



# Short-Run Costs and Postal Pricing

**January 9, 2013**

**Prepared by U.S. Postal Service Office of Inspector General  
Report Number: RARC-WP-13-004**

# Short-Run Costs and Postal Pricing

## Executive Summary

The Postal Accountability and Enhancement Act (PAEA) altered the previous relationship between costing and pricing at the U.S. Postal Service, but it still maintained the requirement that each product cover its attributable cost.<sup>1</sup> Therefore, the Postal Service's estimate of attributable cost for each product could be considered its price floor. The current product cost system is based on estimates of long-run costs; however, recently, there have been suggestions that the Postal Service could consider pricing at least some of its products to cover short-run costs. The U.S. Postal Service Office of Inspector General (OIG) briefly discussed this issue in its white paper, *A Primer on Postal Costing Issues*.<sup>2</sup>

The OIG asked Professor Michael D. Bradley of the Economics Department of George Washington University, an expert in postal economics, to co-author this paper on pricing and short-run costs. This paper defines what is meant by short-run and long-run costs, explores the issues associated with using short-run costs when developing prices, outlines what information is needed to measure short-run costs, and develops a multistep algorithm for estimating short-run costs that is consistent with the existing Postal Service cost system. It first focuses on explaining why short-run costs arise and how they can be used in pricing, in general terms, and then investigates their applications for the Postal Service. Developing a specific recommendation regarding the Postal Service using short-run costs to develop prices is beyond the scope of this paper.

The distinction between long run and short run for a firm is defined by its ability to adjust its inputs, mainly labor and physical capital. If a firm is able to adjust inputs to their optimal levels, it is considered to be operating in the long run. If a firm cannot adjust its inputs to an optimal level, it is considered to be operating in the short run.

The Postal Service may be operating under short-run conditions and therefore may want to consider setting prices for at least some of its products using short-run costs. Under the proper conditions, this approach could help mitigate the costs of excess

---

<sup>1</sup> Public Law 109-435, Postal Accountability and Enhancement Act (120 Stat. 3205), requires each competitive product to cover its direct and indirect costs. Public Law 109-435, Postal Accountability and Enhancement Act (120 Stat. 3201), includes a factor that each class or type of mail must bear the direct and indirect postal costs attributable to it. While it has been acknowledged that the price cap requirement trumps this factor, the Postal Regulatory Commission has stated that the Postal Service must demonstrate that it is working towards setting prices that cover each product's costs. See the discussion of this issue starting on page 13 of the fiscal year (FY) 2010 Postal Regulatory Commission's Annual Compliance Determination Report, [http://www.prc.gov/Docs/72/72382/PRC\\_ACD\\_2010.pdf](http://www.prc.gov/Docs/72/72382/PRC_ACD_2010.pdf).

<sup>2</sup> U.S. Postal Service Office of Inspector General, *A Primer on Postal Costing Issues*, Report No. RARC-WP-12-008, March 20, 2012, [http://www.uspsoig.gov/foia\\_files/RARC-WP-12-008.pdf](http://www.uspsoig.gov/foia_files/RARC-WP-12-008.pdf).

capacity by increasing volume through lower prices. However, the improper use of short-run costs to develop prices could worsen the Postal Service's financial condition.

The Postal Service should move towards the use of short-run costs to develop prices if and only if (1) it can reliably determine that the response in demand will generate sufficient additional net revenue and (2) it can measure short-run costs accurately and update them regularly. In addition, there are a number of important issues that must be considered before one could be confident that using short-run costs would not harm the Postal Service:

- Using short-run costs can result in prices that may generate additional revenue in the short term but will still not allow the Postal Service to cover its institutional costs.
- Because short-run costs are more volatile, they can lead to prices that are more volatile. This makes it difficult for consumers to plan their purchases of both postal products and related goods.
- Prices act as signals to consumers. If consumers know that a low price is temporary, they may simply adjust the timing of when they purchase mail instead of increasing the volume of mail purchased. If consumers view the lower prices as being permanent (when they are not), they may inefficiently overinvest in technology needed to produce mail.
- The effectiveness of using prices based on short-run costs depends on consumer response. If consumers do not increase demand in response to a lower price, the Postal Service will just lose revenue as a result of lowering the price.
- Accurate development of the inputs required for calculating short-run costs is not always straightforward and would likely take a significant effort. For example, operational experts would need to identify and quantify which postal inputs cannot be adjusted.
- Finally, the Postal Service may lose incentive to move to the more optimal long-run operational environment if it starts pricing based on short-run costs.

Even if the Postal Service can accurately measure short-run costs, and finds using them to develop prices can lead to additional revenue, the use of short-run costs does not lead to a reduction in total costs to the Postal Service. Setting prices based on short-run costs should be seen only as a temporary solution while the Postal Service continues to move toward a more optimal operating environment.

## Table of Contents

Introduction .....	1
Definition and Theoretical Basis of Short-Run and Long-Run Costs.....	2
Determining Why Short-Run Costs Arise .....	3
The Role of Short-Run Costs in Pricing.....	4
Price Stability.....	6
Price Signaling.....	7
Consumer Response.....	7
Firm Response .....	8
Calculation of Costs.....	8
Short-Run Costs in a Postal Service Context.....	9
Seasonal Short-Run Issues.....	10
Secular Short-Run Issues.....	11
Toward Constructing a System of Short-Run Costs for the Postal Service .....	12
The Analytical Basis for Calculating Short-Run Marginal Cost.....	12
An Algorithm for Calculating Short-Run Marginal Cost for the Postal Service.....	14
Informational Requirements for Calculating Short-Run Marginal Cost .....	19
A Process for Review and Evaluation of Short-Run Costs .....	20
Conclusion .....	21

## Table

Table 1	Hypothetical Postal Example of the Relationship Between Volume and Capacity Utilization .....	17
---------	-----------------------------------------------------------------------------------------------	----

## Figures

Figure 1	Algorithm to Estimate Short-Run Costs .....	15
Figure 2	The Cost-minimizing Choice of Inputs .....	25
Figure 3	Short-Run and Long-Run Response to an Increase in Output.....	26
Figure 4	Short-Run and Long-Run Response to a Decrease in Output.....	27
Figure 5	Total Cost Curves .....	29
Figure 6	Average Cost Curves.....	30
Figure 7	Long-Run and Short-Run Marginal Costs.....	31
Figure 8	Short-Run Cost Curves.....	32

## Appendices

Appendix A	The Theoretical Basis for Short-Run Costs.....	23
Appendix B	Questions Associated with Basing Prices on Short-Run Costs.....	35
Appendix C	Practical Issues Measuring Short-Run Marginal Costs.....	38

## Short-Run Costs and Postal Pricing

### Introduction

Recently, there have been suggestions that the U.S. Postal Service needs to adjust its product costing system to account for excess capacity in the network. The argument is that the current product costing system uses estimates of long-run costs,<sup>3</sup> and that in times of excess capacity, long-run costs unfairly burden products with the cost of the excess capacity through higher prices.

While the Postal Service is no longer under a breakeven requirement, costs still play an important role in the pricing process for several reasons. First, a price cap does not set prices but rather limits their increase. Second, products must still cover their attributable costs, which may be either short-run costs or long-run costs. This is exactly the way the Postal Service based its prices upon costs in the Standard Mail Volume Incentive Pricing Programs. In those programs it demonstrated that the proposed lower prices exceeded short-run marginal cost even if they did not exceed long-run marginal cost. Third, without measuring costs it is impossible to determine whether a product's price is high enough for the product to be making a positive contribution to defraying institutional costs. Any firm would want to know the contributions made by each of its products and would want to adjust costs accordingly. Fourth, when evaluating a new product or service, it is essential to determine if the proposed prices covers the product's marginal cost. Without calculating costs correctly, this analysis cannot be accurately done.

This paper defines what is meant by short-run and long-run costs, explores the possibility of using short-run costs when developing prices, outlines what information is needed to develop short-run costs, and develops a process for review and evaluation of short-run costs. It first focuses on explaining where short-run costs come from and how they can be used in pricing, in general terms, and then investigates their applications for the Postal Service.

This paper finds that the Postal Service may want to consider setting its prices using short-run costs in certain circumstances. However, it must be understood there are risks to doing so. The Postal Service should consider using short-run costs as a basis for price, if and only if (1) it can accurately determine that the response in demand to short-run prices will generate sufficient additional net revenue and (2) it can measure short-run costs accurately.

---

<sup>3</sup> See Appendix C for a discussion of why the costs estimated in the Postal Service's product costing system are considered to be long-run costs.

## Definition and Theoretical Basis of Short-Run and Long-Run Costs

The distinction between long run and short run is not formally defined by time. Instead, it is defined by a firm's ability to adjust its inputs (labor and capital). A firm is considered to be operating in the long run when it is able to adjust all inputs to their optimal levels — that is, the combination of inputs that result in the lowest cost for a certain level of output.<sup>4</sup> A firm is considered to be operating in the short run when it cannot adjust its inputs to their optimal levels. On a practical basis, although the long-run/short-run distinction is not formally defined by time, it is intimately related to time, because the ability of a firm to adjust its inputs is linked to the time it takes to make the adjustments.

*A firm is in the short run when it cannot adjust at least one of its inputs (labor and capital) to its optimal levels.*

The theoretical basis for short-run costs is explained in detail in Appendix A, but the following is a summary of the key points developed there:

- Short-run total cost for a firm is generally greater and never less than long-run total cost for a firm. This means that it is worse for a firm's financial performance to be in the short run. For example, if a firm is in the short run because of excess capacity in one or more of its inputs, it still has to pay for that capacity even if it does not use it.
- Under certain conditions, a firm's short-run marginal cost can be less than its long-run marginal cost.<sup>5</sup> In other words, the additional cost of supplying more output could be less in the short run than in the long run. This occurs when the firm has excess capacity of a fixed input and does not need as much additional variable inputs to produce additional output.
- A firm for which short-run marginal cost is less than long-run marginal cost may benefit financially by lowering its price so that it only covers short-run marginal cost. However, this will only benefit the firm if the additional revenues are higher than the additional variable cost.<sup>6</sup> In this situation, the firm can generate additional contribution that helps defray the fixed cost of the inflexible input.

These theoretical results raise the possibility that using short-run costs could benefit the Postal Service, but a number of important issues must be resolved before one could be confident that using short-run costs would not hurt the Postal Service. We investigate these issues in this paper.

---

<sup>4</sup> This is not to say that all inputs are completely flexible in the long run. See Appendix A for a discussion of why some inputs will remain fixed in the long run.

<sup>5</sup> Marginal cost is the additional cost caused when the firm produces an additional unit of output. For the Postal Service, the marginal cost is the cost of handling an additional piece of mail.

<sup>6</sup> Variable cost is cost that changes when the firm's level of output changes.

## Determining Why Short-Run Costs Arise

There are a number of reasons why a firm might find itself in a short-run position, i.e., it is in a situation where it cannot adjust at least one of its inputs to its cost-minimizing level. The implications for how a firm should respond to being in the short run will vary based on the reason the firm cannot adjust the input(s). Below we describe several reasons why short-run costs arise (why a firm may be unable to adjust all of its inputs), but these reasons are not meant to be mutually exclusive. A firm could be in the short run for more than one of these reasons at the same time.

- Intentional overcapacity – When the transaction costs of acquiring additional amounts of a particular input are very high, it may be rational for a firm to purchase “too much” of that input relative to current production levels. For example, perhaps a firm is building a traffic tunnel or a gas pipeline for which adding capacity in the future is extremely expensive and the firm anticipates that demand will be increasing in the future. In this situation, the firm may intentionally acquire more capacity than it needs in the present or near future and will be operating in a position of short-run excess capacity. *How a firm should respond to being in the short run depends, in part, on why the firm is in the short run.*
- Accidental undercapacity or overcapacity – Many times a firm must acquire its inputs in advance of production. If the transaction costs of acquiring or reducing an input are high, or time consuming, the firm may have to commit to the amount of the input it will acquire far in advance of actual production and sale. In those instances, the firm must forecast its future sales. Often those forecasts are in error and the actual amount of the input that the firm needs is greater or less than it had acquired.
- Fluctuating demand – The demand for a firm’s product may fluctuate over the course of a day, a month, or a year. If one of the inputs cannot be easily adjusted, the firm will find that it is in a position of suboptimal capacity during some time periods. If the firm acquires enough of this “lumpy” input to serve the peak periods, then it will have overcapacity during slack periods. In contrast, if the firm acquires just enough of the lumpy input to serve periods of average demand, then it will have insufficient capacity during peak periods.
- Persistent rise or fall in demand – A firm may experience a situation in which the demand for its product is persistently rising or falling. This can cause it to face extended periods of time with a suboptimal level of capacity. Suppose the firm faces a persistent fall in demand. Although the firm may have minimized costs to produce the current level of volume, as demand falls, the capacity will become too large. Eventually the firm will adjust and reduce its capacity, but because the decline in demand is persistent, that new lower level will also eventually become too large. This ongoing process may cause the firm to suffer from repeated periods of excess capacity.



- Indivisibility of the fixed factor – Certain inputs are indivisible, meaning that they cannot be readily broken down into smaller units. For example, the Flat Sequencing System (FSS) machines can only be purchased as a whole; it is not possible to purchase just a part of one. If the demand for the firm’s good is less than the minimum capacity of one unit of the input, the firm will have excess capacity. In addition, if the firm’s demand falls in between the capacities associated with the whole unit amount of the input, it will have to choose between excess and insufficient capacity.
- Legal or regulatory restrictions on adjustment – There can be instances in which the firm stays at a suboptimal level of capacity even when it could adjust its inputs. This is because the firm faces a legal or regulatory requirement to maintain its current level of capacity. A universal service requirement is an example of a regulatory requirement that may force the firm to sustain excess capacity. Alternatively, it may have legally enforceable labor agreements that preclude it from making the desired adjustment of its labor force.
- Internal impediments to adjustment – Sometimes internal barriers hinder a firm’s ability to adjust one or more of its inputs. Resistance to any change in the status quo, lack of acceptance of the changing market demand, decentralized decision-making, and inability to coordinate a reduction in inputs across departments or divisions are all examples of internal impediments that may stop a firm from adjusting its inputs. In addition, there may be reasons why a firm would not make the adjustment even when it was needed to minimize costs. The firm may not want to send public signals (such as closing a facility) that the demand for its product is falling. Or the firm may fear consumer, regulatory, or political backlash against the adjustment.

## The Role of Short-Run Costs in Pricing

When a firm is operating in the long run, there is only one set of product costs that the firm considers when setting prices. However, when the firm is operating in the short run, a choice arises. Should the firm base its prices on long-run costs or short-run costs?

In pure economic theory, there is little uncertainty about when firms should use short-run costs as the basis for setting prices. This is because both the role of short-run costs and the firm’s goals are well defined and unambiguous. In economic theory, the “fixity” in the short run for a profit maximizing firm is usually discussed as a situation in which the firm’s “plant” or “factory” is fixed in size. This reflects the fact that for most firms it is more difficult to adjust its physical capital input — buildings and equipment — than it is to adjust its labor input. Once the firm’s “plant” is fixed, it then attempts to maximize profits given that plant size. To do so, it uses its short-run marginal cost curve in setting price or output.

For a monopoly, the theoretical distinction between the short-run and the long-run cost curves is rarely considered, because a monopoly chooses its price and associated level

*There is not a universal answer to the question of which costs, short-run or long-run, should serve as the basis for a firm's pricing decisions.*

of production. Because of its market power, the monopoly does not have to take its price from the market. This means that, in theory, there is no reason for the monopoly to be off its long-run marginal cost curve. As a result, the theoretical cost analysis of a monopoly is typically the same for short run and long run.

But, this is just in theory. As previously discussed, in practice there are a variety of reasons why a monopoly firm may not be operating at its cost-minimizing plant size. If one of these reasons causes a monopoly firm to be away from its optimal capacity for a material period of time, then the monopoly firm has the option of setting its prices based upon short-run marginal cost rather than long-run marginal cost. This option arises because the monopoly firm is a “price setter” not a “price taker” and typically sets price above marginal cost.<sup>7</sup> For example, this issue arises for utilities such as electric companies or water systems that are often legal monopolies and use large, indivisible units of capital in the production of their goods.

Experience has shown that there is not a universal answer to the question of which costs, short-run or long-run, should serve as the basis for the firm's pricing decisions.<sup>8</sup> There are advantages and disadvantages to each choice. The primary advantage of using long-run marginal cost as the basis for pricing is that its use is consistent with the “best” outcome, in which resources are efficiently allocated given the market structure. This is the efficient outcome because consumers and firms are responding to the set of price signals that lead them to their overall optimal choices of consumption and production. In other words, the use of long-run marginal costs as the basis for prices facilitates moving the market to its long-run, most profitable position.

However, during the transition to its long-run position, the use of long-run costs as the basis for prices can lead to inferior short-run results. This fact points to the primary advantage of using short-run marginal cost pricing — it provides the prices that can better match demand to supply during temporary conditions in which capacity cannot be set to its cost-minimizing value. This provides a potential for the firm to offset its high level of short-run fixed costs.

Simply pricing above a product's short-run marginal cost does not ensure that the price is sufficiently high so that the firm is earning a profit on the units sold. For a positive profit to be earned, the firm must also charge a price that is above its short-run average total cost. However, in the short

***Primary Advantage of Using Short-run and Long-run Costs to Set Price:***

***Long Run — Consistent with achieving its optimal long-run position.***

***Short Run — May allow firm to utilize capacity when it cannot adjust all of its inputs.***

---

<sup>7</sup> Other firms with market power, like “oligopolists” who sell in a market with just a few firms, also have the ability to set prices and could face this pricing issue.

<sup>8</sup> For example, there has been an ongoing debate as to whether short-run or long-run marginal cost pricing should be used in pricing electricity. For a review of this debate, see R. Anderson and M. Bohman, “Short and Long Run Marginal Cost Pricing: On Their Alleged Equivalence,” *Energy Economics*, October 1985, pp. 279-288.

run, the firm may want to produce and sell even if the price does not cover short-run average total cost. If the firm's price is above its short-run average variable cost then its sales will contribute to covering the firm's fixed cost, associated with the firm's fixed input. While this will not allow the firm to cover all of its costs in the long run, it does allow it to reduce its losses in the short run.

From the firm's perspective, the choice between long-run and short-run cost thus depends upon two factors:

- The relative weight that the firm puts on these two, potentially conflicting, goals: achieving its optimal long-run position and utilizing its capacity in the short run.
- Whether and to what extent the use of short-run costs will actually improve capacity utilization in the short run, relative to the extent to which using long-run costs will help the firm achieve its optimal position.

In evaluating these two factors, comparison of the roles of the two types of cost can be focused on a set of specific issues that highlight their relative strengths and weaknesses. Each of these issues is presented and discussed below, along with an indication of the implications for use of short-run and long-run costs.

### ***Price Stability***

Consumers prefer stable, predictable prices and can make more efficient consumption choices when prices are relatively stable through time.<sup>9</sup> Stable prices lead to better choices because price volatility erodes the information content of prices. When there is a lot of price volatility, it is more difficult for consumers to compare the prices they pay for relative goods. The stability of prices is particularly important when the "consumer" (which could be another firm or government agency) is purchasing the good or service as an input into the production of other goods. The purchase of the good at issue is then part of the purchase of a set of inputs, some of which may be non-fungible capital goods. For example, postage might be just one input purchased by firms that produce and distribute catalogs. Those firms might also need to purchase computer software and printing equipment. When the price of one input is highly variable, it makes it more difficult for firms to choose the best amounts of all inputs.

***Pricing based on short run costs can lead to price volatility and customers tend to prefer stable, predictable prices.***

Short-run marginal costs are more variable than long-run marginal costs because they change as the amount of the fixed input changes. Think about a situation in which a firm finds itself in the short run because of persistently declining demand. Because it is difficult or expensive to adjust one of its inputs, this

firm will find itself regularly in the short run, even though it is attempting to adjust capacity as demand continues to fall. This means that each adjustment of capacity only

---

<sup>9</sup> Ibid. for a discussion of the role of price stability and price signaling as they apply to the electric utility industry. For evidence that postal customers value predictable price changes, see U.S. Postal Service and Postal Rate Commission, Ratemaking Summit, May 28, 2002, <http://www.prc.gov/prc-docs/library/archived/day1-transcript.pdf>.

temporarily brings the firm back to its cost-minimizing capacity. Once demand starts to fall again, excess capacity will again increase.

If the firm sets its prices based upon short-run marginal cost, it will lower prices whenever it is in a condition of excess capacity. However, after capacity is adjusted to its optimal level, the firm will need to increase its prices to reflect the higher short-run costs (when a firm is at its optimal capacity, short-run costs equal long-run costs). If demand for the firm's product follows this pattern, the prices that consumers face will vary much more if they are based upon short-run marginal cost than if they are based upon long-run marginal cost. The resulting variation in price makes it more difficult for consumers to plan their purchases of both this good and related goods through time, and represents a potential advantage of basing prices on long-run costs.

### ***Price Signaling***

Prices play an essential informational role in markets, allowing consumers to make informed decisions about which set of purchases best satisfies their consumption needs. In addition to their role in determining current purchases, current prices also provide information about the future costs of acquiring goods.

Because short-run prices will differ from long-run prices, the firm should consider the price signaling implications of using short-run prices. Prices based upon very low short-run marginal costs could lead to both overconsumption and overinvestment in complementary goods. For example, extremely low gasoline prices might lead consumers to buy larger, less fuel efficient vehicles. In contrast, if consumers do not expect the low prices to continue into the future, then the use of low short-run prices could lead consumers to simply change when they purchase a good, shifting planned future purchases to the present. Such a response would cause a lower average purchase price for the firm, possibly cancelling out the potential benefit to the firm from making better use of capacity in the short run.

*Using prices based on short-run costs can result in price signaling that could have implications on the firm's revenue and consumers' satisfaction over time.*

In sum, firms using short-run prices should be aware that their use has price signaling effects that could have implications for both their own revenue and their consumers' satisfaction through time.

### ***Consumer Response***

The primary advantage of using short-run costs is that the resulting prices can match demand to supply during a period of suboptimal capacity. However, such a matching is dependent upon the responsiveness of consumers to price changes. Consider the situation in which the firm has excess capacity, so that the short-run marginal cost is less than the long-run marginal cost, and the short-run price will be below the long-run price. If consumers respond to the lower price by expanding purchases in a timely

fashion, then output will expand and the firm will be able to make better use of its excess capacity.

*If consumers are generally unresponsive to the price change, then the firm will just lose revenue as a result of lowering its price.*

However, if consumers are generally unresponsive to the price change, then the firm will just lose revenue as a result of lowering the price.<sup>10</sup> In this instance, it is better for the firm to use long-run prices rather than short-run prices.

Similarly, if consumers' response to price changes is very slow, there may not be enough time for them to respond to a price decrease before the excess capacity is eliminated. This also leads to a preference for long-run prices.

The responsiveness of consumers to a price change may be unknown or difficult to measure. If so, the risk associated with using lower short-run prices produces an advantage for sticking with long-run prices.<sup>11</sup>

### ***Firm Response***

In the realm of economic theory, firms always move to the long run by adjusting the fixed input as soon as it is practicable. This is because adjusting the amount of the fixed input is profit enhancing and it is assumed that there are no transaction costs of making the adjustment. Actual firms face a more complex situation, and adjusting the amount of fixed input may have a number of impediments, both internal and external, that delay the adjustment from taking place. These impediments not only cause the firm to be in the short run but also increase the likelihood that the firm will stay in the short run for a longer time.

In an environment in which the adjustment to the long run may not be made rapidly, the impact of using short-run costs to set prices *on the behavior of the firm* needs to be considered. Firms are motivated to move to the long run because being in the short run is causing them to incur losses or have smaller profits than are otherwise possible.

*Using short-run costs to set price may remove part of the incentive for the firm to move to its optimal environment.*

To the extent that using short-run costs leads to prices that better match demand to supply, the short-run losses may be smaller than if the price stayed at its long-run value. While this is obviously an improvement in the firm's short-run financial position, it does remove part of the incentive for the firm to move to its efficient long-run combination of inputs. As a result, short-run pricing could harm the firm's long-run performance.

### ***Calculation of Costs***

A final practical consideration in using short-run costs to set prices is the firm's ability to calculate short-run costs in an accurate and timely way. Short-run costs are more

<sup>10</sup> The condition in which consumers are unresponsive to price changes is known as "inelastic demand." When demand is inelastic, a lower price leads to a fall in revenue for the firm.

<sup>11</sup> The opposite occurs during a period of insufficient capacity. Then, a lack of consumer responsiveness is a good thing for the firm because the higher short-run prices increase firm revenues without driving customers away.

difficult to calculate than long-run costs because they include the additional requirements of identifying the input that is not being adjusted and calculating the

*Short-run costs are more difficult to calculate and change more often than long-run costs. Before embarking on a short-run pricing strategy, the firm should be confident that it could accurately estimate and regularly update short-run costs.*

impact of that fixity on product costs. These are relatively difficult tasks and this difficulty could force the firm to use approximations to the actual short-run costs. These approximations reduce the accuracy and reliability of short-run costs and raise a caution with regard to their use.

Also, short-run costs change when the capacity is adjusted and would need to be calculated more often than long-run costs. If the firm is not able to recalculate its short-run costs on a timely basis, it may find itself in a situation in which its estimated long-run costs are closer to its true short-run costs than are its estimated short-run costs. Before embarking on a pricing strategy based upon short-run costs, the firm should be confident that it can accurately estimate those costs.

## Short-Run Costs in a Postal Service Context

The Postal Service would not seem to be a candidate for a short-run analysis because it has characteristics that typically result in a firm being able to adjust its inputs — most notably it is labor intensive and its capital is not concentrated in a few large plants. Despite these characteristics, the Postal Service does have difficulties adjusting some of its inputs, particularly in response to volume declines. If the decline in volume is sufficiently large, these impediments to input adjustment make a short-run analysis a reasonable option for the Postal Service to consider. These adjustment difficulties arise for several reasons:

- Labor agreements limit the Postal Service's ability to adjust its workforce in response to variations in volume. While it has obtained improved flexibility in recent labor negotiations, the Postal Service still faces constraints on adjusting labor both within a given year and across years.
- The Postal Service faces impediments to closing Post Offices, particularly in rural areas. Despite having tens of thousands of Post Offices, the Postal Service has been able to close only a relatively small number of them in response to recent volume declines. Recently it has been reducing hours at a number of smaller Post Offices but this does not reduce the actual numbers of offices it must sustain.
- There is often resistance to closing mail processing plants. The recent Mail Processing Network Rationalization Service Change case, which was litigated at

*Despite being labor intensive, the Postal Service faces a number of difficulties adjusting some of its inputs, particularly in response to volume declines.*

the Postal Regulatory Commission (PRC), illustrates the challenges the Postal Service faces when attempting to downsize its mail processing network.<sup>12</sup>

- Because of its universal service obligation, the Postal Service has substantial fixity in its delivery network and must currently provide the same frequency of delivery to all areas of the country regardless of mail volume. Similarly, because of the need to satisfy service standards, the Postal Service could have fixity in a number of its networks including transportation and delivery.
- In addition to external constraints, the Postal Service has a number of internal barriers that hinder it from adjusting its inputs. These include things like the lack of a well-articulated vision and a lack of data about potential savings.<sup>13</sup>

These input inflexibilities preclude the Postal Service from always being able to operate at optimal capacity. During periods of low volume, it may have excess capacity and during periods of high volume it may have insufficient capacity. Moreover, because the demand for postal products varies both within the year and across years, the Postal Service experiences both seasonal and secular short-run issues. Both of these are introduced and discussed below.

It is important to make a distinction between unused capacity that exists due to normal operating conditions and unused capacity that exists due to the inability to adjust inputs in reaction to changing market conditions. The impediments discussed in this section preclude the Postal Service making adjustments to its inputs in reaction to the substantially lower volumes. Note, however, that even if the Postal Service were to achieve the optimal input mix, that does not mean that all of its resources will be fully utilized at all times. Because of its service standards and fluctuations in demand, the Postal Service will regularly have unused capacity in certain parts of its network on certain days. This unused capacity is different from the excess capacity that occurs in the short run because the Postal Service cannot adjust all of its inputs.

### ***Seasonal Short-Run Issues***

Mail volume has a pronounced seasonal pattern, with the peak volume occurring in the first postal quarter (October through December) and the trough volume occurring in the fourth postal quarter (July through September).<sup>14</sup> Because of inability to sufficiently adjust its inputs within the year, the Postal Service faces a potential short-run condition, with suboptimal capacity during both peak and trough periods. During the peak period, the Postal Service faces inadequate capacity and must use additional resources to process, transport, and deliver the mail. To the extent it must use less efficient or more expensive methods of operation during this period, its short-run marginal cost will be

---

<sup>12</sup> U.S. Postal Regulatory Commission, "Advisory Opinion on Mail Processing Network Rationalization Service Changes," Docket No. N2012-1, September 28, 2012, [http://www.prc.gov/Docs/85/85269/Advisory Opinion %20PDF%20 09282012.pdf](http://www.prc.gov/Docs/85/85269/Advisory%20Opinion%20PDF%2009282012.pdf).

<sup>13</sup> U.S. Postal Service Office of the Inspector General, *Barriers to Retail Network Optimization*, RARC-WP-11-005, June 9, 2011, [http://www.uspsoiq.gov/foia\\_files/RARC-WP-11-005.pdf](http://www.uspsoiq.gov/foia_files/RARC-WP-11-005.pdf).

<sup>14</sup> There is some variation on when specific products have their troughs and peaks.

above its long-run marginal cost. In other words, it will cost more to handle additional mail volume than it would if the Postal Service were able to adjust all of its inputs.

*Using short run costs to set prices in response to seasonal variation in volume implies charging higher prices during the peak volume times and charging lower prices during the slack volume times.*

In contrast, during the summer volume decline, the Postal Service will find itself in a position of excess capacity because it cannot reduce some of its inputs. During this time, its short-run marginal cost will be below its long-run marginal cost. In other words, because there is excess capacity, the additional cost of handling additional mail volume is less than it would be if the Postal Service were

under optimal capacity. And the inverse is also true. As volume declines, the Postal Service sheds fewer costs than it would if it were able to adjust its inputs to be at optimal capacity.

Using short-run costs to set prices in response to this seasonal variation in volume would imply charging higher prices during the peak volume times and charging lower prices during the slack volume times. This pricing approach would discourage additional volume during the peak period, when it is expensive to handle, and encourage additional volume during the slack periods, when it can be handled more cheaply.

This pricing mechanism is often accomplished through a system of surcharges and discounts. Although the Postal Service has not charged surcharges during the peak volume period, it has experimented with providing discounts for certain products during the slack volume period. For more information on the Postal Service's seasonal discounts, see the brief discussion of previous work calculating short-run marginal costs in Appendix C.

### ***Secular Short-Run Issues***

In recent years, in addition to its seasonal demand fluctuations, the Postal Service has faced an ongoing secular demand for its products, a decline that is expected to continue into the future. Because the volume decline is large and persistent, the Postal Service has worked to reduce its retail, mail processing, transportation, and delivery networks in an attempt to become fiscally viable. However, its ongoing operational deficits suggest that it has been unable to reduce capacity to its long-run cost-minimizing position.

*The existence of ongoing excess capacity raises the possibility that the Postal Service may want to consider whether its pricing should be based on short-run or long-run marginal costs.*

To the extent that this is true, the Postal Service is operating in a short-run situation of excess capacity, caused mainly by the persistent decline in demand. This existence of excess capacity raises the possibility that the Postal Service's short-run marginal costs may be below its long-run marginal costs and suggests that that the Postal Service may want to consider whether its pricing should be based upon short-run or long-run marginal costs.



Note that this secular condition is separate but potentially related to the seasonal variations discussed above. The secular and seasonal conditions are related through the impact of the secular excess capacity on seasonal variations in capacity. For instance, to the extent a secular decline in demand has created excess capacity in parts of the Postal Service network throughout the entire year, it would affect the seasonal capacity issues. When the Postal Service has secular excess capacity, it is less likely that it will have insufficient capacity during the seasonal peak period and more likely that it will have greater excess capacity during the seasonal trough. This would affect the relative sizes of the appropriate surcharges and discounts.

## **Toward Constructing a System of Short-Run Costs for the Postal Service**

An important part of a firm's determination as to whether it wants to estimate and use short-run costs is assessing the difficulties it faces in calculating them. Estimation of short-run costs requires developing the appropriate algorithm to calculate those costs, and it also creates additional informational requirements for the firm. In this section we provide the analytical structure, computational algorithm, and information requirements for calculating short-run costs for the Postal Service. In addition, in Appendix C, we discuss a number of important practical issues the Postal Service faces in measuring short-run costs.<sup>15</sup>

*An important part of a firm's decision to estimate and use short-run costs is assessing the difficulties it faces in calculating them.*

### ***The Analytical Basis for Calculating Short-Run Marginal Cost***

The Postal Service already has a vetted costing system for calculating long-run marginal cost, so it is efficient to embed the calculation of short-run marginal cost into that long-run costing system.<sup>16</sup> Therefore, before deriving the analytical basis for a short-run cost calculation, we start with a description of the computational relationship between short-run and long-run marginal costs for the Postal Service.

The Postal Service estimates long-run marginal cost through calculating a series of "component specific" long-run unit volume variable costs. A "component" or "cost component" is an activity within the postal value chain that has an identifiable cost driver (the main attribute that is affecting how much cost is incurred) and a relatively homogenous relationship between the cost driver and the incurred cost. For example, the delivery of mail on the street is a cost component, and the related cost driver is the

---

<sup>15</sup> Ideally, a short-run costing system would provide estimates of both short-run marginal cost and short-run average incremental cost. Incremental costs are relatively complex and difficult to measure, so the Postal Service currently does not measure long-run incremental cost for most products but rather uses a proxy called long-run "attributable" cost. Consistent with that approach, we will discuss only the calculation of short-run marginal costs and will leave a discussion of short-run average incremental cost to future research.

<sup>16</sup> One could develop a methodology to calculate short-run costs that does not rely on the current long-run costing system.

number of delivered letters, flats, and parcels pieces.<sup>17</sup> Similarly, the sorting of letters on a delivery barcode sorter is another cost component and the cost driver is the pieces of mail sorted.

With this structure in mind, we can now write the formula for calculating the unit volume variable cost for the “*i*th” product in the “*j*th” component. The formula starts with the accrued cost for that component,  $C_j$ . The cost is then multiplied by the long-run elasticity of cost with respect to the cost driver,  $\varepsilon_j^{LR}$ , to form the total volume variable cost for the component.<sup>18</sup> The “*i*th” product then receives its appropriate share of that volume variable cost, which is found by multiplying the total volume variable cost by the “*i*th” product’s proportion of the distribution key,  $\delta_{ij}$ .<sup>19</sup> This produces the “*j*th” component’s long-run volume variable cost for the “*i*th” product, which is then divided by the product’s originating volume,  $V_i$ , to get the component long-run unit volume variable cost:<sup>20</sup>

$$UVVC_{ij}^{LR} = \frac{C_j^{LR} \varepsilon_j^{LR} \delta_{ij}}{V_i}.$$

The “*i*th” product’s overall long-run unit volume variable cost is just the sum of the component unit long-run volume variable costs:

$$UVVC_i^{LR} = \sum_{j=1}^m UVVC_{ij}^{LR} = \sum_{j=1}^m \frac{C_j^{LR} \varepsilon_j^{LR} \delta_{ij}}{V_i}.$$

Using this general approach, the Postal Service can estimate the short-run marginal cost for a product by using the calculation of its short-run unit volume variable cost as a proxy.<sup>21</sup> This requires calculating and substituting short-run component cost elasticities for the existing long-run component cost elasticities. Otherwise, the analytical basis is the same as the one used for calculating long-run marginal costs, as demonstrated by the following equation:<sup>22</sup>

<sup>17</sup> As the volume of mail increases, the cost of delivering mail on the street also increases. Shape is also a cost driver because letters are less expensive to deliver than flats, and flats are less expensive to delivery than parcels.

<sup>18</sup> The elasticity of cost is not the same thing as price elasticity. Cost elasticities measure how much costs change with respect to changes in the cost driver. Price elasticities measure how much the quantity demanded for the product changes in response to a change in price.

<sup>19</sup> A distribution key is a measure of what proportion of the cost component’s cost driver is caused by each product.

<sup>20</sup> For a mathematical demonstration of the equivalence between unit volume variable cost calculated in this way and marginal cost see Michael Bradley, Jeff Colvin, and John Panzar, “On Setting Prices and Testing Cross-Subsidy Using Accounting Data,” *Journal of Regulatory Economics* 16 (1999): 83-100.

<sup>21</sup> Marginal cost is a theoretical concept that is difficult to calculate. It is generally accepted in postal economics that unit volume variable cost, which can be calculated, is a good proxy for marginal costs.

<sup>22</sup> In theory, short-run total cost and long-run total cost can be different. However, as a practical matter, it is not possible for the Postal Service to have two sets of total accrued costs. This is because at a point in time the Postal Service is either in the long run or in the short run, but not both. Formally speaking, one cannot calculate both long-run and short-run marginal cost from one set of actual cost data. Nevertheless, as a practical matter both measures are computable. One simply needs to start from an assumption as to whether the accrued costs for a given time period represent long-run or short-run total cost.

$$UVVC_i^{SR} = \sum_{j=1}^m UVVC_{ij}^{SR} = \sum_{j=1}^m \frac{C_j^{SR} \varepsilon_j^{SR} \delta_{ij}}{V_i}.$$

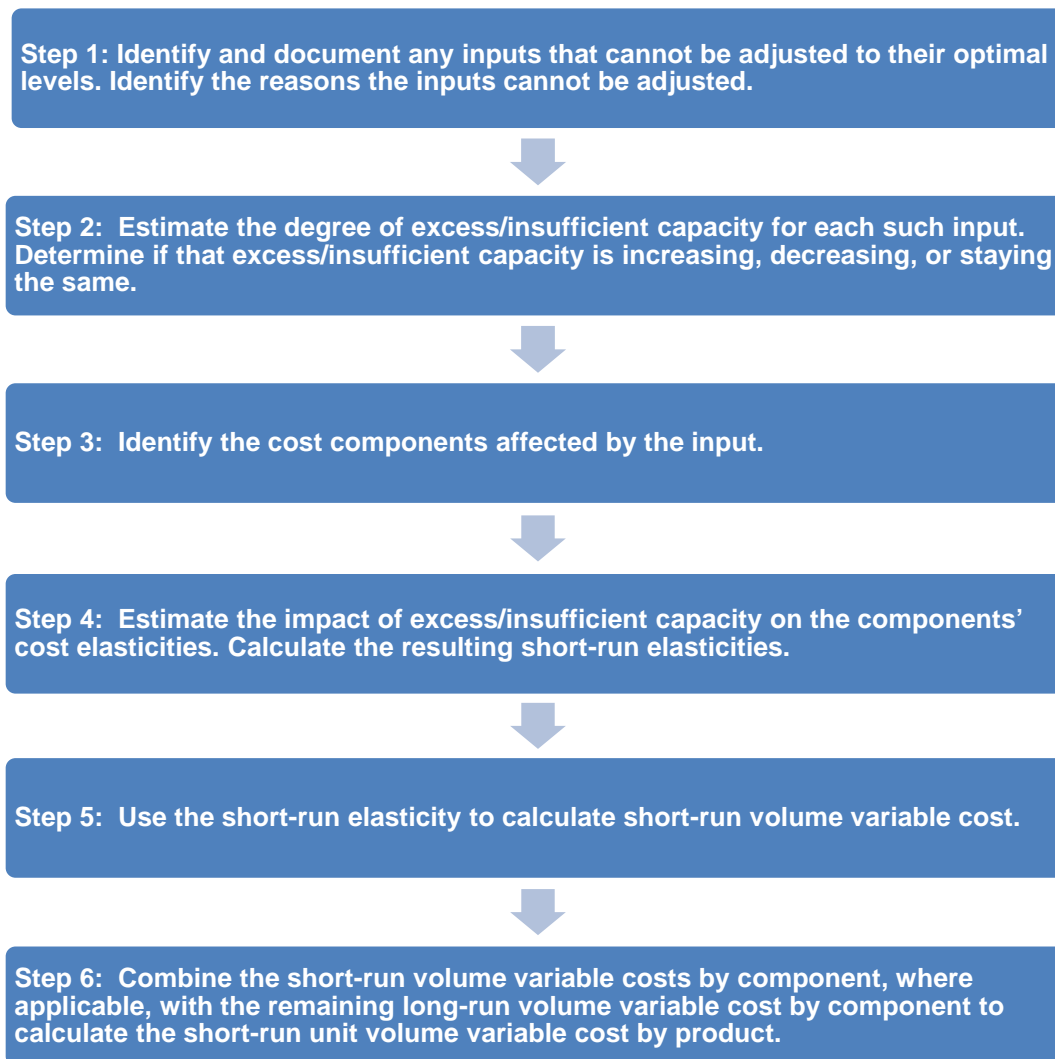
This equation demonstrates that accurately calculating short-run cost elasticities is the primary computational challenge for the Postal Service in calculating short-run costs. There are a variety of potential methods for estimating the short-run elasticities and the method of estimation may well differ from the method currently used to estimate a long-run elasticity. The central role played by calculating these elasticities is also highlighted by a review of a computational algorithm that the Postal Service could follow in applying the analytical structure.

*The primary computational challenge for the Postal Service in estimating short-run costs is accurately calculating short-run cost elasticities.*

### ***An Algorithm for Calculating Short-Run Marginal Cost for the Postal Service***

Just as it is logical to investigate the analytical structure for short-run marginal cost within the Postal Service's analytical structure for long-run costs, so too it is logical to develop a computational algorithm for short-run costs that makes use, as much as possible, of existing Postal Service data sources and computational methods. Such an approach also makes it more likely that short-run costs can be updated and modified when necessary. The multistep computational algorithm presented in this section implements an analytical basis for short-run marginal cost calculation within the general structure of the Postal Service's long-run product-costing system.

Figure 1: Algorithm to Estimate Short-Run Costs



Note that the proposed algorithm is discussed in terms of calculating short-run costs arising from a secular decline in volume but that it is equally applicable for short-run costs arising from seasonal variations in volume. To simplify this discussion, the remainder of this description will use the term “excess” capacity but the algorithm could also apply to “insufficient” capacity.

**Step 1: Identify and document any inputs that cannot be adjusted to their optimal levels. Identify the reasons the inputs cannot be adjusted.**

As discussed above, the difference between the long run and the short run is the fact that in the short run at least one of the firm’s inputs cannot be adjusted. Therefore, the calculation of short-run costs starts with identification and documentation of any inputs

that have not been adjusted to their optimal levels. This would necessitate significant input from experts in postal operations.

A place to start this analysis is with the decline in volume that has occurred since the fiscal year (FY) 2006 volume peak. Since then, the Postal Service has lost nearly 22 percent of its volume, averaging a 5.7 percent annual decline over the period from FY 2008 through FY 2011. Because of its persistence, and because it has continued after the recession ended in mid-2009, this volume decline looks to be a secular decline. It is thus possible to use the volume decline as a benchmark for identifying which inputs have or have not been proportionately reduced.<sup>23</sup>

In addition, the institutional structure of the Postal Service can be reviewed in order to identify where it faces impediments in adjusting inputs. For example, labor agreements can be reviewed to identify any contract rules that slow the Postal Service from reducing the number of employees or hours worked. In addition, the regulatory and political processes can be reviewed to identify any regulatory or political constraints that hinder closing offices or downsizing networks. This process can also provide insights into which inputs are difficult for the Postal Service to adjust.

**Step 2: Estimate the degree of excess/insufficient capacity for each such input. Determine if that excess/insufficient capacity is increasing, decreasing, or staying the same.**

Once it has been determined which inputs are difficult to adjust, the next step is to estimate how far the input is away from the optimal level. That can be measured by the degree of excess capacity in an input. This is often measured with a statistic called “capacity utilization,” which is the ratio of the firm’s current production to its capacity to produce.

To understand how this measure is useful, consider a simple hypothetical example in which the Postal Service volume is assumed to be originally 200 billion pieces and the handling of that volume required 40 million labor hours. Assume that the Postal Service was in long-run equilibrium with respect to this labor input, and therefore was initially at full capacity. In other words, 40 million postal labor hours have the capacity to handle 200 billion pieces of mail, so the ratio of actual production to capacity is 100 percent.

Now assume, hypothetically, that volume declined by 20 percent to 160 billion pieces but the number of hours for this labor input stayed the same. The Postal Service’s capacity to handle mail would remain at 200 billion pieces, although its actual “production” would be only 160 pieces. Its hypothetical capacity utilization would be 160 billion / 200 billion or 80 percent, indicating that excess capacity exists in this activity.

---

<sup>23</sup> Note that the identification of appropriate proportional reduction in an input is not as easy as comparing the percentage decline in volume and the percentage decline in the various inputs as reported in the Revenue, Pieces and Weight (RPW) report. First, not all types of volume declined at the same rate and different products may make use of different mixes of inputs. Thus, one needs to be aware of the possible impact of changes in the mix of mail. Second, long run equilibrium may not require a proportional response in all inputs. To the extent the Postal Service has long run economies of scale or density, the percentage decline in inputs will be less than the percentage decline in volume even in long run equilibrium.

Note that unutilized capacity can arise even if the Postal Service is able to partly adjust its inputs. Suppose we revise the hypothetical example to allow the Postal Service to reduce its amount of the input by 4 million hours, reducing its total labor hours to 36 million. Now its capacity to handle mail volume is 180 billion pieces.<sup>24</sup> After the adjustment, the capacity utilization is 160 billion / 180 billion or 88.9 percent. This simple example also demonstrates the point that the lower the capacity utilization, the greater will be the excess capacity.

**Table 1: Hypothetical Postal Example of the Relationship between Volume and Capacity Utilization**

	Sorted Volume (Billions)	Labor Hours (Millions)	Productivity (Pieces/Hour)	Sorting Capacity (Billions)	Capacity Utilization
Baseline	200	40	5,000	200	100.0%
Scenario 1: Volume Decline, No Adjustment	160	40	4,000	200	80.0%
Scenario 2: Volume Decline, Partial Adjustment	160	36	4,444	180	88.9%

**Step 3: Identify the cost components affected by the input.**

The excess capacity must ultimately be translated into the implications for how cost responds to volume changes. The first step in that process is to identify which cost components are affected by the input that cannot be adjusted to its optimal level and has excess capacity. The Postal Service’s product cost system assigns its general ledger costs to a series of cost components that reflect the various activities required to accept, transport, sort and deliver the mail.<sup>25</sup> By identifying which general ledger accounts are assigned to each component, one can find which components include the inputs being investigated.

**Step 4: Estimate the impact of excess/insufficient capacity on the components’ cost elasticities. Calculate the resulting short-run cost elasticities.**

Each cost component in the Postal Service’s product cost model has a long-run cost elasticity, which measures how quickly the component’s cost changes as the cost driver

<sup>24</sup> We calculate the Postal Service’s current capacity by multiplying its current hours by productivity of hours in long run equilibrium. In the long run, volume was 200 billion pieces and labor hours were 40 million. This means that hourly productivity was 5,000 pieces per hour. When the total number of hours is reduced to 36 million, throughput capacity is found by multiplying those 36 million hours by 5,000 pieces per hour. That multiplication yields a capacity of 180 billion pieces.

<sup>25</sup> For a detailed description of the cost components see the U.S. Postal Service, *Summary Description of USPS Development of Costs by Segments and Components*, filed at the Postal Regulatory Commission periodically, recently on July 2, 2012, <http://www.prc.gov/prc-pages/library/detail.aspx?docketId=&docketPart=Documents&docid=83424>.

increases or decreases. The cost elasticities are determined in a number of ways, such as econometric estimation, functional or engineering analysis, or operational assumption. When calculating short-run marginal cost, the Postal Service must consider, on a component-by-component basis, if these long-run cost elasticities are still applicable. In any instance in which the long-run cost elasticity is not applicable, a short-run cost elasticity must be estimated.

Short-run cost elasticities can be estimated by the same methods used for long-run cost elasticities: econometric estimation, functional or engineering analysis, and operational assumption. The key difference is that the short-run cost elasticity must take into account the fixed input. For example, in econometric estimation, a long-run cost function will estimate the relationship between a component's total cost and its level of the cost driver, controlling for changes in input prices, where appropriate. Such a long-run cost function would have the general form of  $C = f(V, w)$ , where  $C$  is a component's total cost,  $V$  is the amount of the cost driver produced in the component, and  $w$  is the vector of input prices, such as wages. A short-run cost function would also have to take into account that one of the inputs, such as capital,  $K$ , is fixed. A short-run cost function would include this constraint explicitly in the equation and would have the general form of  $C = f(V, w, K)$ .

A functional or engineering analysis approach looks at the individual activities within a component and makes an operational determination as to which of those activities vary as the amount of the cost driver handled varies. A short-run approach would require reevaluating that functional determination to see if the fixed input causes a change in the activities that vary with changes in the cost driver.

Finally, operational experience and judgment is sometimes used to determine how a cost varies with changes in the amount of the cost driver. In estimating short-run cost elasticity, these assumptions need to be revisited to see if they need to be revised to take into account the fixed input. The existence of excess capacity would cause the short-run cost elasticity to be less than the long-run cost elasticity. Just the opposite condition would occur if there were to be insufficient capacity.

**Step 5: Use the short-run cost elasticity to calculate short-run volume variable cost.**

Once the short-run cost elasticity has been determined for a component, it can be used to estimate the component's short-run volume variable cost. This step is relatively straightforward because it is embodied directly in the Postal Service's product cost model and does not require any restructuring of the model. The short-run volume variable cost in a component is just the product of the accrued cost in the component and the short-run cost elasticity. The existing distribution key can be used to attribute the short-run volume variable cost to the products that caused it to arise.

**Step 6: Combine the short-run volume variable costs by component, where applicable, with the remaining long-run volume variable cost by component to calculate the short-run unit volume variable cost by product.**

It is quite likely that not all components will need adjustment to account for an inflexible input. If so, those components will still make use of their long-run cost elasticities in calculating unit volume variable cost. This means that the calculated short-run marginal cost will reflect a mixture of components in which there is a short-run cost elasticity required to estimate volume variable cost and those in which the long-run cost elasticity continues to be used. Once the volume variable costs by product are added across components, their sum can be divided by the respective volume to find the short-run unit volume variable cost.<sup>26</sup>

***Informational Requirements for Calculating Short-Run Marginal Cost***

The calculation of short-run marginal cost (unit volume variable cost) places additional informational requirements on the Postal Service costing system as the current system was not developed to estimate short-run costs.

Those additional requirements would at a minimum include:

- Identification of the constraints that preclude the optimal adjustment of inputs.
- Identification of the inputs that have not been optimally adjusted to volume changes and an enumeration of the reasons that adjustment did not take place.
- Empirical evidence demonstrating that certain inputs have not been adjusted optimally in response to volume changes.
- Data on the amount of the inputs that have not been completely adjusted along with historical data on their values in higher volume periods.
- Data on the capacity utilization and/or productivity for the components affected by lack of adjustment including identification of the long-run capacity and/or productivity.
- Data and information required to estimate short-run cost elasticities.

*The calculation of short-run marginal costs will place additional informational requirements on the Postal Service costing system and some of these data may be challenging to obtain.*

A review of these data requirements reveals that some of them may be challenging to obtain. That is because they require operational information that might not be collected in the normal course of business. In addition, the list includes some historical data that may be currently unavailable or difficult to obtain from a data archive. These issues

---

<sup>26</sup> This is mathematically equivalent to calculating the unit volume variable costs by component and then summing them.



suggest that the calculation of short-run costs would take significant effort to obtain essential inputs from field operation experts as well as costing experts.<sup>27</sup>

### ***A Process for Review and Evaluation of Short-Run Costs***

By definition, the short run is a “temporary” condition in the sense that adjustment in either the firm’s input mix or a change in the demand for its product will return the firm to its long-run equilibrium condition. In addition, the nature of the short-run deviation from the firm’s long-run equilibrium depends upon the same two factors: the firm’s input mix and the demand for its products. Because both of these factors have the potential to change through time, the short run is a relatively fluid condition and the resulting short-run costs may change frequently through time. This suggests that a two-prong process of review and evaluation of short-run costs is appropriate.

*A two-prong process of review and evaluation of short-run costs is appropriate. The first prong is annual review of the assumptions and calculations used for measuring short-run costs. The second prong is an “early warning system” to identify material changes in inputs or product demands.*

The first prong is an annual review of the assumptions and calculations associated with constructing short-run costs. As the Postal Service’s input mix changes and as the volumes it handles change, the nature and degree of its “short run” will also change.

Consequently, the Postal Service should carefully review each step in its computational algorithm for short-run marginal cost and should estimate the

degree to which a change in conditions warrants a recalculation of costs. The Postal Service currently calculates its long-run cost on an annual basis. It would be logical for the Postal Service to review and possibly recalculate its short-run cost at the same frequency. However, it is not necessary, and may not be preferable, to calculate both short-run and long-run costs at the same time. This depends upon the resources available and the difficulty in obtaining the information needed to calculate short-run costs.

The second prong of the short-run cost review process is an “early warning system” designed to identify any material changes in inputs or product demands that could affect short-run costs. For example, if the Postal Service is able to consolidate a number of mail processing plants over a six-month period, it may be able to move toward its long-run equilibrium position. If so, a reevaluation of the assumptions and computations underlying the calculation of short-run costs would be appropriate. The early warning system would consist of a listing of the key impediments, amounts of unadjusted inputs and determinations of excess capacities that would serve as the basis for an evaluation of any changes in input mix and product demand. Any operational or regulatory changes that would materially affect any of the quantities on the list should trigger a review of short-run marginal cost calculation.

---

<sup>27</sup> While the Postal Service did provide estimates of short-run costs in its summer sales, these estimates did not involve the level of sophistication discussed in this section. See Appendix C for a discussion on the Summer Sale methodology.

## Conclusion

Firms that are unable to adjust all of their inputs to their cost-minimizing levels can find it useful to consider using short-run costs as a basis for price setting. The Postal Service appears to be operating under short-run conditions, as it has multiple impediments to moving to its optimal operating environment and the demand for its products has been declining through time. Therefore, it might make sense for the Postal Service to consider calculating and using short-run marginal costs to set prices for at least some of its products.

However, there are some material risks associated with this approach. It is important that the Postal Service is confident that the change in prices associated with short-run costs will actually increase total revenue and that this increase in revenue will be high enough to cover short-run variable costs. If the price change would not meet these two conditions, it is quite possible that the Postal Service's overall financial condition could be worsened by the use of short-run costs. This means that if the Postal Service decides to move towards short-run costing and pricing, it is essential that it devote sufficient resources to accurately measure both short-run costs and consumer response to the price change.

In addition, it is important to keep in mind that even if implemented correctly, using short-run costs and prices is only a temporary fix and not a long-term solution. The use of short-run prices does not solve the more fundamental problems that the Postal Service faces. It is critical that the Postal Service still move towards optimizing its network in order to improve its financial position.

## Appendices

## Appendix A      The Theoretical Basis for Short-Run Costs

### *The Role of Time in Defining Costs*

The distinction between long run and short run is not formally defined by time. Instead it is defined by a firm's ability to adjust its inputs. A firm is in the long run when it is able to adjust all inputs to their optimal levels. However, on a practical basis, the distinction between long run and short run is intimately related to time. The ability of a firm to adjust its inputs is linked to the time it takes to make the adjustments. The greater the time span, the more inputs are adjustable. For example, the more facilities can be added or downsized, the more machines can be added to or taken out of a plant. Theoretically, if enough time is allowed to pass, the firm will be able to adjust all of its inputs. This simple idea is the basis for the following statements about the long run, which appear in most microeconomics textbooks, and other places:

- In the long run, all inputs are variable.
- In the long run, there are no fixed costs.
- The long run is defined as a situation in which a firm can adjust all of its inputs.

However, recent analysis has shown that the issues surrounding the adjustment of inputs are a bit more complex than suggested by the textbook treatment.<sup>28</sup> To reflect this additional complexity, the traditional dichotomy between fixed and variable costs has been expanded. In this more thorough classification, costs are first determined to be either "sunk" or "avoidable." A sunk cost is irrevocable and must be paid whether the firm operates or not. More formally, if we let  $y$  represent a firm's level of output and  $k$  represent a constant, then a sunk cost is defined by the following condition:

$$C(y) = k \text{ for all } y \geq 0.$$

Note that the definition implies that a sunk cost must be paid even if the firm produces no output.<sup>29</sup>

An avoidable cost, by contrast, is only paid when the firm produces a positive level of output. Avoidable costs can be fixed or variable. A fixed avoidable cost is the cost created by the use of a fixed input whose value does not vary with the level of output.

---

<sup>28</sup> See for example, R.L. Sexton, P.E. Graves, and D.R. Lee, "The Short and Long Run Marginal Cost Curve and Pedagogical Note," *Journal of Economic Education*, Winter 1993, pp. 34-37 or X.H. Wang and B.Z. Yang, "Fixed and Sunk Costs Revisited," *Journal of Economic Education*, Spring 2001, pp. 178-185.

<sup>29</sup> It is true that the firm may eventually be released from the cost through the passage of time (e.g., the term of a fixed lease may run out), but this has nothing to do with whether the firm produces any output.

An avoidable fixed cost is defined as:

$$C(y) = k \text{ for all } y > 0.$$

Finally, an avoidable variable cost has zero value when the firm does not produce any output and varies with changes in the level of output produced.

This broader approach to cost classification recognizes that some of the costs that a firm incurs only when it produces a positive level of output will not vary with level of output no matter how much time passes. This means that it allows for costs to be fixed even in the long run. This contradicts the traditional statement that there can be no fixed costs in the long run and rejects the lack of sufficient time to adjust an input as the sole basis for justifying its fixity. Instead, it suggests that the amount of an input may not be perfectly adjustable in the long run because it is indivisible. For example a restaurant may need to obtain a license to operate, the cost of which is independent of the number of meals served. Because this license must be obtained every year the restaurant operates, it is a long-run cost. Because it can be avoided if the restaurant shuts down it is an avoidable cost and because it does not depend upon the level of production it is a fixed cost. In sum, it is a long-run fixed cost.

In addition, indivisibility refers to the situation in which an input cannot be broken into sufficiently small increments so as to provide the optimal amount of the input. For example, the operator of a single-plane jet service between two cities would not be able to adjust the size of the jet as the number of passengers varies because the jet is indivisible. In this example the cost of the airplane is a fixed cost. But a fixed cost may be sunk or avoidable and additional investigation is required to determine which. In the case of the jet service operator, the fixed cost would likely be avoidable because if the jet were not used for service between the two designated cities, it is likely to be useful for another purpose. This means the operator could sell the jet and avoid incurring its cost. If, on the other hand, the jet could not be sold, then the fixed cost becomes sunk as the operator must pay the cost whether or not it provides the jet service.

In sum, all sunk costs are fixed but not all fixed costs are sunk. Total costs are defined as the sum of sunk fixed costs (SFC), avoidable fixed costs (AFC) and variable costs (VC). This definition of fixed costs can also be used to show the two-way classification of costs, by technology (fixed vs. variable) and by incurrence (sunk vs. avoidable):

$$\begin{array}{c}
 \text{Fixed} \qquad \qquad \text{Variable} \\
 \text{TC} = \underbrace{\text{SFC} + \text{AFC}}_{\text{Fixed}} + \underbrace{\text{VC}}_{\text{Variable}} \\
 \underbrace{\text{SFC}}_{\text{Sunk}} \quad \underbrace{\text{AFC} + \text{VC}}_{\text{Avoidable}}
 \end{array}$$

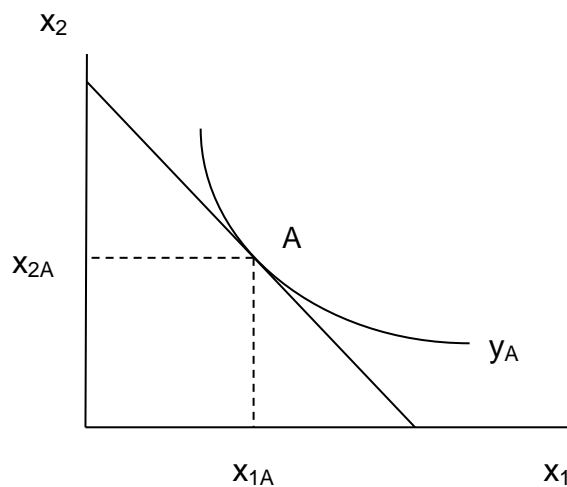
### ***The Short-Run and Long-Run in Production and the Implications for Cost***

Long-run production is the situation in which a firm produces its vector of outputs at the lowest total cost, because it is free to adjust all of its inputs to their optimal levels. In

technical terms this occurs at the tangency of the firm's isoquants (combinations of inputs that produce the same level of output) and isocost lines (combinations of inputs that cost the same amount). In other words, the firm is choosing the combination of inputs that cost the least (smallest isocost line) and can still produce the desired level of output (the firm is on the desired isoquant).

Short-run production involves fixing the quantity of at least one of the inputs and then choosing the optimal amount of the remaining inputs. This means that the short-run total cost (SRTC) can be no smaller than the long-run total cost (LRTC) and that the two cost measures will be equal only when the amount of the fixed input happens to equal the optimum amount for the level of output being produced. Such a situation is illustrated in Figure 2 in which the desired level of output,  $y_A$ , is being produced with the cost-minimizing set of inputs  $x_{1A}$  and  $x_{2A}$ .

**Figure 2: The Cost-Minimizing Choice of Inputs**



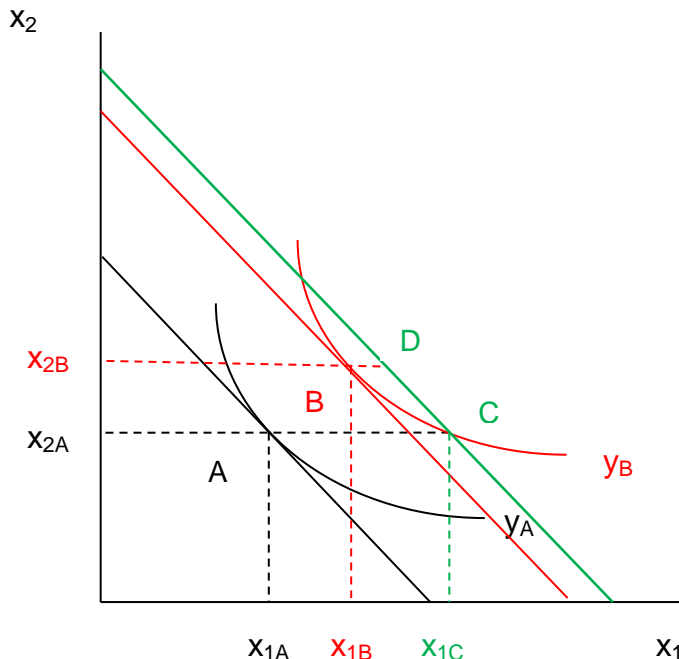
If the amount of  $x_2$  happens to be fixed at  $x_{2A}$ , then the short-run and long-run costs of production will coincide. However, this is true only for output level  $y_A$ . If the desired level of output varies, then SRTC will exceed the LRTC. Suppose that the desired level of output expands beyond  $y_A$  but the level of the second input remains fixed at  $x_{2A}$ . The impact on cost is illustrated in Figure 3.

Suppose that the desired level of output is increased to  $y_B$ . The long-run response is to increase the amounts of both inputs until once again there is a tangency between the new (higher) isocost line and the new (higher) isoquant. In Figure 3, the new combination of inputs occurs at point B, with the amounts of the two inputs set at  $x_{1B}$  and  $x_{2B}$ , respectively.

Now consider the case in which the level of  $x_2$  is fixed at amount  $x_{2A}$ . The firm now adds  $x_1$  until it reaches the desired isoquant and is producing  $y_B$ . In Figure 3, this combination of inputs occurs at point C. The important point for cost incurrence is that while points B and C both occur on the same isoquant, point C occurs on a higher isocost line than

point B.<sup>30</sup> This demonstrates that the SRTC of producing  $y_B$  is greater than the LRTC of producing the same level of output.

Figure 3: Short-Run and Long-Run Response to an Increase in Output



Also note that the increase in cost is smaller for the long-run response. Figure 3 demonstrates that the long-run response requires moving to the isocost line that runs through point B. The short-run response requires moving to the isocost line that runs through point C. This isocost line also runs through point D, which involves a higher cost than point B.<sup>31</sup> This larger movement demonstrates that the short-run marginal cost (SRMC) is greater than the long-run marginal cost (LRMC). Thus, when the short run occurs because there is insufficient capacity (too little of the fixed input) both the SRTC and the SRMC are greater than their long-run counterparts.

Less intuitive is the situation in which there is excess capacity of the fixed input. For example, this situation would occur if the firm started at an output level with the optimal combination of  $x_1$  and  $x_2$ , experienced a drop in demand for its output and was unable to adjust the amount of  $x_2$  that it uses. This situation is illustrated in Figure 4. Again, we start the firm at output level  $y_A$ , at which the amount of input  $x_2$  being used just equals its optimal amount. We then analyze the situation in which the level of output drops below  $y_A$ , as indicated by an inward shift of the isoquant to output level  $y_B$ . As was the case for an increase in output, the long-run response to a decline in output involves

<sup>30</sup> The further a parallel isocost line is from the origin, the higher its associated cost of production. With fixed input prices, total cost strictly increases as the amounts of inputs used increases.

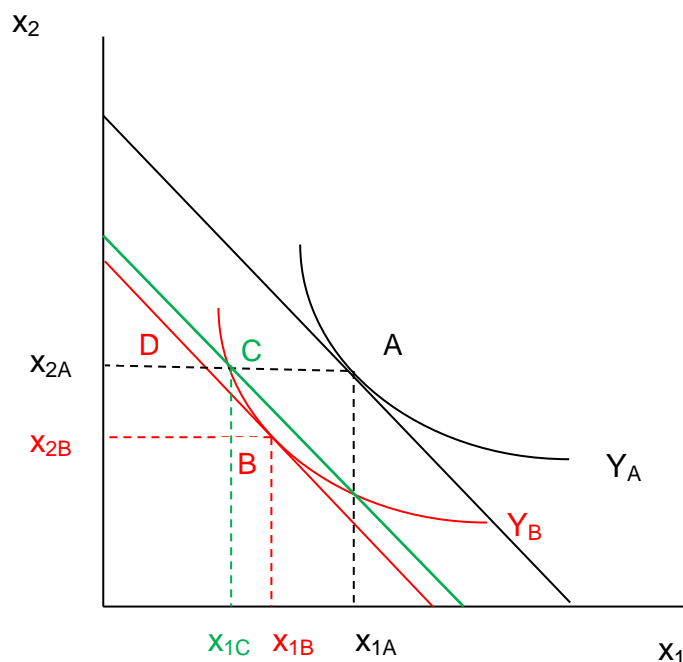
<sup>31</sup> Point D requires the same amount of input  $x_2$  and more of the input  $x_1$  than does Point B.

using the new, reduced, optimal amounts of both  $x_1$  and  $x_2$ . In Figure 4, the new long-run input combination occurs at point B.

In contrast, if the firm is constrained to continue to use the fixed amount of the second input,  $x_{2A}$ , the short-run response occurs at point C, which is the intersection of input level  $x_{2A}$  and the new isoquant,  $y_B$ . Because the firm cannot reduce the amount of  $x_2$  that it is using, it responds with a large reduction in  $x_1$ .

The isocost line associated with the short-run response, running through point C, lies further from the origin than the isocost line associated with the long-run response, which runs through point B. This means, like the case of expanding output, the SRTC for producing  $y_B$  will exceed the LRTC for producing the same level of output. Note that the short-run response requires using less of the first input ( $x_{1C}$ ) than does the long-run response ( $x_{1B}$ ). This is because the firm is using more of the second input in the short-run response.

**Figure 4: Short-Run and Long-Run Response to a Decrease in Output**



Although the SRTC exceeds its long-run counterpart, the short-run marginal cost is less than the long-run marginal cost in the case of excess capacity. The short-run response occurs at point C and the implied short-run change in cost is the difference between the total cost implied by the isocost line running through point A and the isocost line running through point C. The long-run change in cost is the difference between the cost implied by the isocost line running through point A and the isocost line running through point B. But that isocost line also runs through point D, which is further away (horizontally) from



point A than is point C. Thus, the change in cost in the long run is greater than the change in cost in the short run.

The intuition behind this result is that by being able to reduce both inputs the firm can save more in the long run, following a volume decline, than it can in the short run. Thus, the long-run marginal cost is larger than the short-run marginal cost. But the result works in the opposite direction, too. After an increase in output (Figure 3), it is cheaper for the firm to move from point C to point A (short run) than it is to move from point B to point A (long run). This means in a condition of excess capacity the short-run marginal cost is less than the long-run marginal cost for increases as well as decreases in output.

### ***The Relationship among Long-Run and Short-Run Cost Measures***

The usefulness of short-run cost concepts as a basis for pricing depends, in part, upon their relationship to long-run cost concepts. In addition, determining the profitability of a pricing scheme based upon short-run costs may depend on the relationship among short-run cost measures. For these reasons it is important to examine and articulate the relationship among the familiar cost measures in the short run and long run.

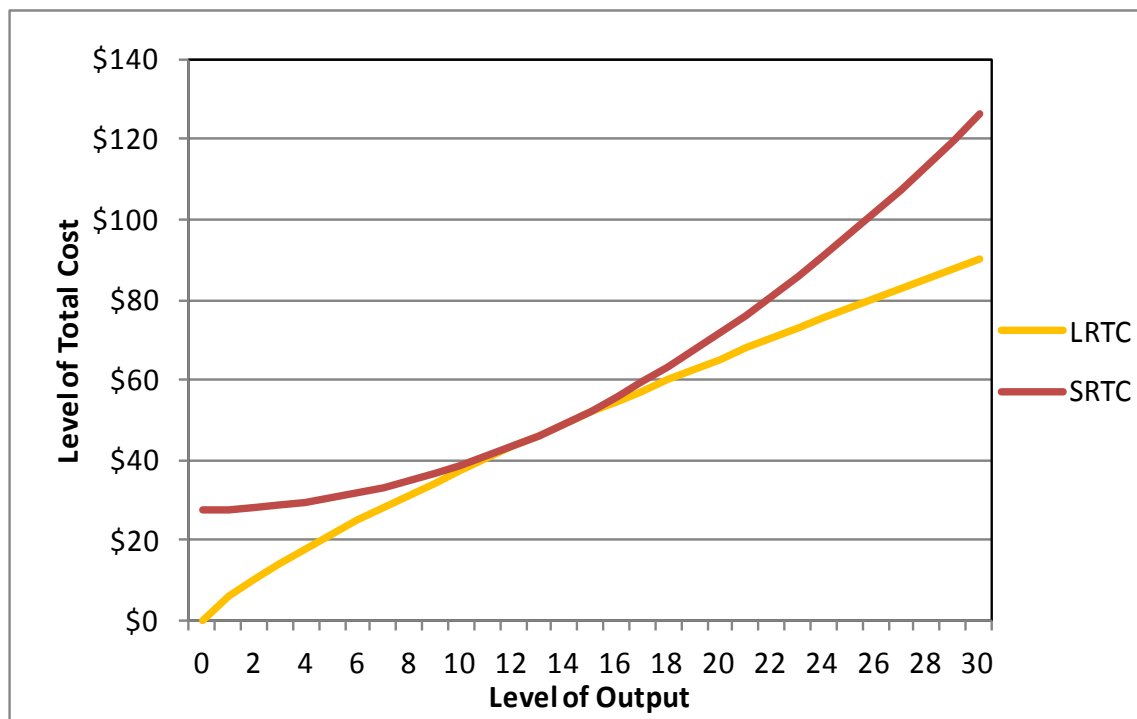
As demonstrated above, the long-run cost of producing a level of output is associated with the firm using the cost-minimizing combination of inputs. This means that LRTC is the lowest possible cost for producing any level of output. The “best” that a short-run cost function can do is to match the long-run cost function at the level of output for which the amount of the fixed input just equals the optimal amount. For all other levels of output, the SRTC function will generate a higher level of cost and SRTC will exceed LRTC. This means that SRTC is never less than LRTC.

This relationship is highlighted for a firm with Cobb Douglas production and increasing returns to scale in Figure 5,<sup>32</sup> where the level of output at which the amount of the fixed input equals its optimal value is 13 units. At that level of output the two cost curves are tangent; at all other points the SRTC lies above the LRTC.

---

<sup>32</sup> This will occur when the sum of the coefficients on the inputs in the production function exceeds one. The Cobb Douglas production function is presented in the mathematical derivation at the end of this appendix. That derivation also presents the short-run and long-run cost functions associated with Cobb Douglas production.

Figure 5: Total Cost Curves

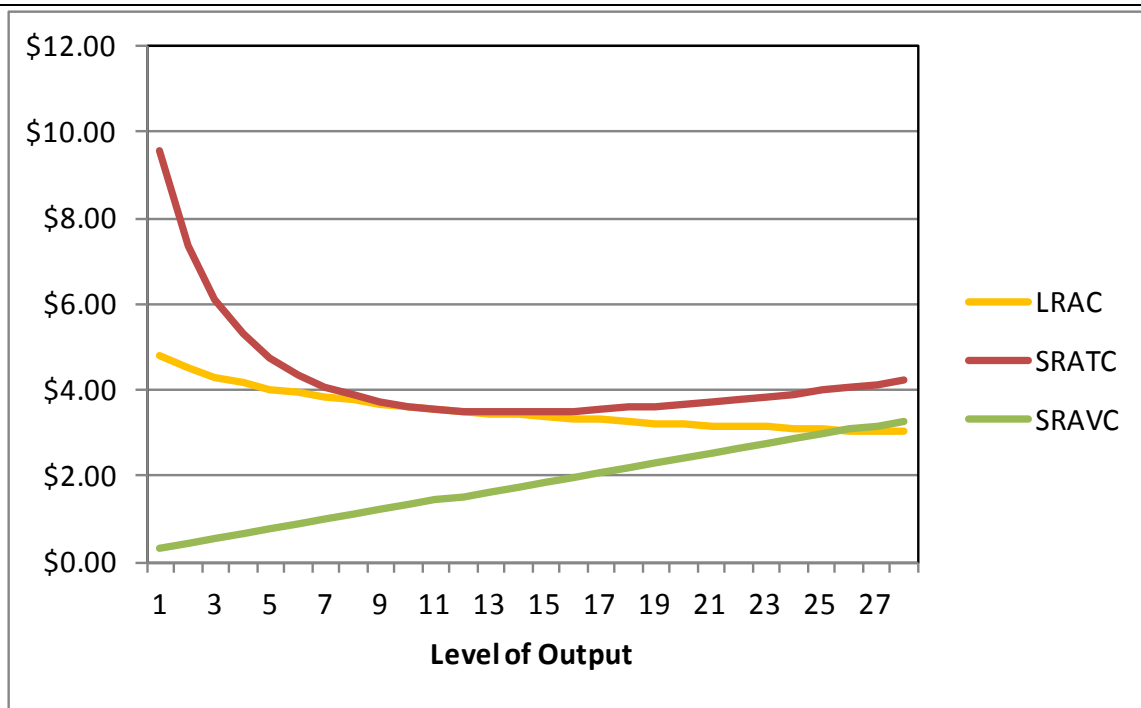


The relationship between the SRTC and LRTC curves also determines the relationship between average total cost curves. Average total cost is just the ratio of total cost to output, so whenever SRTC exceeds LRTC, short-run average total cost (SRATC) also exceeds long-run average total cost (LRATC).

Less obvious is the relationship between long-run average variable cost (LRAVC) and short-run average variable costs (SRAVC). This relationship is most easily seen in a case where there are no long-run fixed costs. In this case, long-run average total costs equal long-run average variable costs and can be labeled just long-run average costs (LRAC). Because the short run arises due to the fact that at least one input is fixed, it is possible that the short-run average variable cost (SRAVC) could be less than the long-run average cost (LRAC). This would occur, for example, when the amount of the fixed input is large relative to the level of output and only a small amount of the variable inputs would be needed to produce the desired level of output in the short run.

This possibility is illustrated in the Figure 6 for the case of the Cobb Douglas production with one fixed input. At low levels of output, the short-run average variable cost is below the long-run average cost while the short-run average total cost is well above the long-run average cost. This illustrates the role of the fixed input in determining the nature of short-run costs.

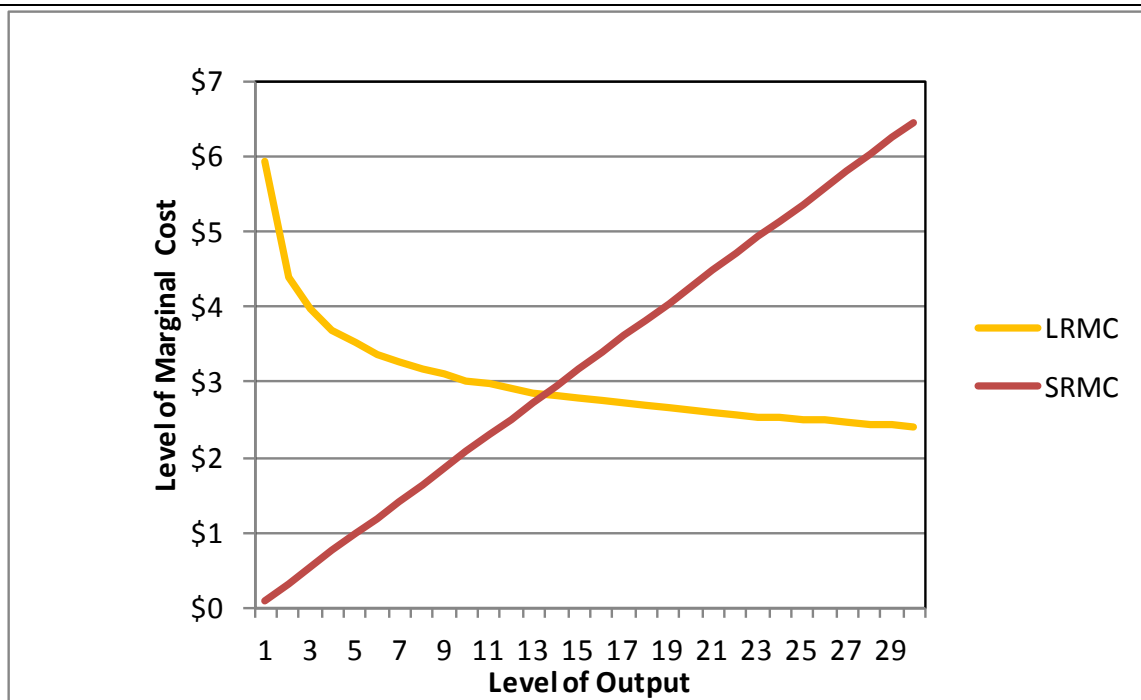
Figure 6: Average Cost Curves



Because of the role that marginal cost plays in pricing, the cost relationship that gets the most attention is the one between long-run marginal cost (LRMC) and short-run marginal cost (SRMC). As demonstrated above, when the amount of one input is constrained, the relationship between the short-run and long-run marginal costs depends upon whether output is above or below that associated with the optimal capacity of the fixed input. In cases where the output is above where the fixed input would be optimal, the short-run marginal cost will exceed the long-run marginal cost but in cases of where the output is below where the fixed input would be optimal, the long-run marginal cost will exceed the short-run marginal cost.

Graphically, this means for any output level below the one for which the amount of the fixed input is optimal, the SRMC will lie below the LRMC curve. This is roughly associated with the condition of “excess capacity.” The opposite condition will occur for output levels above the one for which the amount of the fixed input is optimal. This is roughly associated with the condition of insufficient capacity. These conditions are illustrated in Figure 7.

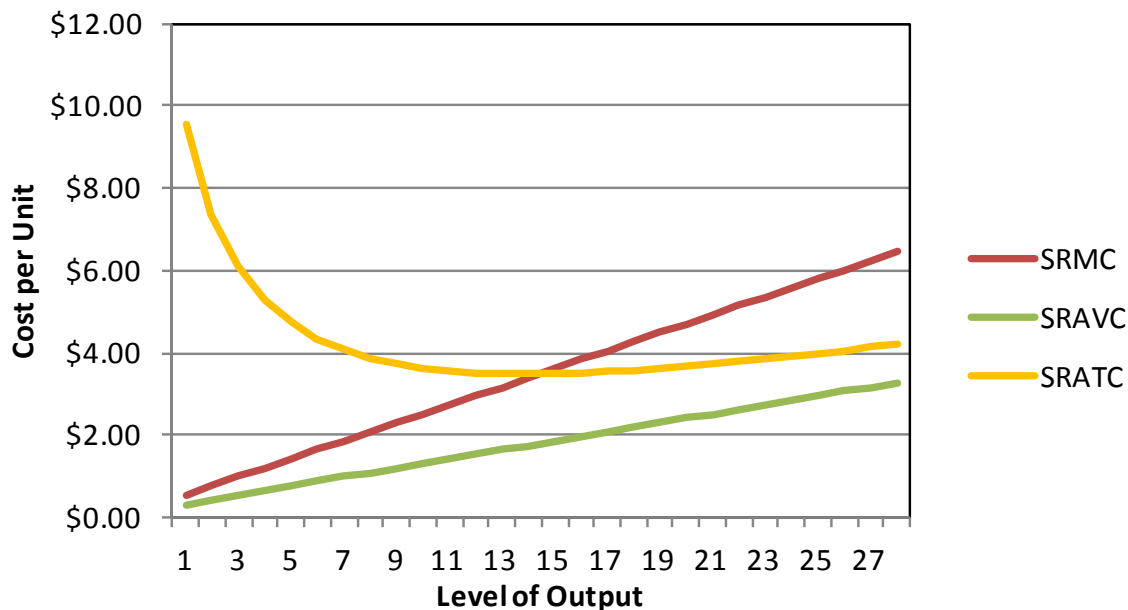
Figure 7: Long-Run and Short-Run Marginal Costs



The last set of conditions that needs to be reviewed is the set that describes the relationships among the short-run cost curves. These conditions are important because they determine whether or not pricing at SRMC will lead to an increase or decrease in profits for the firm. Simply because a firm's price is above its SRMC does not ensure that the price is sufficiently high so that the firm is earning a profit on the units sold. For a positive profit to be earned, the firm must also charge a price that is above its short-run average total cost (SRATC). However, in the short run, the firm may want to produce and sell even if the price does not quite reach this level. If the firm's price is above its short-run average variable cost (SRAVC) then its sales will be making a contribution to covering the firm's fixed cost, associated with the firm's fixed input. While this does not provide a sustainable price for the firm in the long run, it does allow it to reduce its losses in the short run.

The relationship among these three short-run curves depends upon the nature of the production process. In most instances, the existence of a fixed input will subject the other inputs to diminishing returns (declining marginal product). If so, the SRMC curve will lie below the SRATC at output levels below its minimum and will lie above the SRATC at output levels above its minimum. In addition, because it is increasing everywhere, the SRMC curve will lie above the SRAVC curve. This case is illustrated in the Figure 8 for Cobb Douglas production. Note that in this case, if the firm sets a price equal to or above its SRMC, it is assured of at least covering its short-run variable costs.

Figure 8: Short-Run Cost Curves



While the existence of a fixed factor means that diminishing returns will eventually set in, it may not occur over the relevant range of production. If the technology is such that the variable factors are experiencing increasing marginal product over the relevant range, then it is possible for the SRMC to lie below both the SRATC and the SRAVC cost curves.

### **Mathematical Derivation**

The general form for long-run production with two inputs is given by:

$$y = f(x_1 x_2).$$

Using the Cobb-Douglas function form yields<sup>33</sup>:

$$y = x_1^\alpha x_2^\beta.$$

The long-run cost function is derived from minimizing the cost of any level of output given the Cobb Douglas technology:

$$\min C = \omega_1 x_1 + \omega_2 x_2, \text{ s.t. } y = x_1^\alpha x_2^\beta.$$

Solving this problem yields the following set of first order conditions:

$$\frac{\partial C}{\partial x_1} = \omega_1 - \lambda \alpha x_1^{\alpha-1} x_2^\beta = 0.$$

$$\frac{\partial C}{\partial x_2} = \omega_2 - \lambda \beta x_1^\alpha x_2^{\beta-1} = 0.$$

$$\frac{\partial C}{\partial \lambda} = y - x_1^\alpha x_2^\beta = 0.$$

Solving the first order conditions yields the input demand functions:

$$x_1 = y^{\frac{1}{\alpha+\beta}} \left[ \frac{\alpha \omega_2}{\beta \omega_1} \right]^{\frac{\beta}{\alpha+\beta}}.$$

$$x_2 = y^{\frac{1}{\alpha+\beta}} \left[ \frac{\beta \omega_1}{\alpha \omega_2} \right]^{\frac{\alpha}{\alpha+\beta}}.$$

The input demand functions can be substituted back into the cost equation to produce the long-run cost function that relates total cost to the level of output and input prices:

$$C(y) = \left[ \omega_1 \left( \frac{\alpha \omega_2}{\beta \omega_1} \right)^{\frac{\beta}{\alpha+\beta}} + \omega_2 \left( \frac{\beta \omega_1}{\alpha \omega_2} \right)^{\frac{\alpha}{\alpha+\beta}} \right] y^{\frac{1}{\alpha+\beta}}.$$

---

<sup>33</sup> Examination of the Cobb-Douglas functional form is useful because it permits analysis of both constant returns to scale and increasing returns to scale, both of which occur in the postal context.

Now suppose the amount of  $x_2$  that the firm uses is fixed. The production function will reflect this fixity (the bar indicates that the amount of  $x_2$  is constant):

$$y = x_1^\alpha \bar{x}_2^\beta.$$

With just one variable input, there is just one input demand function derived through cost minimization:

$$x_1 = y^{\frac{1}{\alpha}} (\bar{x}_2)^{-\frac{\beta}{\alpha}}.$$

This expression can be used to derive the short-run cost function:

$$C(y) = \omega_1 y^{\frac{1}{\alpha}} (\bar{x}_2)^{-\frac{\beta}{\alpha}} + \omega_2 \bar{x}_2.$$

These results can be used to derive expressions for the familiar short-run and long-run cost measures.

$$LRMC = \left( \frac{1}{\alpha + \beta} \right) y^{\frac{1-\alpha-\beta}{\alpha+\beta}} \left[ \omega_1 \left( \frac{\alpha\omega_2}{\beta\omega_1} \right)^{\frac{\beta}{\alpha+\beta}} + \omega_2 \left( \frac{\beta\omega_1}{\alpha\omega_2} \right)^{\frac{\alpha}{\alpha+\beta}} \right].$$

$$LRAC = y^{\frac{1-\alpha-\beta}{\alpha+\beta}} \left[ \omega_1 \left( \frac{\alpha\omega_2}{\beta\omega_1} \right)^{\frac{\beta}{\alpha+\beta}} + \omega_2 \left( \frac{\beta\omega_1}{\alpha\omega_2} \right)^{\frac{\alpha}{\alpha+\beta}} \right].$$

$$SRMC = \frac{1}{\alpha} \omega_1 y^{\frac{1-\alpha}{\alpha}} (\bar{x}_2)^{-\frac{\beta}{\alpha}}.$$

$$SRAVC = \omega_1 y^{\frac{1-\alpha}{\alpha}} (\bar{x}_2)^{-\frac{\beta}{\alpha}}.$$

$$SRATC = \omega_1 y^{\frac{1-\alpha}{\alpha}} (\bar{x}_2)^{-\frac{\beta}{\alpha}} + \omega_2 \bar{x}_2 y^{-1}$$

## **Appendix B      Questions Associated with Basing Prices on Short-Run Costs**

When considering the use of short-run costs as a basis for prices, firms must consider both the size of the potential benefits from their use along with the likelihood of achieving those benefits. That likelihood depends upon factors like how observable and measurable is the firm's excess capacity and how much is known about consumer response to price changes. To make an informed decision, the firm must carefully analyze and assess these important issues. Yet that is not the end of the inquiry. Suppose that a firm does a careful evaluation and decides it would be in its benefit to set prices based upon short-run costs. There still are a number of important questions related to implementing the approach and these questions are addressed in this section.

### ***How Do Short-Run Costs Fit Into an Overall Pricing Strategy?***

Short-run costs are potentially useful when a firm is facing a period of time during which it cannot "right-size" at least one of its inputs. This can be a period of either excess capacity or insufficient capacity. In either case, when a firm is in this situation, it faces the choice of changing its price in an attempt to match the demand for its product to its current capacity to produce that product. In general, this is a temporary price change that will last only until the firm's optimal capacity meets the market demand for its product.

Short-run costs provide a basis for determining and evaluating those short-run prices. In addition, short-run costs provide a method of quantifying the impact of excess or insufficient capacity and provide the firm with information about the opportunity to mitigate the negative financial consequences of not being able to adjust all of its inputs. Short-run costs are also useful for tracking the change in unit costs that take place as the firm changes prices and its level of output. For example, short-run marginal cost tracks how a product's unit cost changes as the firm makes better use of its excess capacity. This is essential information for deciding if short-run prices should be adjusted. In sum, short-run costs are essential pieces of information for helping a firm decide if it is in its interest to deviate from its long-run prices.

### ***What Measures of Short-Run Cost Should Be Used?***

Short-run marginal cost is the starting point for short-run pricing as it measures the additional resources required to produce and sell any additional output. At a minimum, short-run prices must exceed short-run marginal costs or the firm should not sell any additional output. However, it is possible that a firm's prices could meet this standard and it still should not produce and sell any additional output. This would occur if the firm's price exceeds its short-run marginal cost but does not exceed its short-run average variable cost. In this situation, the price being charged is so low that the revenue earned on the product is less than the variable cost incurred in providing it.



This means that the firm is increasing, not reducing, its losses by providing the product.<sup>34</sup>

In a multiproduct firm, average variable cost does not exist, so the firm needs to check whether the short-run price exceeds the product's short-run average incremental cost (SRAIC). The SRAIC is the total additional cost that the firm incurs in providing the total amount of a product. It can be found by comparing what it costs the firm, in the short run, to provide its full set of products with what it costs the firm, in the short run, to provide the total volume ( $V$ ) of all its products except the one being studied. This is the difference between its SRTC of all products and its SRTC of all products except the one being studied and provides the product's short-run incremental cost (SRIC). To find the product's SRAIC, the SRIC must be divided by the product's volume.

Mathematically, the firm's SRAIC for product "A" is given by:

$$SRAIC_A = \frac{SRTC(V) - SRTC(V - V_A)}{V_A}$$

In sum, the firm should use a product's short-run marginal cost as the basis for setting price and then check to ensure that the chosen price exceeds the product's short-run average incremental cost.

### ***What Is the Correct Time Horizon for Using Short-Run Costs?***

Short-run costs arise because of the inability of the firm to adjust one of its inputs. Therefore, theoretically, the short run exists only until that input is adjusted to its optimal level. This means that the correct time horizon for using short-run costs is the period until the firm has been able to adjust all of its inputs.

On a practical level, determining the appropriate time horizon depends upon the reason that the short run occurred. For example, if the short run arises because of fluctuating demand, then the short run is defined by the period over which the fluctuations take place. Suppose there is a seasonal pattern to demand so that demand is abnormally low in the summer and the firm cannot adjust capacity to match demand for that period. Then, the time horizon for short-run costs would be the summer period until volume returns to its normal level. Note that in this situation, the short run would recur each year.

In contrast, suppose the short run arose because of accidental overcapacity that arose because the firm overestimated how quickly the demand for its product would grow when purchasing a large capital input. In this case, the time horizon for the short run is

---

<sup>34</sup> Note that because of fixed cost, the firm might not be earning a positive profit on a product even if its price exceeds its average variable cost. To make a positive profit, a product's price must exceed its average total cost, which includes its average fixed cost. However, as long as a product's price exceeds its average variable cost, the firm should produce and sell it, even if the price does not cover the product's average total cost. This is because when the price exceeds average variable cost, the firm is earning a surplus that can be used to pay the fixed cost. Simply because that surplus is not big enough to cover the entire fixed cost does not mean that it is not beneficial for the firm.

the period until capacity is adjusted because either demand grows sufficiently to meet the firm's capacity or the firm reduces its capacity to meet demand. Lastly, if excess capacity is caused by declining demand, the short run may be ongoing if the firm does not adjust capacity or does not adjust capacity enough to match the fall in demand.

### ***Is the Short Run the Same for All Products?***

The short run is defined by a condition – the inability to adjust – that relates to one of the firm's inputs, not to its outputs. This means that any product that uses this inflexible input will be in the short run. Whether or not this is all of the firm's products depends upon the technology of production and the degree to which they share inputs. It is possible for one subset of the firm's products to make use of the affected input and another subset to not use it. In that case, only some of the firm's products would be in the short run, while others would be in the long run.

### ***Is It Appropriate to Base Some Product Prices on Short-Run Costs and Other Product Prices on Long-Run Prices?***

The firm's choice to use short-run costs to set prices depends upon both conditions within the firm and conditions relating to the firm's customers. If the internal or external conditions differ by product, then it could be appropriate to base some of the firm's prices on short-run costs while other prices are based upon long-run costs. One such situation arises when only a subset of the firm's products makes use of the inflexible input. The prices for that subset of products could then be based on short-run costs while the rest of the firm's products are based upon long-run costs.

Even if all the firm's products share the same technology, it is still possible to envision a situation in which just a subset of the firm's prices was based on short-run costs. Suppose a firm had two products, one of which had very price-sensitive customers and the other of which had very price insensitive customers. Further suppose the firm has too much of an input that is used by both goods. That means each good's short-run marginal cost would be less than the corresponding long-run marginal cost. But, the firm might use short-run pricing only for the first good, which has price-sensitive customers. Lowering the price for the price-sensitive good would expand demand and help the firm make use of its excess capacity.<sup>35</sup> It would also help defray the associated fixed cost. However, lowering the price for the other good would simply lower the firm's revenue and would not expand output. The price for that good should continue to be based upon its long-run cost.

---

<sup>35</sup> To the extent that the Postal Service is able to lower a price for a particular subset of customers that are sensitive to price declines (in order to maximize its revenue gain from this change), it will have to be able to effectively limit the price change to just the price-sensitive subset of customers.

## Appendix C Practical Issues Measuring Short-Run Marginal Costs

In the main paper, we provided the analytical basis for calculating short-run marginal cost and presented and explained a computational algorithm for doing so. However, there are a number of practical issues related to calculating these costs that bear discussion. We present these practical issues in this appendix.

### ***What Is the Relationship Between Short-Run Marginal Costs and the Postal Service's Volume Variable Costs?***

The Postal Service's product cost model calculates a measure called "volume variable cost per piece" for postal products.<sup>36</sup> It can be shown that this measure provides an estimate of the marginal cost for those products. However, an important issue for this analysis is whether the Postal Service is estimating short-run or long-run marginal cost with its volume variable cost measure.

The Postal Service normally calculates its volume variable costs over a three- to four-year time horizon, which allows it to vary a wide range of inputs in response to demand changes. In addition, a review of the inputs which the Postal Service assumes respond to changes in volume include buildings, vehicles, supervisory labor, and equipment. These are the types of inputs that are typically held fixed in a short-run analysis. On the other hand, the Postal Service's cost system does include some fixed costs and only about 60 percent of its total cost is considered volume variable.

Taken together, the facts indicate that the Postal Service's cost system is measuring long-run marginal cost.<sup>37</sup> Although 40 percent of total cost is not attributed to products, very little of that amount is due to fixed cost. Nearly all of the 40 percent is made up of "network costs" generated by economies of scale or density, which is consistent with long-run costing. Moreover, those costs that are assumed to be fixed with respect to volume (like the fixed part of its delivery network) would continue to be fixed even over any time horizon. Finally, both the Postal Service and the Postal Rate Commission have identified the Postal Service's product costs as estimating long-run marginal costs.<sup>38</sup>

### ***Previous Work Calculating Short-Run Marginal Costs***

The Postal Service has already calculated a version of short-run marginal costs. In May 2009, the Postal Service filed a notice with the Postal Regulatory Commission of its intention to change prices for Standard Mail in the form of a program entitled, "Standard

---

<sup>36</sup> For a more detailed analysis of how these costs are calculated see U.S. Postal Services Office of Inspector General, *A Primer on Postal Costing Issues*, RARC-WP-12-008, March 20, 2012, [http://www.uspsoig.gov/foia\\_files/RARC-WP-12-008.pdf](http://www.uspsoig.gov/foia_files/RARC-WP-12-008.pdf).

<sup>37</sup> We note that the Postal Service estimates practical long-run costs, not the theoretically ideal long-run costs. The Postal Service approach is to allow for adjustments in its labor and capital inputs, but there is no guarantee that the adjustments lead to the exact optimal amounts.

<sup>38</sup> See Postal Regulatory Commission, "Response of the United States Postal Service to Chairman's Information Request No.1," Docket No. R2009-3 at 17, May 15, 2009.

Mail Volume Incentive Pricing Program” (Summer Sale program). The program was a series of price reductions for Standard Mail that would last for the period between July 1, 2009 and September 30, 2009. The Postal Service partly justified these price reductions by suggesting it was experiencing a period of excess capacity, primarily caused by the cyclical decline in volume induced by the recession of 2008-2009. Further, the Postal Service argued that this excess capacity caused the concurrent short-run marginal costs to be less than its traditional long-run marginal costs. It then calculated short-run marginal cost for Standard Mail products to show that the reduced prices still exceeded their short-run marginal costs.

The Postal Service’s approach to calculating short-run marginal cost was essentially to “zero out” any cost segment elasticities (variabilities) for which it determined it had experienced excess capacity and then add together the remaining volume variable costs from the other cost segments.<sup>39</sup> Setting a cost elasticity (variability) equal to zero means that none of the component’s accrued cost will be attributed to products, so the resulting short-run marginal cost, calculated in this way, is necessarily less than the associated long-run marginal cost. For the sake of the specific price incentive program, the Postal Regulatory Commission accepted the Postal Services’ approach to calculating short-run marginal costs but indicated that the use of short-run marginal costs raises larger methodological issues that it may wish to visit at a future point:<sup>40</sup>

The Commission agrees that the Postal Service’s use of short-run marginal cost to justify a change in price raises methodological issues which may impact future pricing incentive programs and the Commission’s Annual Compliance Determination. The methodological change has broad implications which warrant examination. This, however, is not the forum for that examination. The Summer Sale program is unique on several levels and can, for purposes of this proceeding, be addressed on the merits. After this program has been completed and its results analyzed, the Commission may better evaluate whether to initiate a proceeding to consider the methodological issues surrounding short-run marginal cost pricing.

The Postal Service again filed a notice of price adjustment for Standard Mail in 2010 that was similar to the one approved by the PRC in 2009. The short-run costing methodology in the 2010 filing was similar to the one used in the 2009 filing, with the exception that the 2010 filing no longer assumed there was excess capacity in city carriers and the associated cost segments returned to using their long-run elasticities.<sup>41</sup>

---

<sup>39</sup> A “cost segment” is an aggregation of cost components. The Postal Service’s product cost model has 18 active cost segments. It has dozens of cost components. The one exception to the general approach to calculating short-run marginal cost was for cost segment 6, “City Delivery Carriers - Office Activity” for which the Postal Service assumed that the short-run elasticity was 75 percent of the long run elasticity.

<sup>40</sup> Postal Regulatory Commission, “Order Approving Standard Mail Volume Incentive Pricing Program,” Docket No. R2009-3, June 4, 2009. <http://www.prc.gov/Docs/63/63340/Order219.pdf>, starting on page 8.

<sup>41</sup> Postal Regulatory Commission, “United States Postal Service Notice of Market-Dominant Price Adjustment.” Docket No. R2010-3, February 26, 2010. <http://www.prc.gov/Docs/66/66935/Not.MD.Pr.Ad.Sum.Sal.2010.pdf>.

### ***Preliminary Identification of Where Adjustment Constraints Have Arisen***

The Postal Service is different from other utilities that calculate short-run marginal cost in that it does not have a single large capital input, like a power generating plant, which is difficult to adjust to changes in demand. In addition, the Postal Service must consider both its labor inputs along with its capital inputs when analyzing inflexibilities. While the actual identification of inflexibilities is a task the Postal Service must accomplish using its detailed knowledge of its operations and legal structure, a preliminary list of possible locations provides an example of the analysis required.

A first place to investigate is the set of labor contracts that bind the Postal Service's relationship with its unionized employees. To the extent work rules and layoff restrictions constrain the Postal Service from right-sizing its workforce, either in terms of hours or employees, adjustment inflexibilities arise. Inflexibilities may vary by worker craft and analysis of short-run marginal costs should consider these differences.

Another place that bears investigation is the set of political and regulatory constraints on Postal Service business decisions, such as its universal service obligation. For example, if this obligation requires the Postal Service to have more Post Offices open than justified by a business case, an inflexibility that generates excess capacity arises.

A third set of areas to investigate inflexibilities covers the Postal Service's four networks: retail, mail processing, transportation, and delivery. Each of these networks should be carefully examined to identify any structures required to keep the network running that make it difficult to downsize in response to volume declines. For example, if a portion of the transportation network must be maintained to achieve service standards, regardless of volume, excess capacity in the transportation network could arise. Similarly, the requirement to provide delivery to almost every address in the country six days a week could generate capacity inflexibilities in an era of declining volume.

### ***Challenges in Calculating Short-Run Marginal Costs***

Because the proposed computational algorithm for calculating short-run marginal cost is embedded in the Postal Service's existing product cost system, some parts of its implementation are relatively straightforward. Other parts are more challenging because they require a deviation from ongoing methods of cost calculation. The challenging aspects include:

- Documenting the impediments to adjustment in inputs – The Postal Service's current product costing system generally assumes that inputs can be adjusted to volume changes, so it has not accumulated documentation on impediments to adjustment. Generating the documentation to support the identification of inflexibilities will require a new effort and will require significant input from experts in postal operations.

- Measuring the degree of inflexibility in affected inputs – To calculate short-run marginal costs, it is not enough just to identify where inflexibilities will occur. A measurement of the degree of those inflexibilities is also required. This means that an examination of operational methods and historical data will be needed to estimate how much of an impact the identified inflexibilities are having.
- Calculating the amount of excess capacity created by the inflexibility – During a period of secular decline, the inability to adjust inputs means that excess capacity arises and the resulting excess capacity gives rise to short-run marginal costs. An important step in calculating short-run marginal costs is the measurement of the degree to which the current mix of inputs is suboptimal through the measurement of excess capacity. This requires establishing a baseline capacity as well as measuring current throughputs.
- Linking the excess capacity to the short-run variability and estimating the resulting variability – The last step in the short-run marginal cost computational chain is estimating the impact of the excess capacity on short-run marginal cost. If only one input is used and it is in excess capacity in all applications across the entire postal network, then the short-run cost elasticity is zero because the amount of the input remains the same regardless of what happens to volume. However, if multiple inputs are used or if the degree of excess capacity varies across the network the computation of the effect on the short-run cost elasticity is more difficult and econometric estimation, functional analysis, or operational assumptions must be applied.

### ***Should Short-Run Marginal Costs Be Estimated Separately by Each Product?***

Short-run marginal costs arise because a firm cannot optimally adjust its inputs to changes in demand for its products. For the Postal Service's current situation, this means it cannot reduce the amount of certain inputs in response to ongoing demand declines. This lack of adjustment arises from either internal or external constraints on its operations and affects the cost of its providing its products.

While the impact of inflexibilities is not necessarily the same for all products, the effect takes place at the same time. If two products make use of an input that is in excess capacity, the short-run cost of providing both of those inputs will be affected as long as the input is in excess supply. Consequently, it is appropriate to estimate the short-run marginal costs for all products at the same time. The efforts required to implement the proposed computational algorithm generally apply to all products so few additional resources are required to cover all products.

Note that the Postal Service does not have to use the calculated short-run marginal costs to set prices for all products. As explained above, differences in the demand for individual products may lead to a choice of using long-run marginal costs as the basis for pricing some products while using short-run marginal costs as the basis of pricing for others. Nevertheless, the calculation of short-run marginal costs for all products would provide useful information in making important business decisions.