Service Quality, Price Caps and the USO under Entry*

Michael A. Crew and Paul R. Kleindorfer
Rutgers University and INSEAD

1. INTRODUCTION

The problem of quality of service in regulated industries has been the subject of an extensive and developing literature as surveyed by Sappington (2005). Service quality has also been an issue for postal operators (POs) under monopoly, with increasing attention paid in recent years to its measurement. It will become even more important under competitive entry, as it contributes to both the definition and cost of the USO as well as to product and marketing strategies for business products that will be the focus of competition. Indeed, with full market opening (FMO) approved to take effect in the European Union (EU) in 2011 and 2013, service quality along with maintaining the Universal Service Obligation (USO) will take on new dimensions. For example, the quality of competitive products may be affected by the quality levels determined by regulators for USO products. Similarly, the technological innovations spurred by the Internet will impact service quality and its regulation. The topic is complicated and too broad to be covered in one paper. The focus of this paper will be on integrating price-cap regulation with service quality in the postal sector under entry. While service quality under entry has an extensive literature for regulated industries, it is relatively new ground in the postal sector, because, until only recently, entry was not a major issue. With 2011 looming in Europe and postal reform taking root around the world, this is changing fast and quality is beginning to get the attention of postal economists, for example, Correia da Silva, Mautino, Dudley, and Payling (2008).

The paper proceeds as follows. Section 2 provides a very brief review of the literature on the nature of the problem of service quality in regulated industries concentrating on the postal sector. The review of other industries has been summarized thoroughly by Sappington (2005). Section 3 sets out the basic model of service quality under entry in the postal sector. It addresses both the cost of increasing quality and its benefits for postal customers. On the cost side, the zonal structure of postal operations and the interaction of quality between USO products and competitive products is examined. The regulator, in setting service standards for USO products, affects the service standards of the PO’s competitive products and consequently its ability to compete. In the model competitive products are available to business customers only while USO products are available to both business and households. The model also captures the impact of quality standards that might be imposed in terms of the density of collection points and retail outlets, which are frequently imposed as part of the USO. Section 4 extends the theoretical framework to reflect price cap regulation (PCR) and the problem facing the PO and the regulator in providing incentives for service quality of USO products. Section 5 provides a brief summary, conclusions and some ideas on the direction of future research and policy.

*The authors gratefully acknowledge the support of Royal Mail Group in the preparation of this chapter. The views expressed are solely those of the authors and do not necessarily reflect those of Royal Mail Group. Comments from Paul Dudley, Ian Leigh, David Levy and Stephen Littlechild on previous versions of the paper are gratefully acknowledged.
2. SERVICE QUALITY IN THE POSTAL SECTOR

Sappington (2005) provides an extensive and critical survey of the literature of regulating quality in regulated industries other than the postal sector. This paper will not attempt to replicate his work but will attempt to hit some of the highlights as they relate to the postal sector. Quality regulation is more complex than price regulation for a number of reasons. It is multidimensional. Quality is much more difficult to measure than price. Indeed, consumers may have the ability to perceive quality not only \textit{ex ante} but \textit{ex post}. So, not surprisingly, one of the lessons of Sappington (2005) is that the topic of quality in regulated industries normally does not yield easy answers. Often the answer, including the results in this paper, is “it depends…” This can be illustrated by Sappington’s discussion of the basic case of a single-product monopoly. He shows that an unregulated monopoly does not necessarily provide lower quality service than the welfare maximizing level. This is where the incremental willingness to pay for an increase in quality is equal to the incremental cost (of quality). This result is intuitive along the lines of other much more common marginal conditions in economics. Sappington shows that a monopolist might set quality at a level that is higher, lower or equal to the welfare maximizing level. This is somewhat counter intuitive in that the striking feature of an unregulated monopoly is its restriction of output below the competitive level. Whether the monopolist’s level of quality is greater, less or equal to the welfare optimal level depends on the marginal valuation of quality by consumers. If marginal valuation of quality decreases as the number of units purchased increases then the monopolist’s level of quality is lower than the welfare optimal. Arguably, this is the most likely case but it cannot be stated as a general rule. In addition to these demand effects, service quality can also affect costs and the interaction of these cost effects with regulation and break even operations can be complicated. This brief discussion gives a hint as to the difficulty of deriving results about quality that are simple, clear and unambiguous.

Quality in network industries has aroused considerable interest for many years. Because of its multidimensional nature, quality has been and will continue to be a source of debate. An early contribution to the debate on quality was by Telson (1975), who argued that the level of reliability provided by the electric utility industry was excessive relative to a number of practical benchmarks. Telson’s paper expressed a view that was common in the 70s that cost of service or rate of return regulation (ROR) resulted in inefficiently high levels of quality. The theoretic foundations for this notion go back to Averch-Johnson (1962) (AJ). The AJ Effect argued that under ROR firms were likely to have a bias toward excessive capital. Given the link between capital and quality (e.g., in terms of the reliability of service), the idea that quality would be in excess of efficient levels seemed a likely consequence of the AJ effect.

With ROR in mind and its incentives to provide inefficiently high levels of quality, the adaption of price-cap regulation (PCR) in the 80s led to concerns that PCR might lead to quality levels that were too low in regulated monopolies. (Sappington 2005, p131) noted that a price ceiling will provide an incentive to produce lower than the efficient level of quality “because a price ceiling prevents the firm from capturing any of the incremental consumers’ surplus that the higher service quality would engender.” Thus, the introduction of PCR to motivate cost and productivity improvements also led to a concern about potential reductions in the level of quality below efficient levels in regulated monopolies.
By contrast concern in the postal sector has primarily been that of avoiding an inefficiently low level of quality, for example, Reay (1993), and Balogh, Moriarty, Smith, Doherty, and Leigh (2006). Interestingly, this traditional concern with low quality does not support the hypothesis of Sappington and Sidak (2003), which would seem to imply that public enterprises would produce higher quality than the efficient level.\(^1\) POs have taken a number of different approaches to the problem of quality that recognize the multidimensional nature of quality, its perception by customers and measurement problems.\(^2\)

The multidimensional nature of quality arises from the fact that products can have different attributes that determine quality. For example, drinking water’s attributes include the absence of certain noxious substances and the presence of some beneficial minerals. One attribute of quality in postal service is accessibility to PO retail outlets. Similarly, different attributes determine quality in letter mail. Service might be differentiated by particular attributes. In the United Kingdom Royal Mail has traditionally offered First Class and Second Class service. Just as a First Class seat on an aircraft is differentiated by the obvious attribute of being larger than an economy seat, First and Second Class letters are differentiated by speed of delivery. By offering both classes of service a PO can raise revenue and reduce costs compared to offering one service. It can obtain more revenue from customers with a higher valuation of quality and it can lower costs by reducing peak demand and improving capacity utilization. This aspect of quality and its relation to the traditional theory of peak-load pricing has been examined extensively over the last twenty years, for example Crew Kleindorfer and Smith (1990) and Calzada (2008).

Beyond these effects of service-differentiated pricing to smooth demand and peak-load effects, there are also significant complexities associated with uncertainty in demand and the problem of assuring a given level of reliability of service. For example, although First Class mail may offer delivery within a day, because of random effects, delivery within a day is not certain. Clearly the probability of delivery (a.k.a. reliability) is an important attribute of quality. If the service offers next day delivery but the probability of delivery within this window is only 50%, then this is a much lower quality service than where the probability of delivery is say 95%.\(^3\)

So the multidimensional nature of quality arises not only from deterministic attributes but also because of random effects. The random dimensions make the problem of quality more complex in a number of respects. Pricing is more complex when random effects are taken into

---

\(^1\) SS argue that a public enterprise would likely be sales revenue maximizer, which would presumably imply higher quality. See Crew and Kleindorfer (2008) for a review of the SS hypothesis and related other models of the objectives that might be pursued by a PO, especially in the case of public enterprise, the most common organizational form for the incumbent PO in the postal sector.

\(^2\) For example, Buser, Jaag, and Trinkner (forthcoming) provide a measure of quality of service based on the accessibility of PO retail outlets and other attributes. Swinand and Jones (2006) review previous studies on postal quality and estimate willingness to pay for increased quality for international mail.

\(^3\) Crew and Kleindorfer (1992) examine the role of reliability on postal pricing and on efficient choices of mail processing technology. For example, as analyzed in Crew and Kleindorfer, flexible workers are an important element of mail processing operations precisely because they provide at lower cost than full-time workers the ability to meet reliability standards in the face of random shocks in mail demand.
account and measurement is also complicated by random effects. Simple mechanisms to control quality may not be adequate. Take the notion of a minimum quality standard, where the firm is penalized for failing to achieve the minimum standard and rewarded for exceeding it. Some random effects may involve factors completely exogenous to the firm, for example, weather, or natural disasters. The structure of efficient incentives to reward or penalize a PO based on minimum quality standards is clearly more complex in the presence of such exogenous random effects. Measurement is also more complex as a result of random effects.

These multidimensional attributes of quality explain why regulating quality is not a simple process. The problem of regulating quality for a monopoly can be summarized as a multi-stage process. The first stage is for the regulator to determine the attributes of the quality that will apply to USO products. This may mean setting different standards for urban areas from those of rural areas. The regulator will also have to take into account the random effects that are present. This amounts to setting reliability levels, e.g. the probability of delivery within the time standard specified for the area concerned. (Rural areas might also have lower reliability standards than urban areas.) The regulator also has to provide incentives to achieve these prescribed standards, perhaps in terms of the price-cap formula applied to the PO. The second stage is where the regulator monitors the achievement of the reliability levels and levies/pays appropriate penalties/bonuses. The third stage is where the regulator reviews the process and periodically may modify attributes, reliability levels and incentives. At each stage the regulator and the PO must address a number of complex issues. At the first stage there are information asymmetries, the firm having better knowledge of the relationship between quality and cost than the regulator. The regulator could over-reward or under-reward quality as a result of this information asymmetry. In either case in the third stage adjustments may be made by the regulator. Depending on the expectations of the PO on the regulator’s actions in the third stage the incentives for quality may be attenuated. For example, if the PO makes greater than expected profits the regulator may consider changing the formula. If this is anticipated by the PO, optimal quality may not be supplied.

The next section develops a basic model intended to inform the above three-stage process of quality of service regulation in the postal sector. The model attempts to break new ground by concentrating on the impact of entry on quality and the regulatory problem in the postal sector. In the process, it will be evident that the structure of the postal sector leads to a number of interesting problems under entry that complicate considerably the analysis of regulatory interventions to assure the attainment of efficient quality levels.

3. MODELING POSTAL SERVICE QUALITY UNDER COMPETITION

Two types of customers are considered: 1) consumers of single-piece letter mail, henceforth referred to as households, and 2) bulk mailers, henceforth referred to as business. A number of constraints are imposed on the Universal Service Provider (USP, assumed to be the incumbent PO): 1) a uniform price for all delivery zones for end-to-end (E2E) service for single-piece mail; 2) a constant mark-up constraint imposed on bulk mail (where the mark-up may be varied to achieve breakeven, to meet regulatory restrictions, or to maximize profits); 3) a requirement for locating post office outlets and services according to population density in each zone; 4)
constraints on the service quality of the USP; and 5) a constraint on prices. The last named constraint will take the form of either a zero-profit constraint on prices of all products provided by the USP (in the Ramsey formulation of the problem) or the form of a global price cap (in the regulated version of the problem). The intent here is to capture the spirit of the USO as constraining USP services. These constraints restrict the ability of the USP to compete (especially for high-margin products in the low-cost delivery areas). However, they may, nonetheless, be valuable for customers by reducing transactions costs, providing ubiquity of coverage or in assuring other potentially valued service attributes.

Only the case of completely inelastic demand is analyzed, with demand for zone \( t \in [0, T] \) given by \( D_H(t) \) for household customers (which includes small business mailers) and \( D_B(t) \) for business customers. Denote by \( z \in \mathbb{R}^+ \) the quality of delivery service offered, which is assumed to be the same for household and business products. Quality may differ between the incumbent/USP (I) and entrants (E) with \( z_i \) as the quality of \( i \in \{I,E\} \). In keeping with the notion of FMO, there is no reserved area for the incumbent. So it assumed Bertrand competition between the USP and entrants, who are sufficiently numerous and homogeneous that they can be treated as a competitive fringe. The consumer surplus generated per unit of demand at price \( P \) is \( V_j - P + \rho_j(z) \), \( j \in \{H,B\} \), where \( V_j \) is the reservation price per unit and where \( P - \rho_j(z) \) may be thought of as the quality-adjusted full price of the product (see Gal-Or (1983) for a similar treatment). The quality adjustment \( \rho_j(z) \) is assumed increasing and strictly concave with \( \rho_j(0) = 0 \) and \( z = 0 \) is “base-line quality”. Entrants are assumed to set quality in each zone independently, representing quality in zone \( t \) offered by E as \( z_E(t) \). By contrast, I sets a uniform quality level \( z_I(t) = z_I \) for all \( t \in [0,T] \).

The USP provides all services to the household customers, which are priced at a uniform price \( P_H \). Marginal delivery costs for each zone \( t \in [0, T] \) are assumed to constant and are denoted, respectively, by \( C_{DI}(t, z) \) for the incumbent and \( C_{DE}(t, z) \) for entrants, both of which are assumed continuous and strictly increasing functions of their arguments, and convex in quality level \( z \). Upstream collection and pre-sortation costs are denoted \( C_{UI} \) for the incumbent PO and \( C_{UE} \) for entrants. From the assumption that \( C_{UE} < C_{UI} \), entrants always process their own mail upstream.

Delivery zones \( t \in [0, T] \) are indexed in increasing order of \( C_{DE}(t, z) \), whose ordering is assumed for be unaffected by changes in \( z \). It is assumed that \( C_{EB}(0, 0) < C_{IB}(0, 0) \), so that entrants enjoy a delivery cost advantage in some of the low-cost areas for baseline quality \( z = 0 \). The USP is assumed to price business access services in zone \( t \in [0, T] \) at a constant markup over cost (but capped by the quality-adjusted single-piece uniform price \( P_H \) for all delivery zones).

---

4 The QoS offered by I to business and household customers could be different and could also depend on the delivery zone \( t \in T \), but a simpler, scalar representation is assumed. This scalar representation could capture prespecified relative quality levels across zones which could be all adjusted in sync as \( z \) varies.

5 This assumption is made primarily for expositional purposes. Weaker assumptions suffice for the results below. All the required assumptions are satisfied if, for example, \( C_{DE}(t, z) \) has the form \( C_{DE}(t, z) = a(z)C(t) + b(z) \), where \( a(z) \) and \( b(z) \) are non-negative, increasing and convex functions of \( z \), and \( C(t) \) is a non-negative, increasing function of \( t \).
Business customers in zone $t \in [0,T]$ will choose the supplier (E or I) with the lower quality adjusted price, denoted $P_E(t)$ or $P_I(t)$ for zone $t$, where:

$$P_E(t) = \min[C_{UE} + C_{DE}(t,z_E) - \rho_B(z_E), C_{UE} + (1+M)C_{DI}(t,z_I) - \rho_B(z_I)] \quad (1)$$

$$P_I(t) = C_{UI} + (1+M)C_{DI}(t,z_I) - \rho_B(z_I) \quad (2)$$

$P_E(t)$ equals the first term in brackets in $(1)$ if E bypasses in zone $t$ and equals the second term if the E uses access in zone $t$. See the proof of Lemma 1 in the Appendix.

Lemma 1: Competition among entrants will assure that, for every zone $t$ for which entrants provide delivery, quality $z_E(t)$ will be set to minimize $C_{DE}(t,z) - \rho_B(z)$. Moreover, entrants will serve those customers E2E precisely for those $t \in [0,T]$ for which

$$\hat{z}_E(t) = C_{DE}(t,\hat{z}_E(t)) - \rho_B(\hat{z}_E(t)) \leq (1+M)C_{DI}(t,z_I) - \rho_B(z_I) \quad (3)$$

Given our assumptions on $C_{DE}(t,z)$ and $\rho_B(z)$, there is a unique $z \in \mathbb{R}^+$ solving the minimization problem in Lemma 1. Denote this solution by $\hat{z}_E(t)$. With an eye on $(1)$, assume for any $t \in [0,T]$, $z_I \in \mathbb{R}^+$ and $M \in \mathbb{R}^+$ that if there is an intersection between $C_{DE}(t,\hat{z}_E(t))$ and I’s access price $(1+M)C_{DI}(t,z_I)$, it is unique. This intersection defines the zone $t_A(M, z_I)$, to the left of which entrants will deliver all business mail and to the right of which the incumbent USP will be the sole supplier of delivery services.\(^6\) It is further assumed that this intersection is below the uniform price $P_H - \rho_B(z)$ (in other words that the USP serves some of the business customers at rates below the single-piece rate). $t_U(P_H, M, z_I)$ is defined as the zone at which E’s E2E price using I’s access service $C_{UE} + (1+M)C_{DI}(t,z_I)$ just equals the quality-adjusted single-piece letter price $P_H$. Clearly, whenever this E2E price exceeds $P_H$, E will use I’s E2E service rather than access.\(^7\)

Given the above discussion and Lemma 1, the zones $t \in [0,T]$ can be divided into three subsets (see Figure 1):

- **Zones** $t \in [0,t_A]$ : Entrants provide E2E service in direct competition with I
- **Zones** $t \in (t_A,t_U]$ : Entrants workshare their mail and use I’s access services for delivery
- **Zones** $t \in (t_U,T]$ : Entrants retain any mail they collect and use I’s E2E service

---

\(^6\) More complex models have been developed after Crew and Kleindorfer (2000) which show outcomes that have the incumbent and entrants sharing delivery of mail on routes, rather than the simpler model here that shows a winner-take-all outcome in which either a delivery zone is fully supplied by entrants or fully supplied by the incumbent. These richer models take a customer-specific focus or introduce imperfect product differentiation. See also Crew and Kleindorfer (2007) for a discussion.

\(^7\) It is assumed implicitly that (at least for high-cost zones) the single-piece letter and business services are perfect substitutes for business customers. It is also assumed that E does not undertake any worksharing activities for mail delivered using I’s E2E service, and does not receive any discounts for such service. Other assumptions on access pricing and worksharing discounts for the high-cost areas would not change the results here in any substantial way.
From (1)-(2), the following two identities characterize the cutoff zones \( t_A(M, z_I) \) and \( t_U(P_H, M, z_I) \), determining zones where \( I \)'s access services are used by \( E \):

\[
\hat{C}_{DE}(t_A) = C_{DE}(t_A, \hat{z}_E(t_A)) - \rho_B(\hat{z}_E(t_A)) = (1 + M)C_{DI}(t_A, z_I) - \rho_B(z_I) \tag{4}
\]

\[
P_H = C_{UE} + (1 + M)C_{DI}(t_U, z_I) \tag{5}
\]

Figure 1 illustrates the outcome of these assumptions for business customers, following Crew and Kleindorfer (2007). The horizontal axis in Figure 1 represents the delivery zone, arranged in increasing order of entrant’s unit delivery costs (evaluated at the entrants quality adjusted delivery cost, as this is the only cost that matters in determining market shares). The vertical axis is quality-adjusted price (or unit cost). Table 1 provides a summary of notation.

**Figure 1: Illustrating Business Mail Zones Served**

By Entrants and by the Incumbent under Entry \((C_{UE} < C_{UI})\)

![Diagram illustrating business mail zones served by entrants and by the incumbent under entry](image)

The USO requirement that post offices and collection points are located in reasonable proximity to the population of households\(^8\) is assumed to give rise to (say, annual) fixed costs of:

\[
F(u) = u \int_0^T D_H(t)dt \tag{6}
\]

where \( u \in Y^+ \) reflects the cost of accessibility decisions related to the density of post office coverage. The product \( u = \gamma n \), where \( \gamma > 0 \) is the average fixed cost per post office, and \( n \) is the required density of coverage, expressed as the required number of post offices as a fraction of annual letter volumes. The decision \( u \in Y^+ \) is one element defining the scope of the USO, the other elements being uniformity of price for single-piece mail and viability. The viability constraint for the USO will be understood here to mean that the single-piece price \((P_H)\), markup

---

\(^8\) While the details of this requirement are not specified, the First Postal Directive requires accessibility and affordability as part of the USO.
(M) and quality level $z_I$ (if such exist) should assure breakeven operations for the incumbent while meeting pricing uniformity and other USO constraints represented by $u$ in (6).

<table>
<thead>
<tr>
<th>Table 1: Notation for End-to-End (E2E) Service and Access Service</th>
</tr>
</thead>
<tbody>
<tr>
<td>Using E2E Service to Zone $t$</td>
</tr>
<tr>
<td>Using E Upstream and I for Access Service to Zone $t$</td>
</tr>
<tr>
<td>Household Demand Served by the I (the USP)</td>
</tr>
<tr>
<td>$D_H(t)$ for all $t \in T$</td>
</tr>
<tr>
<td>NA</td>
</tr>
<tr>
<td>Business Demand Served by E</td>
</tr>
<tr>
<td>$D_B(t)$ for $0 &lt; t \leq t_A$</td>
</tr>
<tr>
<td>$D_B(t)$ for $t_A &lt; t \leq t_U$</td>
</tr>
<tr>
<td>Business Demand Served by I</td>
</tr>
<tr>
<td>$D_B(t)$ for $t_U &lt; t \leq T$</td>
</tr>
<tr>
<td>$D_B(t)$ for $t_A &lt; t \leq t_U$</td>
</tr>
<tr>
<td>Quality-adjusted Price of E’s Service</td>
</tr>
<tr>
<td>$C_{UE} + C_{DE}(t, \hat{z}_E(t))$</td>
</tr>
<tr>
<td>$-\rho_B(\hat{z}_E(t))$</td>
</tr>
<tr>
<td>$C_{UE} + (1 + M)C_{DI}(t, z_I)$</td>
</tr>
<tr>
<td>$-\rho_B(z_I)$</td>
</tr>
<tr>
<td>Quality-adjusted Price of I’s Service</td>
</tr>
<tr>
<td>$P_H - \rho(z_I)$</td>
</tr>
<tr>
<td>$\rho(z_I)$, $j \in {H, B}$</td>
</tr>
<tr>
<td>NA</td>
</tr>
<tr>
<td>Cost if Customer Consumes E’s Service</td>
</tr>
<tr>
<td>$C_{UE} + C_{DE}(t, \hat{z}_E(t))$</td>
</tr>
<tr>
<td>$C_{UE} + C_{DI}(t, z_I)$</td>
</tr>
<tr>
<td>Cost if Customer Consumes I’s Service</td>
</tr>
<tr>
<td>$C_{UI} + C_{DI}(t, z_I)$</td>
</tr>
<tr>
<td>NA</td>
</tr>
</tbody>
</table>

To capture the benefits of the USO, it is assumed that household transactions costs (per letter) in accessing post offices are proportional to the average distance to a post office,\(^9\) which are assumed to be inversely related to the total number of such post offices $N$. Let $\kappa/N$ be the transactions cost per single-piece letter (with annual volumes represented by $D_H(t)$ for zone $t \in [0, T]$). Using the notation defined in (6), total annual transactions costs/losses related to postal network access, denoted $L(u)$, are represented by

$$L(u) = \frac{\kappa}{N} \int_0^T D_H(t) \, dt = \frac{\kappa}{n} \int_0^T D_H(t) \, dt = \frac{\gamma \kappa}{\gamma n} = \frac{K}{u} \tag{7}$$

where $u$ defines the scope of the USO, and where $K = \gamma \kappa$ is a proportionality factor measuring the value of time and inconvenience of accessing a post office.

From (1)-(7), USP profits $\Pi(P_H, M, z_I, u)$ can be expressed as:

\(^9\)In Crew and Kleindorfer (1998), transactions cost savings for customers arising from uniform pricing by the PO are also considered. These are neglected here.
\[ \Pi(P_H, M, z, u) = \int_{t_U(P_H, M, z)}^{T} [MC(t, z) + C\!(t, z)] D(t) dt \]
\[ + \int_{t_U(P_H, M, z)}^{T} [P_H - (C_{UI} + C\!(t, z))] D(t) dt \]
\[ + \int_{0}^{T} [P_H - (C_{UI} + C\!(t, z))] D_H(t) dt - F(u) \]

The first two terms in (8) represent profit from the business sector, the third term profit from the household sector, and the final term the fixed costs of retail outlets and collection points. Note the implicit assumption here (in the third term) that mail deposited with the USP on an E2E basis for delivery in zones \( t > t_U \) still incurs all the upstream costs \( C_{UI} \) for the PO. Issues of the GYS (Crew and Kleindorfer, 2000, 2005) are not excluded here, so that the required \( u \in \mathbb{R}^+ \) is not so onerous as to preclude a solution \((P_H(u), M(u), z_I(u))\) to (8) that allows the USP to break even.

From (1)-(8), the welfare associated with a given USO policy as the weighted sum of consumer and producer surpluses is:

\[ W(P_H, M, u, z) = \alpha S_H(P_H, u, z) + S_B(P_H, M, z) + \int_{0}^{T} [V_H + \rho_H(z_I) - P_H] D_H(t) dt - L(u) \]

where \( S_H \) and \( S_B \) are the respective consumer surplus measures for sectors H and B, \( \Pi \) is the profit function for the USP in (7), and \( \alpha \geq 1 \) is the weight associated with household surplus relative to that of business customers and the USP (the welfare weights for the USP profits and business customer surplus are assumed equal). From the above definitions, household surplus is given by:

\[ S_H(P_H, u, z) = \int_{0}^{T} [V_H + \rho_H(z_I) - P_H] D_H(t) dt - L(u) \]

Also, recalling that \( \hat{C}_{DE}(t) = C_{DE}(t, z_E(t)) - \rho_B(z_E(t)) \), business surplus is given by

\[ S_B(P_H, M, z) = \int_{0}^{T} [V_B + \hat{C}_{DE}(t)] D_B(t) dt \]
\[ + \int_{t_U(P_H, M, z)}^{T} [V_B + \rho_B(z_I) - (C_{UE} + (1 + M)C\!(t, z))] D_B(t) dt \]
\[ + \int_{0}^{T} [V_B + \rho_B(z_I) - P_H] D_B(t) dt \]

Before proceeding, some assumptions on costs and demands are stated, which amount to assuming that the Incumbent is disadvantaged in the low-cost zones, but has cost advantages in the high-cost delivery zones.
Regularity Conditions (RC): \(C_{DI}(t, z)\) and \(\dot{C}_{DE}(t)\) are continuous and strictly increasing functions in their arguments, with \(C_{DE}(0,0) > \dot{C}_{DE}(0)\) and \(C_{DI}(T,0) < \dot{C}_{DE}(T)\). For any \(M, z_1 \in \mathbb{R}^+\), if there is a solution \(t_A(M,z_1) \in [0, T]\) solving the access pricing equality \(C_{DE}(t, \hat{z}_E(t)) = (1+M)C_{DI}(t,z_1)\), it is unique. Moreover, \(D_H(t)\) and \(D_B(t)\) are strictly positive on \([0, T]\).

Consider now the solution \((P^*, M^*, u^*, z_1^*)\) to the Ramsey problem

\[
\text{Maximize } \langle W(P_H, M, u, z_1) \mid \Pi(P_H, M, u, z_1) \geq 0; P_H \geq 0; M \geq 0; u \geq 0; z_1 \geq 0 \rangle \quad (12)
\]

The first-best solution (under entry) is obtained as a by-product of the Ramsey solution by neglecting the profit constraint (i.e., by assuming that its associated Lagrange multiplier \(\lambda = 0\)). To rule out myriad special cases, it is assumed that the Ramsey solution (12) entails that both access and E2E service are provided for business customers, i.e. that the Ramsey-optimal solution solving (12) satisfies \(0 < t_A(M^*, z_1^*) < t_U(P_H, M^*, z_1^*) < 1\).

Ramsey Optimal Solution: Denote by \(Q_H\) total household mail volume \(Q_H\) and by \(Q_B^U\) the total business volumes sold under single-piece rates \(P_H\), so that

\[
Q_H = \int_0^T D_H(t)dt \quad \text{and} \quad Q_B^U(P_H, M, z_1) = \int_{t_U}^T D_B(t)dt \quad (13)
\]

where \(t_U = t_U(P_H, M, z_1)\) is characterized by (5). Then, the Ramsey-optimal solution \(\langle P_H^*, M^*, u^*, z_1^* \rangle\) solving (12) satisfies the following:

\[
\frac{(\alpha - \lambda - 1)Q_H - \lambda Q_B^U}{1 + \lambda} = \left(\frac{C_{UI} - C_{UE}}{\partial C_{DI}(t_U, z_1)}\right)D_B(t_U) = \lambda \int C_{DI}(t, z_1)D_B(t) + [(1 + \lambda)(C_{UI} - C_{UE})D_B(t_U)[\partial t_U/\partial P_H] \quad (14a)
\]

\[
M^* = \left[ \frac{t_U}{t_A} \left(\frac{\lambda C_{DI}(t_A, z_1)D_B(t_A) + [(1 + \lambda)(C_{UI} - C_{UE})D_B(t_A)[\partial t_A/\partial M]}}{(1 + \lambda)C_{DI}(t_A, z_1)D_B(t_A)[\partial t_A/\partial M]} \right)^+ \right] \quad (14b)
\]

\[
u^* = \frac{\alpha K}{(1 + \lambda)Q_H} \quad (14c)
\]
\[
\begin{align*}
&\left[\frac{1}{T} \int_0^T \frac{\partial}{\partial z_1} (\alpha p_H(z_1) - C_{DI}(t,z_1)) D_H(t) dt\right] + \left[\frac{1}{T} \int_{t_A}^T \frac{\partial}{\partial z_1} (\rho_B(z_1) - C_{DI}(t,z_1)) D_B(t) dt\right] \\
&= -(C_{UI} - C_{UE}) D_B(t_U) \frac{\partial C_{DI}(t_U,z_1)}{\partial z_1} + MC_{DI}(t_A) D(t_A) \frac{\partial t_A}{\partial z_1} \\
&\quad - \lambda \frac{\partial \Pi(p_H, M, z_1, u)}{\partial z_1}
\end{align*}
\]

(14d)

where the “+” operator in (14b) is defined for any real number \(x\) as \(x^+ = \max (x, 0)\), so that the expression in brackets is truncated at 0.

While these conditions are somewhat complicated, they can be interpreted. For example, the lhs of (14a) represents the welfare gain for households relative to lost profits from a decrease in \(P_H\). The rhs represents the gain in welfare (paid for by business mailers) from decreases in the cost of upstream operations resulting from changes in \(P_H\) that decrease the coverage of E2E zones served by I (since an increase in \(P_H\) will lead to an increase in \(t_U\), with \(\partial t_U / \partial P_H > 0\) computed from (5) as shown in (14a)). Similarly, (14d) reflects the balance at optimum between the welfare changes in net benefits from changes in quality across household and business customers and the increased costs of upstream operations resulting from a decrease in \(t_U\) (where again \(\partial t_U / \partial z_1 < 0\) is computed from (5)).

It is interesting to compare the Ramsey-optimal solutions \((P^*_H(\lambda), M^*(\lambda), u^*(\lambda), z^*_1(\lambda))\) including the profit-maximizing solution, to the Welfare-optimal solution, denoted as \((P^0_H, M^0, u^0, z^0_1)\). The profit-maximizing solution corresponds to (14) when \(\lambda \to \infty\) and the welfare-optimal solution corresponds to (14) when \(\lambda = 0\). Most of the following results are directly evident from (14), comparing the case where \(\lambda = 0\) to where \(\lambda > 0\).

\[
\forall \lambda > 0, \alpha > 1: \quad P^*_H(\lambda) > P^0_H, \quad M^*(\lambda) \geq M^0 = 0, \quad u^*(\lambda) < u^0
\]

(15)

\[
\forall \alpha > 1, \exists \lambda > 0, \text{ s.t.} \forall \lambda \leq \lambda: \quad z^*_1(\lambda) \leq z^0_1
\]

(16)

The fact that \(M^0 = 0\) follows from (14b). The fact that \(M^*(\lambda) \geq 0\) for \(\lambda > 0\) follows directly from the feasibility constraint that \(M \geq 0\). More interesting is the fact that (as noted in the Appendix) \(\partial t_A / \partial M > 0\) and \(\partial t_U / \partial M < 0\). From this, it can readily be shown that \(\partial W / \partial M < 0\) for any \(M \geq 0\). With an eye on Figure 1, the intuition for this result is that increases in \(M\) lead to decreases in \(t_U\) leading to the less efficient upstream provider (I) providing services for a greater coverage area, while simultaneously decreasing coverage of the lowest access zone \(t_A\), which should be provided by entrants since their delivery costs are lower at \(t_A\) than the incumbent. Increases in \(M\) therefore lead to welfare losses. Nonetheless, in the second-best Ramsey world, it is generally the case that a positive mark-up \((M > 0)\) remains efficient in funding the USO.
Concerning quality, it turns out to be remarkably difficult to prove anything in general about the Ramsey and profit-maximizing solutions. In general, quality may be either higher or lower under profit maximization relative than under welfare maximization (Sappington 2005). With the more specific assumptions adopted here, more specific results can be derived. However, the details will not be presented, given the complexity of the expressions, e.g., (14) and the fact that pure profit maximization and pure welfare maximization will be mitigated in practice by regulation and financial constraints. A special case that is straightforward and intuitive is where the profit constraint is not severe (so that the optimal dual variable $\lambda$ in the Lagrangean defining the Ramsey problem is small). In this case (see the Appendix for details) quality will be no greater than at the welfare optimal. The intuition behind this fact is that quality has a negative impact on profits unless the mark-up $M$ is very large (so that as the weight $1+\lambda$ on profits increases in the Lagrangean, the optimal choice could be to reduce quality). WTP for customers increases with quality, and costs increase with quality, but the incumbent PO has no benefit from increased quality on the demand side (given the assumption of completely inelastic demand), except for the access products (the mid-region of Figure 1). In this region, increases in quality drive up costs, and these are recovered by the access price. However, the (optimal) mark-up $M$ must be high on access, and the cost sensitivity for access goods to quality high as well, to induce the profit-maximizing firm to want to increase quality relative to the welfare maximizing firm. This is so because the mark-up $M$ applies only to access revenues while the additional costs from quality increases apply to all routes (and for both business and household customers). Thus, when profit is not highly weighted in the Ramsey problem (corresponding to the case where USO fixed costs are not a great burden), lower quality levels will result than for a welfare maximizer.

The major assumption underlying the above results is that postal markets are open to entry with no reserved area. The model results in some interesting tensions between maintaining the USO and opening postal markets to entry. It demonstrates the key role played by the USO in this matter, and the interdependencies between the scope of the USO and pricing decisions. In the spirit of previous contributions on the USO (e.g., Crew and Kleindorfer, 1998, 2000, 2007), the model embodies the standard tradeoff between enlarging the scope of the USO to benefit small customers and its impact on required price increases for bulk mailers to cover the increased USO costs. Large mailers may (as in the model above) face non-uniform tariffs (e.g., zonal tariffs) whose mark-ups must be increased to cover any increase in the scope of the USO. So, as far as this tradeoff is concerned, the introduction of quality of service considerations into the model changes the character of solution very little.

The impact of quality of service in this model comes in three areas: first, is its direct impact on cost of providing service; second, is its impact on the value of the service but not on demand; third, through the interacting affects of quality adjusted cost and price of entrants and the incumbent, is the affect on coverage of I’s access and E2E services. While these effects are largely intuitive, their interdependencies are quite complex, as is evident from the model, which itself is clearly a simplification of the actual tradeoffs. Under FMO, competition occurs through the price of USO products. In addition, it occurs through quality competition between I and E affects the zones over which access and E2E service will be provided by each, as well as profitability for I. These complexities suggest a decentralized approach to price and quality.

---

10 This follows from the assumption that demand is completely inelastic.
regulation, implemented through price caps. Such a regime would leave most of the discretion in setting quality levels under competition to the market to determine, through the incentives provided by a price-cap regime that incorporates incentives for achieving quality of service targets for USO products. This is the focus of the following section.

4. INTEGRATING QUALITY INCENTIVES WITH PRICE CAPS

This section examines the case where incentives for quality are embodied in a price-cap regime. The representation of price-cap regulation is based on Crew and Kleindorfer (2008), but considers only a profit-maximizing incumbent. If the regulator is concerned with motivating the incumbent to increase quality, then it would seem intuitive that making the level of price caps sensitive to the quality chosen and implemented by the incumbent would provide the necessary impetus. In Portugal, for example, there is a history of such a scheme going back to 1995. The PO, CTT and the regulator, ANACOM negotiated a weighted measure of quality, which resulted in a penalty of up to one percent if CTT failed to meet the targets agreed (Castro 2008).

Integrating quality with PCR would proceed as follows. Using the most common form of index, the Laspeyres Index, the PCR constraint would be expressed as follows. For a particular basket with n products, i = 1, 2, ..., n, a quality adjusted PCR rule would require that prices $P_{it+1}$ for year t+1 be set by the regulated company (possibly subject to further constraints on the structure of these prices, as discussed in Crew and Kleindorfer, 2008) to satisfy the following constraint:

$$\sum_{i=1}^{n} P_{it+1} Q_{it} \leq \left( \sum_{i=1}^{n} P_{it} Q_{it} \right) \left[ 1 + \Delta CPI_t - X(z) + Y_t \right]$$  \hspace{1cm} (17)

where $\Delta CPI$ is the change in the consumer price index (or similar index), and the X factor indicates how much consumers can expect to gain in real terms from PCR. As argued in Crew and Kleindorfer (1996) the X factor carries a heavy load. It is initially set by the regulator to give the firm a realistic and attainable standard to improve internal efficiency and achieve benefits from innovation. The $X(z)$ factor in (17) takes into account not only these initial considerations but also provides explicit incentives in the quality dimension by making the X factor a decreasing function of (some vector of) service quality metrics, so that higher levels of quality would lead to a lower X factor, and therefore a more relaxed pricing constraint. The Y factor represents additional adjustments that the regulator may allow based on exogenous factors, e.g. tax law or other effects outside the control of the company. This quality-sensitive form of PCR is referred to by PCR-QoS.

11 Alternative objectives, such as sales maximization, are considered in Crew and Kleindorfer (2008). A discussion of these alternatives is especially relevant for public enterprises, often the ownership structure in the postal sector. The impact of many of envisaged postal reforms has been to restructure incumbents’ governance and incentives so as to provide greater alignment of their inherent objectives with profit. The question of quality regulation for objectives other than profit maximization is left for future research. As we have noted in detail in Crew and Kleindorfer (2008), the assumption of profit maximization, with an active residual claimant, is essential to the logic of price caps and the efficiency claims for them.
In the context of the model of section 3, there are only two products offered by the incumbent: E2E service at price $P_H$ and access service at the (mark-up) price $M$. In this context, the PCR-QoS constraint (17) may be represented in the simpler form:

$$w_P P_H + w_M M \leq \varphi(z_I) \quad (18)$$

where $w_P$ and $w_M$ are the corresponding PCR weights for $P_H$ and $M$ and where we assume $\varphi_z = \partial \varphi / \partial z \geq 0$. Equation (18) is just the static form of the price-cap equation (17) for two products. The quality incentive function $\varphi(z)$ might, for example, take the linear form $\varphi(z) = s_0 + s_1 z$ or the piecewise linear form $\varphi(z) = s_o - s_1(z_o - z)^+ + s_2(z - z_o)^+$ where $z_0$ is some desired or benchmark quality level and where $s_0$, $s_1$, $s_2 \geq 0$ are parameters set by the regulator. Discussion of the specification of the dependence of the price-cap constraint on measured quality $z_I$ will continue below.

As formulated, $\varphi$ does not depend on the density of collection points and retail outlets, as captured in $u$ in (7). The focus is here only on delivery quality $z_I$ and not on $u$. It is assumed that separate constraints are set by the regulator on $u$. Left to its own devices, a profit-maximizing incumbent would set $u$ below efficient levels since it would not fully value the transactions cost savings, captured in the function $L(u)$ above, in choosing $u$). This assumed regulatory constraint on $u$ (say of the form $u \geq \hat{u}$) will then give rise to some level of fixed costs $F = F(\hat{u})$ which must be covered by $I$ through its pricing, subject to the constraint (18).

Given the assumption on $\varphi_z$, increasing $z_I$ relaxes the price-cap constraint and, intuitively, this should lead the incumbent to increase quality levels $z_I$ in response to making the PCR constraint sensitive to changes in quality. The Appendix shows that this intuitive result is, indeed, correct. To prove the result, it is assumed that the Incumbent is first subject to a PCR regime that is not sensitive to quality levels (i.e., the $X$ factor in (17) or the $\varphi$ in (18) do not depend on $z$). Then a PCR-QoS regime is introduced, along the lines of (18), which is sensitive to quality levels achieved. It is further assumed that the pre-PCR-QoS regime entails quality and prices that maximize profits subject to the standard PCR constraint. Under the assumption that these pre-PCR-QoS prices were optimal, the optimal service quality under PCR-QoS must increase, at least weakly, once a quality-sensitive PCR regime is introduced. Thus, making the PCR constraint quality sensitive as in (18) will generally increase quality relative to a PCR regime in which the constraint is not quality sensitive. What is required for this result is that the PCR-QoS regime instituted be such that the existing quality level leads to no more stringent a constraint on prices than the existing PCR constraint. In the context of the standard PCR constraint (17) what this means is that the quality-sensitive $X$-factor $X(z)$ be such that $X(z_0) = X$, the pre-PCR-QoS level of the $X$ factor, where $z_0$ is the pre-PCR-QoS level of quality. This specifies the desired or target level of the consumer sharing parameter $X = X(z)$, given a historic or benchmark level of quality $z_0$. Further constraints on $X(z)$ (e.g., its derivative at $z_0$) will be discussed below.

This result points to several important issues in achieving efficient outcomes through quality-dependent price-caps.
Observable Quality Metrics and Regulation

Deterministic quality attributes, such as frequency of deliveries, can be observed ex ante and regulated as such, either through direct constraints or through the PCR-QoS methods captured in (17)-(18). Stochastic quality attributes, such as percentage of mail meeting specified delivery standards, cannot be completely determined ex ante at the time when prices in (17) are set. A feasible approach in this case is to use the previous period’s achieved quality results to condition the X(z) factor in (17). The measurement and monitoring systems needed to assess quality metrics themselves are already in place (and required under many of the postal reform laws recently enacted—see Crew, Kleindorfer and Campbell (2008) for details).

Determining the Structure of X(z)

The optimal structure of X(z) has not been examined here, but intuitively it seems clear that X(z) should reflect the tradeoff between changes in surplus for household and business customers and the net profit impact for the incumbent of changes in quality induced by the PCR-QoS regime. Thus, in (17), for any given quality metric z on which X(z) depends, the expression

$$N_t(z_t) = -\sum_{i=1}^{n} P_{it} Q_{it} \left( \frac{\partial X(z_t)}{\partial z} \right)$$

will be the effective incentive for the incumbent to increase quality at the margin at time t, since $N_t(z_t)$ reflects the change in the PCR constraint on prices at time (t+1) associated with increases in the indicated metric.\(^\text{12}\) The left-hand side of (17) may be considered an approximation of revenues in period (t+1) (it is only an approximation since the quantities on both sides of (17) are period t quantities). Thus, $N_t(z_t)$ represents the changes in allowed revenues in (t+1) that would be associated with a change in the quality metric $z_t$. Since household and business customers will be the source of this revenue, it seems intuitive that $N_t(z_t)$ should be bounded above\(^\text{13}\) (through the appropriate choice of $\partial X/\partial z$ by the regulator) by the net change in the surplus of household and business customers associated with changes in the metric $z_t$ relative to the current level of this metric is $z_0$. The trade-off between these changes in revenue (or surplus) and the cost of achieving these benefits through changes in quality will be accounted for in the Incumbent’s choice of quality. What needs to be accounted for in the structure of X(z) is the net benefits of such quality changes for postal customers. This is fortunate since the costs of quality changes are likely to be very difficult for regulators to determine. However, even the reduced information requirements noted are not trivial since little seems to be known at this point about the benefits of quality in terms of what customers are prepared to pay for increased quality. This will need to change if the PCR-QoS approach is adopted. Of course, price-caps

\(^\text{12}\) The reader will note that $N_t(z_t)$ in (19) corresponds to $\Phi_{z}(Z_t)$ in (18).

\(^\text{13}\) The regulator may decide to share these benefits with postal customers by setting a lower X(z) factor than what would be implied by equating $N_t(z_t)$ to the net surplus changes from quality changes. However, note that doing so will also erode the incentives for the PO to discover and implement quality improvements. As is the case for normal PCR, judgment is required in balancing all of the competing objectives of the regulator in determining the X factor.
themselves help to discover cost-reducing and revenue-enhancing quality measures. However, if PCR-QoS is supposed to capture benefits from quality that would be neglected by I in the absence of price-cap sensitivity, then the magnitude of these benefits, for postal customers, needs to be understood and measured by the regulator in order to properly design the PCR-QoS mechanism.\(^{14}\)

Note that this argument on the “optimal” structure of \(X(z)\) is only heuristic since, as the model in Section 3 shows, Entrants can be expected to react to quality adjustments by the PO and these reactions are not accounted for in the above argument, which focuses entirely on the incumbent PO. This intuitive logic is, however, likely to be a reasonable guide to setting the initial structure of \(X(t)\) as long as the PO has a dominant market position. After workable competition is established, it would be expected that the regulator would play an increasingly backseat role to competition in regulating quality and prices.

5. CONCLUSIONS

This paper has attempted to provide some insights on incorporating service quality regulation into the discussion of the USO and price regulation under entry. The results of this analysis are strikingly complex, in that quality has significant impacts on the decisions of Entrants concerning their own quality levels, as well their decisions on access and bypass. The results here, however, shed light on the interactions of these choices and on the structure of quality-sensitive price-cap regulation. Nonetheless, a number of questions remain and including a few of these here.

First, the general direction of our analysis has been to show that a profit-oriented firm is likely to under-invest in service quality, whether in terms of USO retail outlets or in delivery quality. This naturally gives rise to an interest in regulatory instruments to counteract this. For quality metrics such as retail outlet and collection point density, these may be best dealt with through explicit constraints, supported by empirical estimates of the efficient levels of such constraints. For stochastic measures of service quality, such as reliability of delivery standards for access or E2E services, the complexity of the tradeoffs involved suggests using a more decentralized approach to provide incentives to the PO to improve quality rather than setting constraints on quality. This gave rise to our discussion of the incorporation of such service quality incentives under a PCR regime and our results pertaining to PCR-QoS.

A related question is the level of information needed by the regulator a) on the cost of PO’s operations and b) on the benefits derived from quality by consumers in order to do a reasonable job of regulating quality. For those elements of service quality determined directly by regulatory constraints (e.g., density of and the scope of services offered at retail outlets), a great deal of information would have to be available to determine the optimal level of these constraints, and obtaining this information and setting these constraints appropriately will be increasingly difficult under competition. For those elements of service quality intended to be determined by PO decisions, as motivated by a PCR-QoS regime, less information is required on the structure

\(^{14}\) Similar comments on the need for empirical assessments of net benefits apply to setting the level of the USO accessibility constraint captured in “\(u\)”. Some initial progress in measuring these benefits is in Buser et al. (2008).
of internal PO costs, but information on the benefits for customers from increased quality will be fundamental in determining the optimal structure of the PCR-QoS constraint. Most of the ensuing work in choosing the appropriate quality levels should then be accomplished by leaving to the PO the complex decisions of investments in quality and the resulting tradeoffs between coverage, cost and profitability, subject to the PCR-QoS constraint.

The model developed does not reflect the impact of quality on demand (except through the level of the reservation price and resulting surplus measures). Obviously, if there are strong demand effects from quality, and if these dominate the cost effects of increasing quality, then regulators will have to worry considerably less about assuring sufficiently high service quality levels. Indeed, under conditions of market dominance, incumbents could over-invest in quality (as was the case in the earlier days of the electricity industry). This does not seem to be the case at present in the postal sector, but as noted in the literature survey very little information is available on either the demand or the cost effects of service quality. Empirical research on this matter therefore remains an important area for future research, both for academics as well as for regulators concerned with implementing QoS constraints efficiently. One consequence of this gap in knowledge is the importance of ex ante negotiations among regulators and the PO in setting or re-setting PCR-QoS parameters. Such negotiations would have a reinforcing character to the QoS standards agreed and the realism and feasibility of incentives for meeting and improving upon these. Until better information is available, or until workable competition becomes the rule maker for quality in the market place, a negotiation framework among relevant stakeholders has much to recommend it.\textsuperscript{15}

This paper has not addressed the practical elements of PCR-QoS implementation, such as quality dimensions to be measured, integration of PCR-QoS with access regulation, appropriate setting of targets and the precise determination of the structure of the X factor as a function of QoS. Thus, while this paper has begun the analysis of the normative underpinning of PCR-QoS, much remains to be done to integrate this with developing practice. Some elements of this integration are evident in the discussions in da Silva et al. (2008) and Castro and Franco (2008).

Finally, the Internet and advances in information technology are not only having an impact on the demand and nature of postal products but also will have an impact on how quality is measured. With initiatives like intelligent barcodes vastly greater amounts of data on quality will be available in the course of mail processing and mail delivery. These data may be available to customers and regulators and not just POs. This much richer dataset will allow not only advances in pricing, going beyond existing notions but also nearly real-time assessments of quality. While these will have major implications for the approach to quality in postal service they are not discussed here. This paper is intended, in part, to start the process of developing some basic analytical foundations that will be needed to design the incentives for achieving an efficient level of quality in postal service. So, while the issue of quality remains a complex

\textsuperscript{15} In their discussant comments on this paper, Stephen Littlechild and Ian Leigh underlined the importance of gaining better knowledge on the net benefits of quality, absent workable competition in the relevant parts of the postal value chain. In the absence of such knowledge, negotiation, consumer research and regulatory discovery processes are all obvious vehicles to promote a better understanding of appropriate targets and achievable improvements. As noted by David Levy, working out the details of how these various approaches to setting, monitoring and rewarding delivered quality of service remains a challenge not just for economic theory, but very much so for regulatory practice in the postal sector.
problem, the ability to obtain much better data that is more widely available, means that prospects for measurement are much better. The enhanced availability of data will clearly improve the *ex post* measurement process. To the extent that it reduces information asymmetries *ex ante*, it opens up the prospect of designing incentive schemes that will be more transparent and offer greater potential for efficient quality setting and monitoring. This is expected to be a significant focus of future research.

**APPENDIX**

**Proof of Lemma 1**

The object is to show that E’s quality $z_E(t)$ for each zone $t$ for which E provides delivery service satisfies:

$$z_E(t) \in \arg\min \left\{ \left. C_{DE}(t, z) - \rho_B(z) \right| z \in \mathbb{R}^+ \right\}$$

The proof is by contradiction. Suppose that E provides delivery service in some zone $t \in [0, T]$ at quality level $z_E = z_E(t)$. This implies that

$$P_E(t) = C_{UE} + C_{DE}(t, z_E) - \rho_B(z_E) < P_I(t) \quad (A1)$$

Suppose, however, that $z_E$ does not minimize $C_{DE}(t, z) - \rho_B(z)$, i.e. there is some $\hat{z}_E$ such that

$$C_{DE}(t, \hat{z}_E) - \rho_B(\hat{z}_E) < C_{DE}(t, z_E) - \rho_B(z_E) \quad (A2)$$

Then there exists an $\varepsilon > 0$ such that some entrant could set the quality-adjusted price in zone $t$ as

$$\hat{P}_E(t) = C_{UE} + C_{DE}(t, \hat{z}_E) - \rho_B(\hat{z}_E) + \varepsilon \quad (A3)$$

such that $\hat{P}_E(t) < P_E(t) < P_I(t)$. Thus, the quality-adjusted price $\hat{P}_E(t)$ would be lower than I’s price as well as entrants with price $P_E(t)$, and would attract all demand in zone $t$. Clearly the original quality-adjusted price $P_E(t)$ would not be sustainable under competition. A similar argument establishes the remaining assertions of Lemma 1. QED

**First-order Conditions for Ramsey Optimization**

The first-order conditions for the Welfare-Optimal and Ramsey results are provided and some of the consequences of these FOCs are reported in the text. For notational convenience, arguments are suppressed on $t_A = t_A(M, z_I)$ and $t_U = t_U(P_H, M, z_I)$. First, there are the comparative statics of $t_A$ and $t_U$ w.r.t. the decision variables, from (4)-(5):
\[ \frac{\partial t_A}{\partial P_H} = 0; \quad \frac{\partial t_U}{\partial P_H} = \frac{1}{(1 + M) \frac{\partial C_{DI}}{\partial t} (t_{A}, z_1)} > 0 \]  
\[ (A4) \]

\[ \frac{\partial t_A}{\partial M} = \frac{C_{DE}(t_A, z_1)}{(1 + M) \frac{\partial C_{DI}}{\partial t} (t_{A}, z_1)} > 0; \]
\[ (A5) \]

\[ \frac{\partial t_U}{\partial M} = -\left( \frac{C_{DE}(t_{U}, z_1)}{(1 + M) \frac{\partial C_{DI}}{\partial t} (t_{U}, z_1)} \right) < 0 \]

\[ (A6) \]

All of these partial derivatives have a unique sign except for \( \frac{\partial t_A}{\partial z_1} \), which depends in a non-simple fashion on all of the decision variables. Concerning (A5), the denominator in the expression defining \( \frac{\partial t_A}{\partial M} \) is positive since \( t_A \) is the unique crossing point of \( \hat{C}_{DE} \) and \( C_{DI} \) (so that, as in Figure 1, the slope of \( \hat{C}_{DE} \) must be greater than that of \( C_{DI} \) at \( t_A \)). A similar argument applies to \( \frac{\partial t_A}{\partial z_1} \) in (A6); however, in this case, the sign of the numerator is not unique (in general, going from negative to positive as \( z \) increases).

Using the notation (13), the following derivatives can be verified:

\[ \frac{\partial S_H}{\partial P_H} = -Q_H; \quad \frac{\partial S_B}{\partial P_H} = -Q_B^U (P_H, M, z_1) \]  
\[ (A7) \]

\[ \frac{\partial \Pi}{\partial P_H} = (C_{UI} - C_{UE})D_B(t_U) \frac{\partial t_U}{\partial P_H} + Q_H + Q_B^U (P_H, M, z_1) \]  
\[ (A8) \]

\[ \frac{\partial S_H}{\partial M} = 0; \quad \frac{\partial S_B}{\partial M} = -\int_{t_A}^{t_U} C_{DI}(t, z_1) D_B(t) dt \]  
\[ (A9) \]
\[
\frac{\partial \Pi}{\partial M} = \int_{t_A}^{t_U} C(t, z_1, t) D_B(t) \, dt - MC(t_A, z_1)D_B(t_A) \frac{\partial t_A}{\partial M} + (C_{UI} - C_{UE})D_B(t_U) \frac{\partial t_U}{\partial M} \tag{A10}
\]

\[
\frac{\partial S_H}{\partial z_1} = \int_0^T \frac{\partial p_H(z_1)}{\partial z_1} D_H(t) \, dt \tag{A11}
\]

\[
\frac{\partial S_B}{\partial z_1} = \int_{t_A}^{t_U} \left[ \frac{\partial p_B(z_1)}{\partial z_1} - (1 + M) \frac{\partial C(t, z_1)}{\partial z_1} \right] D_B(t) \, dt + \int_{t_U}^T \frac{\partial p_B(z_1)}{\partial z_1} D_B(t) \, dt \tag{A12}
\]

\[
\frac{\partial \Pi}{\partial z_1} = \int_{t_A}^{t_U} M \frac{\partial C(t, z_1)}{\partial z_1} D_B(t) \, dt - MC(t_A, z_1)D_B(t_A) \frac{\partial t_A}{\partial z_1} \nonumber \\
- \int_{t_U}^T \frac{\partial C(t, z_1)}{\partial z_1} D_B(t) \, dt - \int_{t_A}^{t_U} \frac{\partial C(t, z_1)}{\partial z_1} D_H(t) \, dt \nonumber \\
+ (C_{UI} - C_{UE})D_B(t_U) \frac{\partial t_U}{\partial z_1} \tag{A13}
\]

\[
\frac{\partial S_H}{\partial u} = -L'(u) = \frac{K}{u^2} \; ; \quad \frac{\partial S_B}{\partial u} = 0 \; ; \quad \frac{\partial \Pi}{\partial u} = -F'(u) = -Q_H \tag{A14}
\]

**Proof of Results (14):**

Given the above derivatives, it is possible to compute the results in (14a-d). For example, forming the Lagrangean
\[
L(P_H, M, u, z_1, \lambda) = W(P_H, M, u, z_1) + \lambda \Pi((P_H, M, u, z_1), \quad \text{condition (14a) is an immediate consequence of the FOC:}
\]

\[
\frac{\partial L}{\partial P_H} = \alpha \frac{\partial S_H}{\partial P_H} + \frac{\partial S_B}{\partial P_H} + (1 + \lambda) \frac{\partial \Pi}{\partial P_H} = 0 \tag{A15}
\]

The “+” sign on (14b) follows since it is easily verified that \( \partial W/\partial M < 0 \) whenever the quantity in square brackets in (14b) is negative (noting the signs of \( \partial t_A / \partial M > 0 \) and \( \partial t_U / \partial M < 0 \) from (A5) above), so that (14b) follows. The FOC for \( u \) solves directly to yield (14c) by setting \( F'(u) + L'(u) = 0 \). Similarly, (14d) follows from (A13).
Results Comparing Ramsey-Optimal and Welfare-Optimal Solutions:

Concerning the comparison of welfare-optimal results and Ramsey results in (15)-(16) (including those for the profit maximizing case, which corresponds to the Ramsey results as $\lambda \to \infty$), these follow directly from (14). Concerning quality, the fact that $z_1^*(\lambda) \leq z_1^0$ for $\lambda$ sufficiently small follows from (14b) and (A13). Indeed, from (A13), $\partial \Pi / \partial z_1 < 0$ when $M$ is small (from (A6), note that $\partial \Pi / \partial z_1 < 0$). Moreover, the third and fourth terms in (A13) are strictly negative so that $\partial \Pi / \partial z_1 < 0$ for a range of sufficiently small values of $M > 0$. Note that $SH$ and $SB$ are concave in $z_t$, and that at the Ramsey optimal solution:

$$\alpha \frac{\partial S_H}{\partial z_1} + \frac{\partial S_B}{\partial z_1} = -(1 + \lambda) \frac{\partial \Pi}{\partial z_1}$$

(A16)

Thus, if the added weight $\lambda \in [0, \lambda]$ given to profits in the Lagrangean is sufficiently small to assure that the optimal solution $M^*(\lambda)$ is such that $\partial \Pi / \partial z_1 < 0$ for all $\lambda \in [0, \lambda)$, then $z_1^*(\lambda)$ must decrease in this $\lambda$-interval, i.e. $z_1^*(\lambda) \leq z_1^0$ at least for $\lambda \in [0, \lambda)$.

The Positive Effect of Quality Incentives in PCR

Assume that quality incentives are embedded in the PCR constraint as in (18). Since the PCR constraint set is relaxed as $z_t$ is increased, it seems intuitive that quality should be non-decreasing under the PCR-QoS regime when compared to the non-quality sensitive PCR regime. To show this, let $(\hat{P}_H, \hat{M}, \hat{z})$ be a solution to:

$$\max \{ \Pi(P_H, M, z) \mid P_H, M, z \geq 0; w_M P_H + w_M M \leq \overline{\Phi} \}$$

(A17)

where $\overline{\Phi}$ may be considered to be the non-quality sensitive PCR cap prior to the implementation of PCR-QoS. The PCR constraint is assumed to be binding and so, from the FOCs for (A16), the (assumed interior) solution to (A16) must satisfy: $\partial \Pi / \partial P_H = \varphi P_H > 0; \partial \Pi / \partial M = \varphi M > 0$, where $\varphi > 0$ is the Lagrange multiplier associated with the PCR constraint in (A17).

Now define the PCR-QoS regime (18) through the function $\varphi(z)$ with $\varphi > 0$ and where the prices and quality in the pre-PCR-QoS regime are assumed to be feasible w.r.t. this PCR-QoS regime, in the sense that $\overline{\Phi} \geq \varphi(z)$. Let $(\hat{P}_H, \hat{M}, \hat{z})$ be a solution to:

$$\max \{ \Pi(P_H, M, z) \mid P_H, M, z \geq 0; w_M P_H + w_M M \leq \varphi(z) \}$$

(A18)

It is necessary to show that $\hat{z} \geq \hat{z}$. Suppose to the contrary that $\hat{z} < \hat{z}$. Since $\varphi > 0$, the feasibility of $(\hat{P}_H, \hat{M}, \hat{z})$ implies that
Thus, \((\hat{\phi}_H, \hat{\mu}, \hat{z})\) is also feasible in the problem (A17), from which it follows that 
\[\Pi(\hat{\phi}_H, \hat{\mu}, \hat{z}) \leq \Pi(\hat{\phi}_H, \hat{\mu}, \hat{z}).\]
In fact, since \(\partial \Pi / \partial H > 0, \partial \Pi / \partial \mu > 0\) at \((\bar{\phi}_H, \bar{\mu}, \bar{z})\), it is clear
from (A16) and (A18) that \(\Pi(\hat{\phi}_H, \hat{\mu}, \hat{z}) < \Pi(\bar{\phi}_H, \bar{\mu}, \bar{z})\), and with this the desired contradiction
to the assumed optimality of \((\bar{\phi}_H, \bar{\mu}, \bar{z})\). Therefore, \(\hat{z} \geq \bar{z}\).

\[w_p \hat{\phi}_H + w_m \hat{\mu} \leq \phi(\hat{z}) < \phi(\bar{z}) \leq \bar{\phi}\]  \hspace{1cm} (A19)

References


Economic Implications of Quality of Service Regulation in a Liberalized Postal Market.” In *Progress Toward Liberalization of the Postal and Delivery Sector*, edited by M.A. Crew and

Buser, Martin, Christian Jaag, and Urs Trinkner. 2008. “Economics of Post Office Networks:
Strategic Issues and the Impact on Mail Demand”. Forthcoming in *Handbook in Postal Reform*,
edited by M.A. Crew, P.R. Kleindorfer and J.I. Campbell, Jr.. Cheltenham, UK: Edward Elgar.

Barcelona and GPRE-IREA.

Castro, João and Agostinho Franco. “Binomial Price and Quality of Service Regulation in
Portugal, and Its Impact on the Market,” 16th Conference on Postal and Delivery Economics,
Portugal May 28-31, 2008


Services,” *Economic Journal*. 100(403), 793-807.

Crew, Michael A., Paul R. Kleindorfer, and James I. Campbell, Jr. 2008. *Handbook of


