2003RP-06

Partage des coûts et tarification des infrastructures Tarification optimale des infrastructures communes

Marcel Boyer, Michel Moreaux, Michel Truchon

Rapport de projet
Project report

Montréal
Mai 2003

CIRANO
Centre interuniversitaire de recherche en analyse des organisations

CIRANO

Le CIRANO est un organisme sans but lucratif constitué en vertu de la Loi des compagnies du Québec. Le financement de son infrastructure et de ses activités de recherche provient des cotisations de ses organisations-membres, d'une subvention d'infrastructure du ministère de la Recherche, de la Science et de la Technologie, de même que des subventions et mandats obtenus par ses équipes de recherche.

CIRANO is a private non-profit organization incorporated under the Québec Companies Act. Its infrastructure and research activities are funded through fees paid by member organizations, an infrastructure grant from the Ministère de la Recherche, de la Science et de la Technologie, and grants and research mandates obtained by its research teams.

Les organisations-partenaires / The Partner Organizations

PARTENAIRE MAJEUR
. Ministère des Finances, de l'Économie et de la Recherche [MFER]

PARTENAIRES
. Alcan inc.
. Axa Canada
. Banque du Canada
. Banque Laurentienne du Canada
. Banque Nationale du Canada
. Banque Royale du Canada
. Bell Canada
. Bombardier
. Bourse de Montréal
. Développement des ressources humaines Canada [DRHC]
. Fédération des caisses Desjardins du Québec
. Gaz Métropolitain
. Hydro-Québec
. Industrie Canada
. Pratt & Whitney Canada Inc.
. Raymond Chabot Grant Thornton
. Ville de Montréal

. École Polytechnique de Montréal
. HEC Montréal
. Université Concordia
. Université de Montréal
. Université du Québec à Montréal
. Université Laval
. Université McGill

ASSOCIÉ AU :
. Institut de Finance Mathématique de Montréal (IFM^2)
. Laboratoires universitaires Bell Canada
. Réseau de calcul et de modélisation mathématique [RCM^2]
. Réseau de centres d'excellence MITACS (Les mathématiques des technologies de l'information et des systèmes complexes)

ISSN 1499-8610 (Version imprimée) / ISSN 1499-8629 (Version en ligne)

Partage des coûts et tarification des infrastructures Tarification optimale des infrastructures communes*

Marcel Boyer†, Michel Moreaux‡, Michel Truchon§

Résumé / *Abstract*

L'intérêt général commanderait qu'on tarife l'usage des infrastructures à leur coût marginal, en prenant soin d'y inclure tous les coûts d'opportunité, dont ceux liés à la congestion et à la pollution. À tout le moins, on devrait abaisser les tarifs de manière à assurer une pleine utilisation de ces infrastructures. Ce résultat est connu par les économistes comme la solution de premier rang. Malheureusement, dans un contexte d'économie d'échelle, ce mode de tarification conduit à un déficit. Pour couvrir les coûts monétaires, les gestionnaires doivent alors élever les tarifs au-dessus des coûts marginaux. Dans l'intérêt général, les recettes totales devraient alors être prélevées grâce à une gestion aussi proche que possible du premier rang, c'est-à-dire en maximisant l'usage de l'infrastructure, pour le plus grand bien-être des citoyens. On cherche alors ce que les économistes appellent la solution de second rang, i.e. la meilleure solution sous la contrainte d'équilibre budgétaire. Le problème est de déterminer les tarifs de manière à atteindre cet objectif. C'est la problématique étudiée dans ce rapport. On débute par la tarification à la Ramsey-Boiteux ou encore linéaire. Dans ce cas, il n'y a qu'un prix par unité de bien ou service, bien qu'il puisse varier d'un bien à un autre. Ensuite, on montre qu'on peut faire mieux avec des tarifs polynômes ou non linéaires, comprenant des charges fixes, des prix d'usage, etc. On termine par un bref survol des applications qui ont été faites de la tarification linéaire et non linéaire.

Mots clés : tarification, Ramsey-Boiteux, prix non linéaires, infrastructures.

The general interest would require that the use of major equipments be priced at their marginal cost, including all opportunity costs such as congestion and pollution costs. At least, prices should be lowered until full utilization of these equipments is obtained. This is a result that economists call the first best solution. Unfortunately, under increasing returns to scale, such prices lead to a deficit. In order to covers all monetary costs, the managers must then fix prices above the marginal costs. In the general interest, they should do so while trying to preserve as much as possible the first best features, i.e. by maximizing the use of these equipments for the largest well being of the citizens. This is what economists call a second best solution, i.e. the best solution given the budget balance constraint. The problem is to find the prices that will meet this objective. These are the questions addressed in this report. We start with the Ramsey-Boiteux or linear pricing. In this case, there is only one price for each good or service, although prices may vary from one category of customers to the other. Next, we show that one can do better with multi-part or non-linear prices, comprising fix charges, usage fees, etc. The report ends with a brief survey of actual applications of linear and non-linear prices.

Keywords*: Pricing, Ramsey-Boiteux, non-linear prices, infrastructures.*

* Cette version du rapport a été remise au Ministère des Finances du Québec (MFQ) dans le cadre d'un partenariat de recherche entre le MFQ et le CIRANO. Les auteurs tiennent à remercier le Ministère pour son soutien financier. Il va de soi qu'ils sont les seuls responsables des opinions et analyses contenues dans ce document, qui ne représentent pas nécessairement celles du CIRANO ou du MFQ. Les auteurs acceptent également la responsabilité de toute erreur qui aurait pu se glisser dans le texte.

† CIRANO et Département de sciences économiques, Université de Montréal.

‡ LEERNA, IDEI, IUF, Université de Toulouse I.

§ CIRANO, CIRPÉE et Département d'économique, Université Laval, **auteur principal**.

Table des matières

Table des figures

1 Introduction

Dans les autres documents de la présente série,[1] on a parfois parlé de tarification (au coût moyen, au coût marginal, à la Aumann-Shapley) mais le problème abordé en était essentiellement un de partage de coûts. On supposait que les quantités demandées par les différents agents ou entités étaient données et il s'agissait de répartir entre ces derniers le coût de les satisfaire de façon conjointe. La question abordée dans le présent rapport est plus générale. On admet que la façon même de répartir les coûts peut avoir une influence sur les demandes elles-mêmes. On suppose donc que les agents, clients ou consommateurs ont une fonction de demande pour les services, qu'on suppose connue avec certitude.[2] La question étudiée est celle de la détermination des tarifs de façon à maximiser le bien-être des agents, tout en couvrant les coûts de production.

De façon assez générale, les infrastructures ont été mises en place à grand frais, alors que leur coût d'opération est relativement faible. L'intérêt général commanderait qu'on tarife l'usage de ces infrastructures à leur coût marginal, en prenant soin d'y inclure tous les coûts d'opportunité, dont ceux liés à la congestion et à la pollution. À tout le moins, on devrait abaisser les tarifs de manière à assurer une pleine utilisation de ces infrastructures. Ce résultat est connu par les économistes comme la solution de *premier rang (first best)*. À l'autre bout du spectre, le gestionnaire de l'infrastructure pourrait exercer tout son pouvoir de marché et maximiser ses profits privés, compte tenu de la demande.

L'objectif visé se situe le plus souvent entre ces deux solutions. Il faut collecter suffisamment de recettes pour couvrir les coûts ou, du moins, une partie d'entre eux. Dans l'intérêt général, les recettes totales doivent alors être prélevées grâce à une gestion aussi proche que possible du premier rang, c'est-à-dire en maximisant l'usage de l'infrastructure, pour le plus grand bien-être des citoyens. On cherche ce que les économistes appellent une solution de *second rang (second best)*, i.e. la meilleure solution étant donnée la contrainte budgétaire

[1] La liste en est donnée en annexe.

[2] Le cas où elle ne l'est pas est abordé en partie dans Boyer, Moreaux et Truchon (ci-après BMT) (2003).

imposée au gestionnaire. Le problème est de déterminer les tarifs de manière à atteindre cet objectif.

C'est la problématique abordée dans ce rapport. On suppose que la contrainte budgétaire ne peut pas être modifiée. On débute par la tarification à la Ramsey-Boiteux, aussi dite *linéaire*. Dans ce cas, il n'y a qu'un prix par unité de bien ou service, bien qu'il puisse varier d'un bien à un autre. Ensuite, on montre qu'on peut faire mieux avec des *tarifs polynômes* ou *non linéaires*, comprenant des charges fixes, des prix d'usage, etc. On termine par un bref survol des applications qui ont été faites de la tarification linéaire et non linéaire.

2 La tarification à la Ramsey-Boiteux

La théorie économique nous enseigne que, pour assurer la maximisation du bien-être des consommateurs, les biens et services doivent être vendus à leur coût marginal social.[3] Cependant, en présence d'économies d'échelle, ce mode de tarification donne un déficit. Une solution possible consiste à combler ce déficit par une subvention, comme on le fait pour le transport en commun et la production de spectacles par exemple. Dans d'autres situations, cela est politiquement impossible et on requiert plutôt que le responsable de la production s'autofinance, au moins en partie. Pour ce faire, il doit alors majorer les prix, du moins certains d'entre eux, au dessus des coûts marginaux.

La règle de Ramsey-Boiteux indique comment opérer cette majoration, tout en générant le moins de distorsions possibles par rapport aux consommations de premier rang obtenues avec la tarification au coût marginal. Elle maximise le bien-être total des consommateurs sous la contrainte budgétaire. Elle suppose la fonction de demande connue ou, du moins, l'élasticité de cette dernière. La définition ainsi que quelques illustrations de ce concept fondamental sont données à l'annexe A.

Il sera plus aisé de comprendre la règle de Ramsey-Boiteux si on comprend bien la manière dont un monopole fixerait les prix pour maximiser son profit. On va donc commencer par

[3] Ce dernier inclut les dommages à l'environnement et les effets pervers pour les autres agents.

ce problème, d'abord avec un catégorie de clients et ensuite deux. On passera ensuite assez naturellement à la règle de Ramsey-Boiteux.

2.1 La maximisation du profit avec une catégorie de consommateurs

Considérons un monopole qui produit un bien à un coût marginal c constant.[4] Ce monopole fait cependant face à un coût fixe C, nécessaire au fonctionnement et à l'entretien de l'infrastructure. C'est le type de configuration de coûts qui justifie l'existence d'un monopole. Dans ce cas, le coût marginal est en effet toujours inférieur au coût moyen et ce dernier est toujours décroissant. La tarification au coût marginal donnerait forcément un profit négatif.

Supposons pour l'instant qu'il n'y a qu'une catégorie de consommateurs, avec une demande d'élasticité $\eta(p)$. Pour maximiser le profit, le prix p doit être choisi de manière à satisfaire la condition suivante :

$$\frac{p-c}{p}=\frac{1}{\eta\left(p\right)} \tag{1}$$

Ainsi, pour maximiser le profit, il faut choisir un prix p de manière à ce que la marge $(p-c)$ réalisée par rapport au prix soit égale à l'inverse de l'élasticité.[5] De plus, le prix ne doit pas être si faible que la demande dépasse les capacités de production.

Une justification intuitive de l'expression ci-dessus est la suivante. La réduction du prix d'une unité ($\Delta p=-1$) a deux effets. Le premier est négatif, puisqu'il correspond à une diminution directe du profit de $\Delta p\times q=-q$. Le deuxième effet est positif et correspond à l'accroissement de la quantité demandée suite à la baisse du prix. En vertu de la définition de l'élasticité, cette dernière est donnée par $\Delta q=\eta\left(p\right)\times\dfrac{q}{p}$.[6] Elle entraîne une augmentation du

[4] Cela signifie que la production de toute unité supplémentaire entraîne un coût additionnel c.

[5] Dans la littérature sur le monopole, on réfère au rapport $\dfrac{p-c}{p}$ sous le vocable d'*indice de Lerner*. Voir Lerner (1934), Dunlop (1939) et Rothschild (1942).

[6] Voir (6) à l'annexe A.

profit d'un montant $(p-c)\times\Delta q=(p-c)\times\eta\,(p)\times\dfrac{q}{p}=\dfrac{(p-c)}{p}\times\eta(p)\times q$. Pour maximiser le profit, il faut que les deux effets s'annulent :

$$\frac{(p-c)}{p}\times\eta(p)\times q=q$$

En divisant les deux parties de cette égalité par $\eta(p)\times q$, on obtient la règle (1).

Comme cas particulier, si le coût marginal de production c est nul, le prix doit être fixé de manière à égaliser l'élasticité à 1 ($\eta\,(p)=1$). De façon générale, le prix p ne doit jamais être fixé de sorte que l'élasticité soit inférieure à 1.[7]

2.2 La maximisation du profit avec deux catégories de consommateurs

Considérons maintenant le cas où il existe deux catégories de consommateurs, avec des élasticités respectives $\eta_1\,(p)$ et $\eta_2\,(p)$, pouvant être servies à des coûts marginaux respectifs c_1 et c_2. Pour les mêmes raisons et suivant les mêmes intuitions que pour la maximisation du profit, la structure optimale des prix est donnée par la formule suivante :

$$\frac{p_1-c_1}{p_1}=\frac{1}{\eta_1\,(p_1)}\qquad\frac{p_2-c_2}{p_2}=\frac{1}{\eta_2\,(p_2)}$$

La marge réalisée par rapport au prix pour chaque type de consommateurs doit être égale à l'inverse de l'élasticité de la demande de ce même type de consommateurs. La marge réalisée par rapport au prix doit être d'autant plus importante que la demande des consommateurs a une faible élasticité. L'intuition derrière ces formules est la suivante. Les consommateurs dont la demande est moins élastique sont moins sensibles aux variations de prix que ceux dont la demande est plus élastique. Ils peuvent faire face à un prix plus élevé, payer une marge plus importante, sans pour autant diminuer sensiblement leur consommation. Les consommateurs ayant une élasticité plus grande paieront un prix plus proche du coût marginal qu'ils imposent au producteur. On veut ainsi qu'ils maintiennent la quantité demandée à un niveau élevé.

[7] Avec $\eta\,(p)<1$, une augmentation de prix d'une unité ($\Delta p=1$) entraînerait une augmentation directe du profit de $\Delta p\times q=q$. La baisse indirecte, due à l'augmentation de prix, serait inférieure à la hausse directe. On aurait en effet $p\times\Delta q=\eta\,(p)\times q>-q$. En termes absolus, $|p\times\Delta q|<q$.

Comme cas particuliers, si les coûts marginaux des deux catégories de consommateurs sont nuls ($c_1 = c_2 = 0$), alors chaque prix p_i doit être fixé de sorte que l'élasticité $\eta_i(p_i)$ soit égale à 1.

2.3 L'optimum de second rang avec deux catégories de consommateurs

Supposons dorénavant que le monopole n'ait pas le droit de maximiser ses profits mais qu'on lui permette simplement de couvrir ses coûts, i.e. d'atteindre l'*équilibre budgétaire*.[8] Comment alors doivent être fixés les prix qui atteignent l'objectif de second rang ? En suivant les mêmes intuitions, la règle est donnée par la formule suivante :

$$\frac{p_1 - c_1}{p_1} = \frac{\lambda}{\eta_1(p_1)} \qquad \frac{p_2 - c_2}{p_2} = \frac{\lambda}{\eta_2(p_2)} \tag{2}$$

Dans cette formule, λ est fixé de façon à satisfaire la condition d'équilibre budgétaire. On trouve cette valeur en résolvant un système d'équations simultanées, par tâtonnement si nécessaire. Si on pose $\lambda = 1$, on retrouve la solution qui maximise le profit. Avec $\lambda = 0$, on obtient les prix de la solution de premier rang, $p_1 = c_1$ et $p_2 = c_2$, qui donnent un déficit. La solution de second-rang, correspondant au cas de l'équilibre budgétaire, est quelque part entre les deux : $0 < \lambda < 1$. Les prix qui sont définis par (2) sont dits de Ramsey-Boiteux.[9] La règle (2) elle-même est souvent appelée la *règle de l'inverse de l'élasticité.*

Si $c_1 = c_2 = 0$, les prix doivent être tels que :

$$\eta_1(p_1) = \eta_2(p_2) = \lambda \tag{3}$$

À l'optimum de second rang, les prix sont alors choisis de sorte que les élasticités des deux catégories de consommateurs soient égales et inférieures à l'unité. Ce résultat montre que la structure des prix est, dans un sens, équitable : les deux types de consommateurs réagissent

[8] S'agissant d'entreprises publiques, on pourrait exiger qu'elles récupèrent un certain pourcentage de leurs coûts. La notion d'équilibre budgétaire peut s'entendre dans ce sens également.

[9] Pour une présentation plus poussée, voir Brown et Sibley (1986). Voir également Ramsey (1927) et Boiteux (1956).

de manière similaire à une augmentation marginale des prix, puisque leurs élasticités sont égales. Cette propriété d'équité est une caractéristique importante de la structure de ces prix.

Dans le cas plus général, les marges réalisées en fixant les prix optima de second rang (les membres gauches de (2)) sont proportionnelles, et non plus égales, à l'inverse des élasticités des demande respectives des consommateurs. Les marges sont donc plus faibles que lorsque le profit de l'entreprise est maximisé. Le facteur λ, inférieur à 1, reflète la contrainte d'équilibre budgétaire imposée au gestionnaire de l'entreprise, qui ne peut pas profiter pleinement de son pouvoir de marché pour maximiser ses profits.

Comme pour la maximisation du profit, les consommateurs à forte élasticité bénéficient d'une marge plus faible, par rapport au coût marginal, que ceux à faible élasticité. L'intuition est la même : si l'entreprise décidait d'une marge trop élevée, ces consommateurs auraient intérêt à se tourner vers un bien substitut. À l'inverse, les marchés où la demande est peu élastique peuvent supporter une marge plus élevée sans que les consommateurs modifient leur choix de façon sensible pour un produit substitut.

La règle de fixation des prix exprimée en (2) permet à l'entreprise d'atteindre l'équilibre budgétaire, tout en minimisant la distorsion par rapport à l'optimum de premier rang (pour lequel les prix sont égaux aux coûts marginaux). Même si les consommateurs à faible élasticité couvrent une plus grande part des coûts de fonctionnement et de maintenance, ils bénéficient tout de même de la présence des consommateurs dont l'élasticité est plus grande, puisque ceux-ci contribuent aussi à l'équilibre budgétaire de l'entreprise.

2.4 Prise en compte des élasticités croisées

Jusqu'à maintenant, il a été question de deux catégories de consommateurs mais d'un seul bien ou service. On peut interpréter la règle (2) comme s'appliquant à la demande pour deux biens différents mais à condition que la demande d'un bien dépende uniquement de son prix. Dans de nombreuses situations, les différents services dont il faut établir les tarifs sont cependant des substituts ou des compléments. Par exemple, l'électricité aux heures creuses est, jusqu'à un certain point, un substitut pour l'électricité aux heures de pointe. La quantité demandée dans une période dépend non seulement du tarif pour cette période mais également de celui qui s'applique aux autres périodes. On peut toujours choisir de faire la lessive durant les heures creuses plutôt qu'en période de pointe, en fonction du tarif pour les différentes périodes. Il importe alors d'en tenir compte dans la dérivation des tarifs optimaux.

Supposons que l'entreprise vende deux services et que la demande pour le service i soit donnée par $q_i = d_i(p_1, p_2)$. Convenons de noter $\eta_{12}(p_1, p_2)$ l'élasticité croisée de la demande du service 1 par rapport au prix du service 2. De même, $\eta_{21}(p_1, p_2)$ est l'élasticité croisée de la demande du service 2 par rapport au prix du service 1. La règle (2) devient alors :

$$\frac{p_1 - c_1}{p_1} = \frac{\lambda}{s_1} \qquad \frac{p_2 - c_2}{p_2} = \frac{\lambda}{s_2} \tag{4}$$

où

$$s_1 = \frac{\eta_2\eta_1 - \eta_{12}\eta_{21}}{\eta_2 - \eta_{12}\dfrac{p_2q_2}{p_1q_1}} \qquad s_2 = \frac{\eta_1\eta_2 - \eta_{21}\eta_{12}}{\eta_1 - \eta_{21}\dfrac{p_1q_1}{p_2q_2}}$$

Les termes s_1 et s_2 sont souvent appelés les super élasticités des demandes pour les deux biens. Ils deviennent respectivement η_1 et η_2 quand $\eta_{21} = \eta_{12} = 0$. À noter que les termes c_1 et c_2 peuvent ne pas être constants et même dépendre de tout le vecteur (q_1, q_2). La formule pour le cas où l'entreprise produit plus de deux biens est donnée à l'annexe B.

3 La tarification non linéaire

Si l'on se contente d'une tarification linéaire, définie par un seul nombre ou monôme, la formule de Ramsey-Boiteux indique comment faire payer les différents types de consommateurs de manière à maximiser le bien-être social. Cependant, la théorie économique nous enseigne qu'il est possible de faire mieux, en offrant aux consommateurs un menu de différents tarifs polynômes, parmi lesquels chacun peut librement choisir. Un tarif polynôme est un tarif non linéaire, défini par différents prix qui s'appliquent à différentes caractéristiques de la demande. Un tarif non linéaire peut, par exemple, être composé d'une charge fixe et de différents prix par unité pour différentes utilisations de l'infrastructure. Cette section débute avec la définition d'un tarif binôme (à deux parties). Ensuite, on définit un menu de tarifs binômes et on généralise enfin cette définition aux tarifs polynômes.[10]

3.1 Menu de tarifs polynômes

Un *tarif binôme* est défini par un couple (f, p), où f est une charge fixe, par période de temps par exemple, et p une charge variable qui s'applique à l'utilisation de l'infrastructure (nombre de passages, nombres de minutes, distance parcourue, etc.). Considérons maintenant une suite de tarifs binômes $(f_1, p_1), (f_2, p_2), (f_3, p_3), \ldots$, telle que :

$$p_1 < p_2 < p_3 < \cdots$$

$$f_1 > f_2 > f_3 > \cdots$$

Cette suite constitue un *menu de tarifs binômes*, si l'on permet aux consommateurs de choisir, dans cette suite, le couple (f_i, p_i) suivant lequel ils vont être facturés. En supposant qu'ils vont choisir le meilleur tarif, c'est-à-dire celui qui minimise leurs dépenses totales, on obtient la courbe de facturation illustrée dans la figure 1. Cette courbe illustre la recette totale obtenue d'un consommateur, en fonction de son utilisation de l'infrastructure. Chaque paire (f_i, p_i) définit une droite d'ordonnée à l'origine f_i et de pente p_i. Cependant, seules les parties pleines de la droite doivent être considérées, puisque tout point des parties en pointillés est dominé, en termes de coût, par un point situé en-dessous, sur une autre droite.

Le concept de tarif binôme peut être généralisé à un nombre quelconque de composantes. Par exemple, un tarif peut prendre la forme (f_i, p_i, r_i, s_i), où f_i est un charge fixe et p_i, r_i et s_i représentent des charges variables, chacune afférente à une caractéristique différente du service. Par exemple, p_i pourrait être le prix à la minute d'une communication Internet, r_i celui du volume d'information téléchargée et s_i celui du volume téléversé. Un tel vecteur est un tarif polynôme. Une suite de tarifs polynômes est un *menu de tarifs polynômes*.

La théorie économique récente nous enseigne qu'il est toujours possible de faire mieux avec un menu de tarifs polynômes qu'avec un tarif linéaire, disons $p_1 > c$. Plus exactement, il est

[10] Pour une présentation plus extensive, voir Wilson (1993).

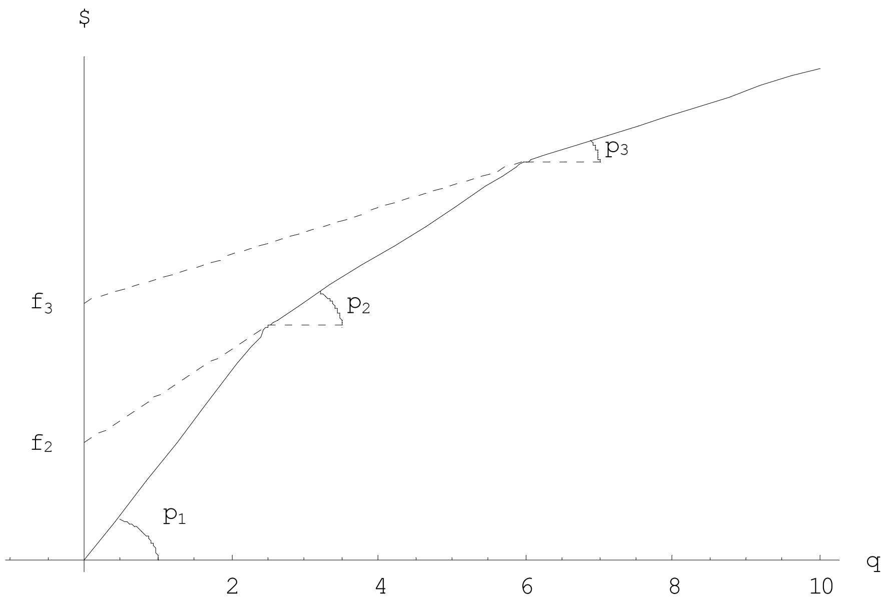

Figure 1 – Courbe des charges avec un menu de tarifs binômes

possible d'avantager certains consommateurs, sans en défavoriser d'autres, tout en augmentant les recettes nettes. On peut comprendre comment en considérant deux consommateurs, l'un à faible demande et l'autre à forte demande. Soit q_1 et q_2 les utilisations respectives de l'infrastructure par ces deux consommateurs à un prix unique p_1. Supposons $q_1 < q_2$ et proposons maintenant le menu $\{(f_1, p_1), (f_2, p_2)\}$ aux consommateurs, avec $p_2 < p_1$ et des charges fixes $f_1 = 0$ et $f_2 = (p_1 - p_2)q_2$. Les consommateurs peuvent choisir la même tarification que précédemment, i.e. p_1, puisque $f_1 = 0$, ou bien ils peuvent choisir (f_2, p_2). La charge fixe f_2 satisfaisant

$$f_2 + p_2 q_2 = p_1 q_2,$$

le consommateur 2, en choisissant (f_2, p_2), paie la même facture qu'avec un prix unique p_1 et consomme la même quantité q_2. Toutefois, étant donné que le prix variable par unité de service est plus faible ($p_2 < p_1$), ce consommateur peut envisager d'augmenter sa demande, au moins à long terme, ce qui se traduira par une augmentation des recettes de l'entreprise.

En ce qui concerne le consommateur 1, il va vraisemblablement choisir de conserver la tarification initiale p_1, puisque, avec $q_1 < q_2$,

$$f_2 > (p_1 - p_2)\, q_1,$$

et donc

$$f_2 + p_2 q_1 > p_1 q_1,$$

à moins qu'il ne trouve plus avantageux de payer la charge fixe f_2 et d'accroître sa consommation au prix p_2. En tout état de cause, il ne se plaindra pas, sa situation ne pouvant évoluer que dans le bon sens. Le fait que toutes les structures tarifaires soient disponibles à tous les consommateurs confère à cette forme de tarification un principe d'équité notable, celui d'*absence d'envie*. Chaque utilisateur peut choisir le tarif qui convient le mieux à ses besoins et ses particularités.

Si le gérant de l'infrastructure n'a pas besoin des recettes supplémentaires qu'apporterait ce menu de tarifs (s'il a déjà atteint l'équilibre budgétaire), il peut le redistribuer aux consommateurs en réduisant les charges f_i et p_i. Ainsi, on a montré comment, en partant d'un prix unique linéaire p_1, il est possible d'améliorer le bien-être de **chaque agent** en introduisant une autre tarification (f_2, p_2), où $p_2 < p_1$.

3.2 Menu optimal

Une question venant naturellement à l'esprit est de savoir s'il existe une structure optimale pour de tels tarifs. La réponse est positive. Cette structure optimale dépend de la nature des différentes demandes, plus particulièrement des élasticités $\eta_1(p)$ et $\eta_2(p)$ des deux consommateurs ou groupes de consommateurs (marchés, demandes). En fait, quand on considère des tarifs polynômes, il faut distinguer l'élasticité par rapport à la charge fixe de l'élasticité par rapport à la charge variable. Ces élasticités sont en principe différentes et la structure tarifaire optimale dépend des deux.

Dans la recherche du menu optimal de tarifs, on doit s'assurer que le consommateur 1 choisira, volontairement et naturellement, le tarif (f_1, p_1), alors que le consommateur 2 préférera

(f_2, p_2). En outre, les différents tarifs (f_i, p_i) doivent être conçus de sorte que la contrainte d'équilibre budgétaire soit vérifiée et que l'infrastructure soit utilisée de façon optimale, i.e. à son maximum.

Si l'on peut améliorer l'utilisation de l'infrastructure, tout en satisfaisant la contrainte d'équilibre budgétaire, en passant d'un tarif linéaire unique p_1 à un menu de tarifs binômes $\{(f_1, p_1), (f_2, p_2)\}$, on peut améliorer davantage l'utilisation de cette infrastructure avec un menu $\{(f_1, p_1), (f_2, p_2), (f_3, p_3)\}$ et ainsi de suite. Le nombre maximum de composantes à un menu de tarifs polynômes correspond au nombre de différents types de consommateurs ou groupes de consommateurs qui utilisent l'infrastructure.

Rappelons que, en construisant ce menu de tarifs, on doit s'assurer que chaque type de consommateur trouvera avantageux de choisir le tarif qui lui est destiné, c'est-à-dire que le consommateur 1 doit préférer (f_1, p_1) à tout autre tarif, le consommateur 2 le tarif (f_2, p_2), le consommateur 3 le tarif (f_3, p_3) et ainsi de suite. On dit alors que les tarifs permettent l'auto-sélection des consommateurs, ou encore que ces tarifs sont auto-sélectifs.

Il faut bien sûr traiter cette complication supplémentaire avec beaucoup d'attention. En particulier, les différents tarifs (f_i, p_i) doivent être déterminés de manière à satisfaire la contrainte d'équilibre budgétaire, tout en maximisant l'utilisation de l'infrastructure. L'élaboration de tarifs polynômes à n-parties est évidemment plus complexe que celle de tarifs binômes.

3.3 Un exemple

On a vu qu'une tarification polynôme peut être plus efficace qu'un tarif linéaire simple. On a aussi insisté sur le problème d'auto-sélection qu'il faut résoudre lorsqu'on établit une telle tarification. Voyons maintenant un exemple illustrant ces idées. En particulier, on va déterminer un couple de tarifs binômes qui permette d'atteindre une solution plus efficace que celle obtenue par les prix de Ramsey-Boiteux.

Supposons qu'il existe deux consommateurs, dont les fonctions de demande respectives sont :

$$\begin{aligned} q_1 &= 100 - 10\,p_1 \\ q_2 &= 10 - 0.5\,p_2 \end{aligned}$$

Le coût total qui doit être couvert est fixe et égal à 75$.

Tarifs de Ramsey-Boiteux

Les prix de Ramsey-Boiteux sont les solutions du système d'équations suivant :

$$\begin{cases} \dfrac{10p_1}{100 - 10p_1} = \dfrac{0.5p_2}{10 - 0.5p_2} \\ 100p_1 - 10p_1^2 + 10p_2 - 0.5p_2^2 = 75 \end{cases}$$

La première équation est équivalente à $\eta_1\left(p_1\right) = \eta_2\left(p_2\right)$ et la deuxième représente la contrainte d'équilibre budgétaire. La solution est :

$$\begin{aligned} p_1 &= 0.6698 \\ p_2 &= 1.3397 \end{aligned}$$

Si ces prix sont proposés aux consommateurs, leurs demandes respectives sont :

$$\begin{aligned} q_1 &= 93.302 \\ q_2 &= 9.3301 \end{aligned}$$

Tarifs binômes auto-sélectifs

Avec les tarifs de Ramsey-Boiteux, le consommateur 2 préférerait payer p_1 plutôt que p_2. Ce résultat est général : la tarification à la Ramsey-Boiteux ne satisfait pas au principe d'absence d'envie. Cela signifie qu'elle ne satisfait pas à la contrainte d'auto-sélection non plus : le consommateur 2 a intérêt à se faire passer pour le consommateur 1.

On va à présent montrer qu'il existe une structure tarifaire (un menu) $\{(f_1, p_1), (f_2, p_2)\}$ permettant d'atteindre une meilleure solution que celle obtenue avec les prix de Ramsey-Boiteux. Rappelons que f_i est une charge fixe, p_i un tarif par unité consommée et que chaque consommateur peut librement choisir entre les deux paires (f_1, p_1) et (f_2, p_2). Cette structure tarifaire est $\{(8, 0.574), (2, 1.1493)\}$. Les deux composantes de ce menu de tarifs binômes définissent les deux droites tracées sur la figure 2. La courbe formée des deux segments pleins donne les dépenses totales en fonction de la quantité demandée.

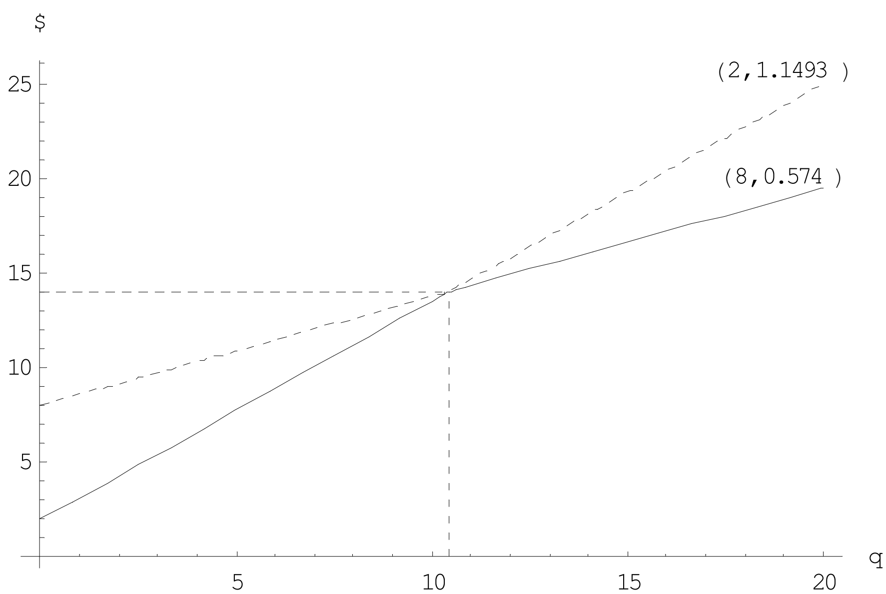

Figure 2 – Deux tarifs binômes

Pour l'instant, faisons l'hypothèse simplificatrice que les charges fixes, 8 ou 2, n'ont pas d'impact sur la demande des consommateurs. Seules les charges variables entrent en compte dans leurs fonctions de demande. Si le consommateur 1 choisit la paire $(8, 0.574)$ et le consommateur 2 la paire $(2, 1.1493)$, leurs demandes respectives sont :

$$
\begin{aligned}
q_1 &= 94.26, \\
q_2 &= 9.425,
\end{aligned}
$$

Ces consommations sont supérieures à celles calculées avec les prix de Ramsey-Boiteux. En outre, les recettes liées à cette tarification sont exactement de 75$.

Il reste à montrer que le consommateur 1 va effectivement choisir la paire $(8, 0.574)$ et le consommateur 2 la paire $(2, 1.1493)$. Sur le graphique 2, on voit que $q_1 = 94.26$ coûte moins cher au consommateur 1 avec le premier tarif qu'avec le deuxième et que $q_2 = 9.425$ coûte moins cher au consommateur 2 avec le deuxième tarif qu'avec le premier. De plus, q_1 se situe à droite du point d'équivalence des deux tarifs, alors que q_2 se situe à gauche de ce point. Cependant, on devrait comparer les surplus des consommateurs plutôt que les dépenses totales. Que sont ces surplus ?

Considérons la fonction de demande représentée dans la figure 3. Si le tarif linéaire est 5$, alors le surplus des consommateurs dans ce marché est donné par l'aire ombrée, moins toute charge fixe f. En effet, on peut interpréter la courbe de demande comme représentant le prix maximum que le consommateur est prêt à payer pour chaque unité de service. Ainsi, s'il paie en réalité 5$ pour une unité qu'il serait prêt à payer 10$, comme pour la 5e unité, le consommateur réalise un surplus de 5$ sur cette unité. En reproduisant ce même raisonnement pour toutes les unités de bien achetées, on obtient le triangle ombré, qui représente le surplus brut du consommateur. Le surplus net s'obtient en retranchant la charge fixe qu'il doit payer pour avoir accès à l'infrastructure. Avec une fonction de demande de la forme $q = b - ap$, le surplus, fonction de f et de p, est facile à calculer. Il s'écrit :

$$S(f, p) = \frac{\left(\frac{b}{a} - p\right)(b - a\,p)}{2} - f$$

Appliquons cette formule pour calculer le surplus du consommateur 1 avec chacun des deux tarifs binômes. On obtient :

$$\begin{aligned} S_1(8,\ 0.574) &= 436.24 \\ S_1(2,\ 1.1493) &= 389.67. \end{aligned}$$

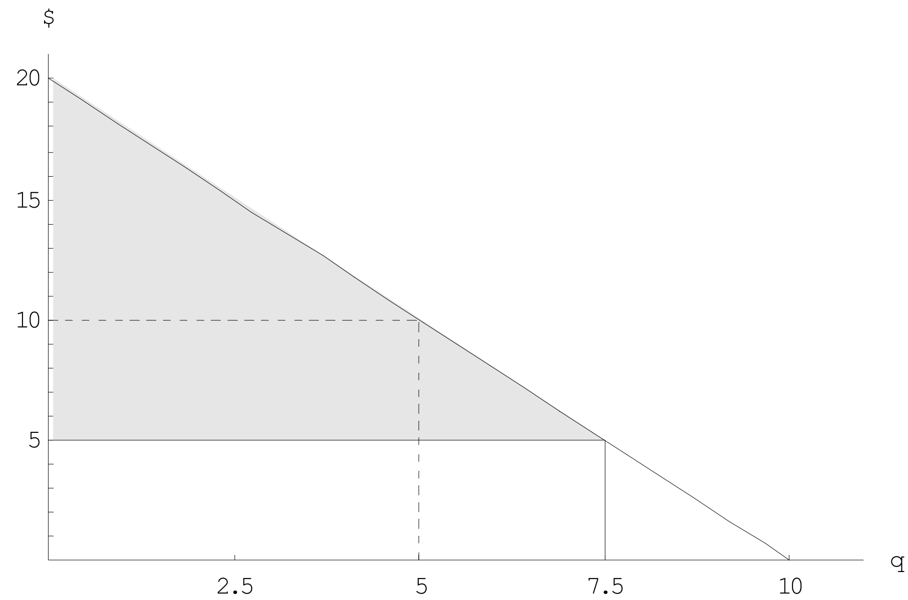

Figure 3 – Le surplus des consommateurs lorsque le prix est fixé à 5$

Il est clair que le consommateur 1 gagne à choisir la première paire de la structure tarifaire, $(8, 0.574)$. Quant au consommateur 2, on obtient :

$$\begin{aligned} S_2(8,\ 0.574) &= 86.34 \\ S_2\,(2,\ 1.1493) &= 86.83. \end{aligned}$$

Le consommateur 2 a donc intérêt à choisir le tarif $(2, 1.1493)$.

En résumé, cette tarification, où les paires (charge fixe, charge variable) ont été spécialement déterminées pour les consommateurs 1 et 2, entraîne une solution meilleure que celle de Ramsey-Boiteux et satisfait au principe de non envie et à la condition d'auto-sélection. Il n'est donc pas nécessaire d'imposer un tarif à un consommateur. Cette tarification permet aussi d'accroître le surplus de chaque consommateur. Remarquons aussi que les charges variables de chaque tarif tiennent compte des différentes élasticités de la demande. On a en effet $p_1 < p_2$ comme avec la règle de Ramsey-Boiteux. En fait, p_1 et p_2 sont les prix optimaux de Ramsey-Boiteux pour $f_1 = 8$ et $f_2 = 2$. Toutefois, ces choix de f_1 et f_2 ne sont pas nécessairement optimaux.

Prise en compte de l'effet de la charge fixe

Jusqu'à maintenant, on a supposé que les charges fixes, f_1 et f_2, n'avaient aucun impact sur la demande. En réalité, elle peut influencer la demande. Supposons donc que la partie fixe du tarif a un effet négatif sur la quantité demandée. Par exemple, supposons que la fonction de demande soit de la forme $d(f,p) = b - ap - cf$. La fonction de surplus d'un consommateur s'écrit alors :

$$S\,(f,p) = \frac{\left(\frac{b-c\,f}{a} - p\right)(b - c\,f - a\,p)}{2} - f$$

Soit les demandes des deux consommateurs :

$$\begin{aligned} q_1 &= 100 - 10\,p_1 - 0.5\,f_1 \\ q_2 &= 10 - 0,5\,p_2 - 0.5\,f_2 \end{aligned}$$

Avec de telles fonctions de demande, voyons que les tarifs non linéaires

$$\{(11.2638,\ 0.574),\ (4.7751,\ 1.1493)\}$$

satisfont la condition d'auto-sélection. En effet, les surplus du consommateur 1 selon ces deux tarifs sont :

$$\begin{aligned} S_1\,(11.2638,\ 0.574) &= 381.4832, \\ S_1\,(4.7751,\ 1.1493) &= 366.052. \end{aligned}$$

Ainsi, il est clair que le consommateur 1 va choisir le premier tarif $(11.2638,\ 0.574)$. Les surplus du consommateur 2 sont :

$$\begin{aligned} S_2\,(11.2638,\ 0.574) &= 5.3915 \\ S_2\,(4.7751,\ 1.1493) &= 44.7555 \end{aligned}$$

Le consommateur 2 préfère donc le deuxième tarif $(4.7751,\ 1.1493)$. En somme, les consommateurs 1 et 2 ont toujours intérêt à choisir les tarifs respectifs $(11.2638, 0.574)$ et $(4.7751, 1.1493)$ qui ont été conçus pour eux.

4 La tarification linéaire et non linéaire en pratique

Nous présentons maintenant une brève revue de certains articles qui étudient la tarification à la Ramsey-Boiteux et la tarification non linéaire (polynôme). Ce survol inclut des analyses empiriques.

4.1 La tarification à la Ramsey-Boiteux

Le secteur de l'électricité fait partie des secteurs où les tarifs optimaux sont utilisés de façon extensive. Dans une étude portant sur un échantillon de fournisseurs privés d'électricité, Naughton (1988) développe un modèle de régulation par les prix où les préférences du régulateur sur les catégories de consommateurs peuvent varier. Cette analyse montre que la structure tarifaire alors en vigueur favorisait davantage les petits consommateurs, résidentiels ou commerciaux, en leur facturant, pour des raisons d'équité, des prix inférieurs aux prix de second rang de Ramsey-Boiteux. En général, les consommateurs commerciaux étaient les moins favorisés (car les moins politiquement organisés). La règle de Ramsey-Boiteux aurait commandé d'augmenter légèrement les prix destinés à la consommation résidentielle, sans modifier les prix à la consommation industrielle, de réduire de façon significative les prix de l'électricité à usage commercial et d'augmenter les charges fixes.

Dans de nombreux pays, on a vu la tarification traditionnelle de l'électricité (tarif uniforme ou tarifs par période) évoluer vers de nouvelles formes qui tiennent compte des coûts marginaux variables de fournir le service et ce de façon dynamique. Une particularité importante de ces nouveaux tarifs est qu'ils sont le plus souvent proposés comme options au tarif de base, de sorte que les consommateurs peuvent choisir entre l'ancien et les nouveaux tarifs. Sur sept pays européens étudiés, six offrent des tarifs par période comme alternative au tarif unique de base.

Dans une expérience de tarification dynamique réalisée en Finlande, Räsänan et al. (1997) proposent un grille de tarifs qui aurait donné au consommateur la possibilité de choisir le tarif le mieux adapté à son type de consommation. Il y a quatre types de tarifs envisagés : un prix uniforme, un tarif avec prix par période, un menu de deux prix uniformes et un menu de deux tarifs avec prix par période. Cette étude montre que l'option de deux tarifs avec prix par période donne les bénéfices les plus élevés. En permettant aux consommateurs de choisir entre plusieurs tarifs, on peut accroître le bien-être social.

Un autre secteur dans lequel on a constaté que la règle de Ramsey-Boiteux était utile est celui des transports. La tarification à la Ramsey-Boiteux a été proposée pour facturer l'accès aux aéroports non congestionnés. S'il existe des excès de capacité, la tarification au coût marginal entraîne un déficit. Cela conduit l'aéroport à imposer des frais fixes supérieurs aux coûts marginaux, ce qui provoque d'autres inefficacités. En général, la base pour les frais d'atterrissage est le poids de l'appareil (poids maximum à l'atterrissage ou poids maximum au décollage). Une étude de Morrison (1982) évalue les prix optimaux de Ramsey-Boiteux pour les aéroports non encombrés et montre que les frais optimaux d'atterrissage doivent dépendre du coût marginal de l'aéroport, de l'élasticité de la demande et du coût du vol. Ces frais doivent croître avec le coût marginal de l'aéroport et le coût du vol et décroître avec l'élasticité de la demande. Ces trois paramètres peuvent être liés à deux particularités de chaque vol : la taille (ou le poids) de l'appareil et la distance qu'il parcourt. Cette étude conclut que les prix d'accès à un aéroport devraient être basés sur la taille (le poids) et la durée du vol.

Des prix à la Ramsey-Boiteux ont aussi été calculés pour le transport en commun dans la partie est de la région de la Baie de San Francisco. Dans cette région, les transporteurs ont un monopole naturel puisque leur coût marginal est inférieur au coût moyen. Cela implique qu'une tarification au coût marginal entraîne un déficit. Train (1977) a montré que, en définissant des prix de Ramsey-Boiteux, les utilisateurs du service par autobus subventionneraient l'usage du rail. Cependant, dans cette région de la Baie de San Francisco, les usagers de l'autobus ont des revenus plus faibles que les usagers du chemin de fer. Du point de vue de l'équité, cette subvention croisée semble tout à fait inopportune. L'étude propose alors que les prix de Ramsey-Boiteux soient ajustés de manière à prendre en compte ces considérations d'équité.

L'application de la tarification à la Ramsey-Boiteux aux services postaux a aussi été abondamment étudiée. Les Services postaux des États-Unis (USPS) ont été créés en 1970 en tant que compagnie gouvernementale semi-autonome. L'USPS est contrainte à l'équilibre budgétaire et chaque catégorie de courrier doit prendre en charge les coûts directs et indirects qui lui sont imputables, ainsi qu'une fraction des coûts institutionnels. Pour partager les coûts institutionnels, non imputables à une catégorie particulière de courrier, les services postaux se sont, pour un temps, servis de la règle de l'inverse de l'élasticité. Mais même si cette règle était utilisée en 1974, il n'existait pas d'estimation de ces élasticités, qui n'ont été disponibles qu'en 1976. La méthode utilisée alors consistait à classer les catégories de courrier suivant les élasticités relatives de la demande.

La règle simple de l'inverse de l'élasticité aurait été appropriée si les postes produisaient des services répondant à des demandes indépendantes. Cependant, il existe des relations de substitution entre les différentes catégories de courrier. L'USPS, reconnaissant l'existence des élasticités-prix croisées non nulles, a tenté d'introduire cette notion dans l'établissement de ses tarifs. En 1978, l'application de la règle de Ramsey-Boiteux a tenu compte des élasticités-prix croisées, sous forme de contrainte, et plus tard, les modèles de demandes utilisés par USPS ont explicitement pris en compte les effets prix croisés.

Une autre particularité, qui distingue encore les services postaux de beaucoup d'autres services publics, est l'existence de services substituts ou compléments dans le secteur privé. Sherman et George (1979) ont remarqué que USPS ne prenait pas correctement en compte la concurrence des produits du secteur privé, ni les effets de ses propres tarifs sur les profits du secteur privé. Ils dérivent un système d'équations permettant de caractériser les prix optimaux en prenant en compte les interdépendances avec le secteur privé. Scott (1986) compare les tarifs appliqués par USPS avec ceux qui seraient obtenus en appliquant la règle de Ramsey-Boiteux généralisée, tout en tenant compte de la concurrence. Le fait que chaque prix soit augmenté de la "valeur du service" que représente la partie du coût institutionnel a pour conséquence que les tarifs choisis par USPS sont assez proches de ceux calculés grâce aux principes généraux de la règle de Ramsey-Boiteux, même si on n'utilise pas explicitement cette dernière règle.

En Angleterre, la poste n'a pas fait d'études statistiques élaborées des demandes de ses produits, même si quelques travaux préliminaires ont été réalisés. Albon (1989) évalue la tarification à la Ramsey-Boiteux sous le régime alors en vigueur de deux niveaux de prix (première et deuxième classes) et montre que la différence de prix entre les courriers de première et deuxième classes est bien plus importante que ce qu'elle devrait être. Cette étude constate aussi que la tarification des envois nationaux ne dépend pas assez explicitement de la distance. L'analyse suggère une structure de prix à quatre niveaux, suivant la catégorie (première et deuxième classes) et la destination (urbain et rural). Une tarification à la Ramsey-Boiteux avec quatre niveaux de tarifs aurait entraîné une augmentation du volume de courrier envoyé.

Dans une autre étude du service téléphonique, Train (1994) calcule les tarifs auto-sélectifs optimaux pour Bell of Pennsylvania, et les compare aux tarifs uniques et à la tarification au coût marginal. L'analyse montre que les tarifs auto-sélectifs conduisent à un surplus social supérieur à celui obtenu avec la tarification obligatoire au coût marginal.

4.2 Tarifs polynômes

Les compagnies d'électricité utilisent fréquemment des tarifs binômes ou polynômes. Par exemple, la compagnie Commonwealth Edison proposait en 1976 un tarif polynôme de la forme suivante : un abonnement de 1.20$ et deux prix unitaires, 0.0418$ et 0.03148$, suivant que la consommation était inférieure ou supérieure à 100 kWh (Brown et Sibley, 1986). Dans cet exemple, l'abonnement est le même, quelle que soit la consommation. Dans d'autres cas, la charge fixe elle-même peut varier avec la consommation. Électricité de France (EDF) propose ce genre de tarification. Les tarifs d'EDF reposent sur un système élaboré de prix non linéaires. Un exemple est le tarif bleu qui offrent trois options aux consommateurs. Chaque option consiste en un abonnement mensuel et un prix par unité d'énergie consommée. L'abonnement est calculé suivant un programme non linéaire et est d'autant plus élevé que les besoins sont grands. Les consommateurs peuvent choisir entre les trois options suivantes : tarif de base, tarif "heures creuses", tarif "périodes critiques" (Wilson, 1993).

Par rapport au tarif de base, l'abonnement du tarif "heures creuses" est plus élevé et celui du tarif "périodes critiques" est plus faible. Avec le tarif de base, le prix unitaire de l'électricité est constant alors que les deux autres tarifs ont chacun un deuxième prix. Avec le tarif "heures creuses", le prix est réduit pendant les périodes creuses alors que le tarif "périodes critiques" impose un prix très élevé pour l'électricité consommée en période de pointe, pénalisant la consommation en période de très forte demande où la capacité est saturée. Ce tarif bleu est en fait une combinaison de trois tarifs binômes, chacun consistant en un abonnement et une charge pour l'énergie consommée. Les trois options offrent à chaque consommateur la possibilité de choisir le tarif binôme qui minimisent ses dépenses étant donné son profil de consommation.

Il existe de surcroît un tarif jaune et un tarif vert, qui sont dans l'ensemble similaires au tarif bleu, si ce n'est qu'ils font dépendre la facturation d'autres facteurs, comme la saison, le mois, la durée de la consommation, etc. Mais, en général, les petits consommateurs ont le choix entre des tarifs moins raffinés et moins complexes que les grands consommateurs (industries, etc.). Avec la mise en œuvre de tels tarifs non linéaires, EDF couvre ses coûts et incite les consommateurs à utiliser les ressources énergétiques de façon plus efficace.

Lorsque le prix payé par les consommateurs dépend de la quantité consommée, il est fréquent que ce prix soit dégressif : plus on consomme, plus le prix unitaire est faible. Dans certains cas cependant, cela peut être le contraire : par exemple, un prix progressif peut être opportun, afin d'encourager l'utilisation efficace de l'énergie. Ainsi, la compagnie China Light and Power de Hong Kong est passée d'un tarif dégressif à un tarif uniforme en 1994, puis à un tarif progressif à trois niveaux en 1996 et enfin à un tarif progressif à quatre niveaux en 1998. Les prix élevés pour les niveaux supérieurs de consommation découragent le gaspillage, alors que les prix faibles pour les consommations des niveaux inférieurs sont destinés à protéger les usagers à faibles revenus qui consomment peu d'énergie. Certaines métropoles, comme Tokyo ou San Francisco, appliquent aussi des tarifs progressifs pour les particuliers.

La facturation des services téléphoniques est une autre illustration de l'application des tarifs polynômes. Aux États Unis, depuis 1984, les fournisseurs de services interurbains paient deux charges (par minute) pour connecter leurs usagers aux réseaux existants : une charge pour la ligne afin de couvrir les coûts d'accès à la boucle locale, invariants par rapport au trafic, et une autre charge, qui dépend du trafic, couvrant les services de commutation de l'appel ainsi que les frais d'interurbain. En général, ces charges sont refilées aux consommateurs par le biais d'un prix d'appel interurbain (par minute) uniforme. Comme ces frais devenaient considérables pour certains grands consommateurs, ces derniers ont commencé à installer des systèmes de connexions directes, afin de contourner la boucle locale et la commutation des appels. En 1986, les opérateurs ont déposé une requête auprès de la Federal Communications Commission (FCC) afin d'obtenir qu'un prix non linéaire soit facturé directement aux consommateurs, dans le but d'empêcher le contournement pratiqué par les grands consommateurs. Les prix proposés auraient avantagé les grands consommateurs, mais auraient désavantagés les usagers consommant peu. Heyman et al. (1987) ont calculé le tarif polynôme (la combinaison de tarifs binômes) qui aurait assuré qu'aucun type de consommateurs ne soit désavantagé par rapport au tarif uniforme. Leurs résultats apparaissent dans le Tableau 1, de même que les prix proposés par la compagnie NYTel.

Utilisation	Tarif non linéaire proposé par NYTel ($/min)	Tarif polynôme proposé par Heyman et al.	
		Abonnement ($/mois)	Prix unitaire ($/min)
0-60	0.0961	0	0.0756
61-1000	0.0713	0	0.0756
1001-2000	0.0484	0.52	0.0752
2001-7000	0.0352	29.18	0.0674
7001-20000	0.0302	342.17	0.0446
20000+	0.0269	3 495.58	0.0238

Tableau 1 – Comparaison d'un prix non linéaire et d'un tarif polynôme

Les usagers, de par leur consommation, choisissent le tarif auquel ils seront soumis. Le tarif polynôme proposé par Heyman et al. maximise le surplus social et avantage les grands consommateurs sans désavantager les usagers à faible consommation.

Dans certains cas, des tarifs binômes ou polynômes se cachent derrière des services prétendument différents. À titre d'illustration, AT&T, dans les années 1980, vendait un service de téléphonie WATS ("Wide Area Telephone Service") à un tarif présentant des ristournes à partir de certaines quantités consommées, alors qu'il proposait un service "Message Toll" à un prix uniforme, consistant en un prix unitaire élevé. L'argument avancé par la compagnie était que ces services étaient distincts et donc qu'ils pouvaient être vendus à des prix différents. La FCC a cependant jugé que ces deux services avaient tellement de points communs qu'on pouvait considérer qu'ils constituaient un même et unique service et que leur différence de prix devait être justifiée par des différences de coûts seulement. D'après la FCC, c'était un cas de discrimination permettant de faire payer moins cher aux grands consommateurs, ce qui d'après le Communications Act était illégal (Brown and Sibley, 1986).[11]

4.3 Remarques

Certaines caractéristiques se retrouvent de façon récurrente dans la mise en œuvre des tarifs optimaux, qu'ils soient simples ou polynômes. En général, les industries où l'on retrouve ce genre de tarifs sont des monopoles naturels qui, par définition, ont des coûts marginaux de production substantiellement plus faibles que leurs coûts moyens. Il existe différentes méthodes de fixation des prix qui ne correspondent pas toujours à l'application explicite de la méthode de Ramsey-Boiteux (la règle de l'inverse de l'élasticité). Cependant, même lorsque d'autres méthodes sont employées, les éléments essentiels de la méthode de Ramsey-Boiteux sont implicitement compris dans les prix et ceux-ci ne divergent jamais trop des prix optimaux de Ramsey-Boiteux. Lorsque différents types de produits sont concernés, la question de l'équité émerge, puisqu'il existe des possibilités de subventions croisées entre les

[11] On a vu que la discrimination dans les prix pouvait se justifier par des considération d'efficacité économique, ce qui ne veut pas dire qu'elle soit équitable.

différents groupes de consommateurs. Les paramètres des fonctions de coût et de demande sont indispensables à la détermination des prix optimaux. Enfin, notons que ces informations peuvent être obtenues avec plus ou moins de facilité suivant les contextes étudiés.

5 Conclusion

La valorisation des infrastructures passe en partie par la répartition efficace de leur coût. De plus, on exige souvent qu'une infrastructure se finance par la tarification, en totalité ou du moins en partie. Dans ce rapport, nous avons voulu montrer comment on peut approcher cette question de la manière la plus efficace possible en fixant les prix de manière à se rapprocher le plus possible d'un optimum de premier rang. En pratique, diverses contraintes, autres que la contrainte budgétaire, peuvent influencer la tarification. L'approche générale et les principes de tarification que nous avons développés dans ce rapport permettent de considérer des contraintes de toutes sorte.

Un des concepts qui ont été explorés est celui de la tarification à la Ramsey-Boiteux, qui consiste à fixer les prix des différents biens ou services en fonction de l'inverse de l'élasticité de leurs fonctions de demande. Les tarifs polynômes semblent encore plus intéressants. Ils peuvent être construits comme des menus de différentes composantes, où chaque composante comprend une charge fixe (abonnement) et de prix relatifs à différentes caractéristiques de la consommation. L'aspect le plus important de cette approche est que les consommateurs ont la liberté de choisir, dans le menu, la composante qui lui convient le mieux.

Par contre cette tarification n'est pas immune aux critiques au plan de l'équité. De plus, dans un contexte concret, il peut être difficile d'avoir toute l'information requise (par exemple l'information sur les demandes) pour calculer les prix de Ramsey-Boiteux. La tâche risque de s'avérer encore plus difficile s'il s'agit d'établir un menu de tarifs non linéaires.

Annexes

A Les élasticités de la demande

Les prix influencent les recettes de deux façons. Un prix appliqué à un bien (ou service) donné engendre une recette précise mais ce prix a vraisemblablement un impact sur la quantité totale demandée. Ainsi, une augmentation du prix peut accroître ou décroître les recettes. Le résultat final dépend de l'amplitude de la réaction de la demande à un changement des prix. Les économistes ont une mesure de cette amplitude : *l'élasticité de la demande*. Commençons par l'exemple d'une hypothétique demande :

prix	quantité échangée	recettes
1$	24	24$
2$	18	36$
3$	12	36$
4$	6	24$

Dans cet exemple, l'augmentation du prix de 1$ induit une baisse de 6 unités dans la quantité demandée, quel que soit le prix initial considéré. Cependant, l'impact sur la recette dépend de la situation de départ. Le concept d'élasticité permet de saisir cette caractéristique.

Pour un prix donné p, l'*élasticité de la demande*, dénotée $\eta(p)$, est définie par :

$$\eta(p) = -\frac{\text{\% variation de la quantité échangée}}{\text{\% variation du prix}}$$

Formellement, la définition est :

$$\eta(p) = -\frac{\frac{\Delta q}{q}}{\frac{\Delta p}{p}} = -\frac{\Delta q}{q} \times \frac{p}{\Delta p} = -\frac{\Delta q}{\Delta p} \times \frac{p}{q}$$

où q représente la quantité demandée, p le prix unitaire et Δ une variation de la variable (prix ou quantité). On peut réécrire ces formules de la manière suivante :

$$\text{\% variation de demande} = -\eta(p) \times \text{ \% variation du prix, i.e.}$$

$$\frac{\Delta q}{q} = -\eta\left(p\right) \times \frac{\Delta p}{p}. \tag{6}$$

Ainsi, l'élasticité indique de quel pourcentage la demande va varier si le prix varie de 1%.

Le tableau 2 donne les résultats du calcul des élasticités de l'exemple de demande, où $\bar{p}$ représente le prix moyen et $\bar{q}$ la quantité moyenne.

Prix	quantité	recette	Δp	$\bar{p}$	$\frac{\Delta p}{\bar{p}}$	Δq	$\bar{q}$	$\frac{\Delta q}{\bar{q}}$	η
1	24	24							
			1	$\frac{3}{2}$	$\frac{2}{3}$	-6	21	$-\frac{2}{7}$	$\frac{3}{7}$
2	18	36							
			1	$\frac{5}{2}$	$\frac{2}{5}$	-6	15	$-\frac{2}{5}$	1
3	12	36							
			1	$\frac{7}{2}$	$\frac{2}{7}$	-6	9	$-\frac{2}{3}$	$\frac{7}{3}$
4	6	24							

Tableau 2 – Calcul des élasticités

Une élasticité de $\frac{3}{7}$ ou 0.43 indique qu'une variation relative du prix de $\frac{2}{3}$ (1$ par rapport à 1.5$, ou 66%) induit une variation relative de signe opposé de $\frac{2}{9}$ (ou de -29%), c'est-à-dire,

$$-0.43 \times 0.66 = -\frac{3}{7} \times \frac{2}{3} = -\frac{2}{7} = -0.29$$

Remarques

L'élasticité n'est généralement pas une constante. Elle varie en fonction du point de référence, c'est-à-dire en fonction du niveau de demande à partir duquel le changement de tarif est effectué ou considéré. C'est pour cette raison que l'on fait dépendre η de p : $\eta(p)$. Comme la demande q décroît, ou du moins n'augmente pas, lorsque le prix augmente, l'élasticité $\eta(p)$ augmente normalement avec le prix p. C'est le cas dans l'exemple précédent, bien que la pente de courbe de demande soit constante quel que soit le prix p : $\frac{\Delta q}{\Delta p} = -6$. L'élasticité croît avec p car le ratio $\frac{p}{q}$ augmente avec p.

La variation de la recette est aussi liée à l'élasticité. Un accroissement du prix entraîne une augmentation de la recette si et seulement si l'élasticité est inférieure à 1, c'est-à-dire si $\eta(p) < 1$. Une diminution du prix augmente les recettes si et seulement si l'élasticité est supérieure à 1, $\eta(p) > 1$.

La valeur de l'élasticité est essentiellement déterminée par trois facteurs :

- l'existence ou non de proches substituts : l'élasticité est d'autant plus grande qu'il existe des biens substituts, puisque les consommateurs peuvent éviter une augmentation du prix en achetant ces biens substituts.
- l'importance du bien dans le budget ou dans un sous-budget (par exemple, le coût du taxi entre la résidence et l'aéroport dans le coût total d'un déplacement Montréal-Paris ; plus cette part est faible, plus l'élasticité de la demande de service de taxi est faible).
- l'échelle de temps considérée : plus la période étudiée est longue, plus l'élasticité est élevée car plus nombreuses sont les possibilités d'ajustement.

Deux cas particuliers sont à mentionner. Le premier est celui de la demande *parfaitement inélastique*. La demande est alors totalement insensible aux variations de prix, au moins dans l'intervalle des prix considérés. Il est plausible que la demande soit inélastique à court terme, mais cela est moins probable à long terme. Le deuxième cas est celui de l'*élasticité constante*. Par exemple, si la demande peut être exprimée par la fonction $q = A\,p^{-\alpha}$ dans l'intervalle considéré, où A est une constante, alors $\eta(p) = \alpha$, quelles que soient les valeurs de p choisies dans l'intervalle. Un exemple d'une telle fonction est $q = \frac{1}{p}$.

Élasticités croisées

On peut définir d'autres élasticités, par rapport à d'autres variables, comme le revenu ou les prix des autres biens. Par exemple, supposons que la demande pour un bien i dépende non seulement de son prix p_i mais également du prix de $m-1$ autres biens. Formellement, la demande pour le bien i est donnée par $q_i = d_i\left(p_1, \ldots, p_m\right)$. On peut alors définir l'élasticité de la demande pour le bien i par rapport au prix du bien j par :

$$\eta_{ij}\left(p\right) = \frac{\Delta q_i}{\Delta p_j} \times \frac{p_j}{q_i}$$

On parle alors d'élasticité croisée.

B Les prix de Ramsey-Boiteux dans le cas général

Une entreprise produit m biens et la demande pour le bien i est donnée par $q_i = d_i(p_1, \ldots, p_m)$, $i = 1, \ldots, m$. On suppose d_i différentiable. Les élasticités croisées peuvent alors s'écrire : $\eta_{ij}(p) = \dfrac{\partial q_i}{\partial p_j}\dfrac{pj}{q_i}$. Supposons qu'on peut inverser le système de fonctions de demande pour obtenir $p_i = f_i(q_1, \ldots, p_m)$. On peut alors définir les élasticités croisées de ces fonctions de demande inverses : $\varphi_{ij}(q) = \dfrac{\partial p_j}{\partial q_i}\dfrac{q_i}{p_j}$. L'usage veut que, dans le cas général (plus de deux biens), on exprime la règle de Ramsey-Boiteux en termes de ces dernières. De façon précise, les prix de Ramsey-Boiteux sont définis par :

$$\frac{p_i - c_i}{p_i} = \frac{\lambda}{s_i},\ i = 1, \ldots, m$$

où

$$s_i = \frac{1}{\sum_{j=1}^{m} \varphi_{ij} \dfrac{p_j q_j}{p_i q_i}}$$

Si $\varphi_{ij} = 0\ \forall j \neq i$, on a alors $\varphi_{ii} = \dfrac{1}{\eta_{ii}}$, si bien que $s_i = \eta_{ii},\ \forall i$. Autrement, la relation entre les η_{ij} et les φ_{ij} est plus complexe, si bien que chercher à exprimer la formule qui précède en termes des élasticités η_{ij} donnerait une formule très lourde et difficile à interpréter. Dans le cas de deux biens, on peut cependant montrer que cette formule est équivalente à celle donnée par (2).

Références

Albon, R., 1989. "Some observations on the efficiency of British postal pricing", *Applied Economics*, 21, 461-73.

Aumann, R.J. et L.S. Shapley, 1974. *Values of Non-Atomic Games*, Princeton, NJ : Princeton University Press.

Boiteux, M., 1956. "Sur la gestion des monopoles astreints à l'équilibre budgétaire," *Econometrica*, 24, 22-40.

Boyer, M., M. Moreaux, et M. Truchon, 2002a. "Les méthodes de partage de coûts : un survol", CIRANO 2002RP-18.

Boyer, M., M. Moreaux, et M. Truchon, 2002b. "Les méthodes de partage de coûts : propriétés", CIRANO 2002RP-19.

Boyer, M., M. Moreaux, et M. Truchon, 2003. "Partage des coûts dans l'entreprise et incitations", CIRANO 2003RP-05.

Brown, S.J., et Sibley, D.S., 1986. *The theory of public utility pricing*, Cambridge University Press, Cambridge.

Clarke, E., 1971. "Multipart Pricing of Public Goods," *Public Choice*, 11, 17-33.

Dunlop, J.T., 1939. "Price Flexibility and 'The Degree of Monopoly'", *Quarterly Journal of Economics,* 53, 522-534.

Heyman, D., Lazorchak, J., Sibley, D., et Taylor, W., 1987, "An Analysis of Tapered Access Charges for End Users", in Trebing, H. (ed.), *Proceedings of the Eighteenth Annual Williamsburg Conference on Regulation*, Michigan State University Press, East Lansing, MI.

Lerner, A.P., 1934. "The Concept of Monopoly and the Measurement of Monopoly Power", *Review of Economic Studies,* 1, 157-175.

Morrison, S.A., 1982. "The Structure of Landing Fees at Uncongested Airports : An Application of Ramsey Pricing", *Journal of Transport Economics and Policy*, 16(2), 151-59.

Naughton, Mc., 1988. "Regulatory Preferences and Two-Part Tariffs : The Case of Electricity", *Southern Economic Journal*, 55, 743-58.

Ramsey, F.P., 1927. "A Contribution to the Theory of Taxation", *Economic Journal*, 37, 47-61.

Räsänen, M., Ruusunen, J., et Hämäläinen, R.P., 1997. "Optimal tariff design under consumer self-selection", *Enegery Economics*, 19(2), 151-67.

Rothschild, K.W., 1942. "The Degree of Monopoly", *Economica,* 9, 24-39.

Scott, F.A.(JR.), 1986. "Assessing USA Postal Ratemaking : An Application of Ramsey Prices", *Journal of Industrial Economics*, 34(3), 279-90.

Sherman, R., et George, A., 1979. "Second-Best Pricing for the U.S. Postal Service", *Southern Economic Journal*, 45, 685-95.

Train, K.E., 1977. "Optimal Transit Prices under Increasing Returns to Scale and a Loss Constraint", *Journal of Transport Economics and Policy*, 11(2), 185-94.

Train, K.E., 1994. "Self-Selecting Tariffs Under Pure Preferences Among Tariffs", *Journal of Regulatory Economics*, 6, 247-64.

Wilson, R.B., 1993. *Nonlinear Pricing*, Oxford University Press, NY.

Documents * CIRANO *
sur
Le partage des coûts communs et la tarification des infrastructures
http ://www.cirano.qc.ca/publications/

****

[1] 2002RP-17 Le partage des coûts communs : enjeux, problématique et pertinence

[2] 2002RP-18 Les méthodes de partage de coûts : un survol

[3] 2002RP-19 Les méthodes de partage de coûts : propriétés

[4] 2002RP-20 Les jeux de coûts : définitions et propriétés souhaitables des solutions

[5] 2002RP-21 Les jeux de coûts : principaux concepts de solution

[6] 2003RP-04 Le cas des réseaux municipaux souterrains

[7] 2003RP-05 Partage des coûts dans l'entreprise et incitations

[8] 2003RP-06 Tarification optimale des infrastructures communes