steel, paper, tire, airplane and chemical industries, to name a few.

What mechanisms were used to split rents? The job classification systems discussed earlier are prevalent throughout manufacturing (though for the most part are less stringent today than a few decades ago). Some observers of these industries hold the view that work rules led to low productivity in plants.

Stringent work rules likely led to low productivity in establishments in many manufacturing industries. In some, they led to other types of distortions and losses not seen in the cement and iron ore industries. As I suggested earlier, work rules in these industries likely led to misallocation—resources being transferred from high- to low-productivity plants. High wages (and stringent work rules) have likely led to another type of misallocation in indus-

tries: a change of technology (in order to escape the wages and work rules).

A similar phenomenon—that is, monopolists splitting (and destroying) rents—occurs in other countries. In Britain, job classification systems (referred to as “job demarcation rules”) are widespread. Demarcation rules are also used in France. In both countries, research suggests that these rules lead to reduced productivity.

I finish this section with a back-of-the-envelope estimate for the within-industry costs of monopoly and tariffs for the United States. This will enable a preliminary stab at the question, are these welfare losses similar in magnitude to Harberger’s losses (0.1 percent of value added), or can I conclude that they may well be significantly larger?

Industries that are known to have strong unions and rigid work rules include mining, utilities, construction, transportation (in particular, airlines and railroads) and parts of manufacturing, in particular, durable manufacturing (steel, airplanes, autos). Assume that work rules had similar negative impacts on productivity in those industries as they did on the industries discussed in detail earlier—again, about 20 percent of industry value added.

Adding together the total value added of these industries thus affected (just over 25 percent of total gross domestic product in 1977) enables an estimate of welfare losses from monopolies and tariffs of roughly 5 percent of GDP (=20 percent loss of 25 percent GDP share). (Further calculation details are in Staff Report 468.) Again, this calculation is obviously extremely crude, but it does suggest that the losses may well be orders of magnitude larger than Harberger’s estimated losses.

Costs of Monopoly: Summary and Observations

Research on the theoretical and quantitative significance of monopoly costs has evolved considerably since the mid-1950s, when Harberger’s influential paper suggested—in contrast to the prevailing view among economists—that in the United States, the costs of monopoly resulting from resource misallocation across industries were actually quite insignificant.

This paper reviews a new stream of research that uses a different approach to analyzing the costs of monopoly. It examines the costs of monopoly and tariffs *within* industries rather than across them. In

particular, it examines the histories of industries in which a monopoly is destroyed (or tariffs greatly reduced) and the industry transitions quickly from monopoly to competition.

Over considerable time spans and a wide range of industries, this research finds that monopoly exacts high costs in two ways: (1) through misallocation of economic resources between high- and low-productivity factories and (2) by decreased productivity at each factory. The historical studies call the Harberger consensus into question. At least in the industries studied thus far, monopoly and tariffs have led to significant welfare losses, on the order of 20 percent of value added.

A common thread runs through these histories, one that suggests a theory. When a monopoly is created, rents are generated. But the distribution of these rents—splitting the spoils—causes conflict among shareholders, managers and employees of the monopoly. These parties devise mechanisms to split the spoils, but the mechanisms often lead, paradoxically, to the destruction of rents.

The implications of this theory of monopoly costs, and of the empirical findings of high costs, deserve serious consideration by policymakers as they design and enforce antitrust measures. Government policies themselves, such as tariffs and other forms of protection, are an important source of monopoly. This review of recent research indicates that the costs of such protectionist policies are considerable and should be fully recognized and appreciated. Furthermore, policy reforms to minimize these costs should be carefully considered.

As for future economic research, a key question is to understand why mechanisms (such as work rules) are used to split rents when they also self-destructively wipe out rents. Why can’t members of the monopoly structure contracts that avoid such large wasted resources? Differences in information? The inability of parties to commit to future actions? Such reasons may well be why mechanisms intended to split rents also destroy them. **R**

¹ Harberger, Arnold C. 1954. “Monopoly and Resource Allocation.” *American Economic Review* 44(2): 77–87.