

An Elementary Price Index for Internet Service Providers in Canada: A Hedonic Study

Kam Yu*

Prices Division, Statistics Canada
and Ph.D. Candidate
Department of Economics
University of British Columbia
e-mail: kyu@econ.ubc.ca

Paper presented at the Sixth Meeting of the
International Working Group on Price Indices
Canberra, Australia
2-6 April 2001

Abstract

It is well known that late introduction of new Products into the CPI leads to an upward bias in the index. Moreover, improper treatment of quality improvement in the product will add to the upward bias. In this paper a hedonic elementary price index for accessing the Internet in Canada is constructed. We find that the quality-adjusted price index declines at about 15% per year. Detailed data are readily available on-line so that costs of implementation are moderate. We discuss the use of different functional forms in the regression, their ease of use and performance, and compare the resulting index with the matched model approach. Some problems in using the Box-Cox transformation and in handling packages with unlimited access are also discussed.

JEL Classification: C43, C52, E31, L86

Keywords: Price index, Internet access, functional forms, spline functions, hedonic regression, Internet service providers.

*The author wishes to thank Erwin Diewert, Shane Greenstein, Alice Nakamura, Marc Prud'homme, and Jack Triplett for helpful comments and suggestions.

1 Introduction

In the last ten years development in telecommunication technology has been growing at a fast pace. Popularity in personal computers and rapid expansion of the Internet have been hailed by some as a third technological revolution after the agricultural revolution in the medieval time and the industrial revolution that began the modern era. Equity values of the so-called e-commerce companies have been sky rocketing to astonishing levels. One small but key player in this transition period are Internet service providers (ISPs), which can be defined as ‘companies or organizations that act as gateways through which businesses, individuals and organizations can access the World Wide Web’ (Hillary and Baldwin, 1999, 36). Canadians, among others, are on the forefront of this information technology revolution. Expenditures on Internet services have been increasing at an exponential rate. For example, average expenditure per household on Internet service was \$14 or 0.04% of total consumption in 1996.¹ In 1997 it increased to \$29 or 0.08% of total consumption, and in 1998 the expenditure was \$48 or 0.13%.² Eurostat recommends the inclusion of any new product in the CPI if it represents at least 0.1% of total consumption.³ From the above data, this threshold was crossed in 1998. It is expected that the expenditure share of Internet services will maintain the upward trend in the following several years before it starts to level off. The purpose of this paper is to study the feasibility of including Internet services in the CPI.

The ISP sector itself has been undergoing rapid change, both in technology and in market structure. The change in quality of the product poses a difficulty in price measurement. If the quality changes are neglected, the resulting price index will be biased. Many statistical agencies use the matched model approach to address this problem. It involves matching products of identical quality between two periods and comparing their prices. For this reason new products in the current period or products that exist in the base period but are obsolete in the current period are not included in the sample. This can become a problem if the quality is changing at a fast pace as in personal computers and Internet services. Therefore price indices using the matched method can also be biased. In recent years there is a vast literature on the hedonic approach for quality change in durable goods. It allows for the effect of quality change of the goods or services by estimating the shadow (implicit) prices of objectively observable characteristics using regression analysis. In the following sections we will examine the issues in applying the hedonic approach to come up with a quality adjusted elementary price index for Internet services.

The organization of the paper is as follows. In Section 2 we look at Internet services in Canada from the demand side and the supply side. A brief discussion

¹All figures in Canadian funds unless otherwise stated.

²See Statistics Canada (1996, 1997, 1998).

³See Bascher and Lacroix (1999).

of the theoretical foundation of the hedonic method is included in Section 3. Section 4 discusses how we collected the data and some descriptive statistics. Then in Section 5 we present the various functional forms employed in this study and empirical findings. We also compare the hedonic price indices with the conventional matched model index. Recommendations on the methodology in regular production of the price index are provided in the concluding Section 6.

2 Internet Services in Canada

Due to the growth in demand for Internet services, the number of providers has been increasing dramatically. In 1994, there were less than 100 ISPs in Canada. By 1999, there are over 1,000 firms of various sizes, ranging from small companies providing services to small rural communities to large phone and cable companies covering all residential areas in Canada. A number of surveys have been carried out to study the market structure on the production side. A brief discussion of the most recent survey carried out by Statistics Canada for Industry Canada in 1997 is provided here.⁴ With a healthy response rate of 60% (389 out of 675), the survey finds that the industry consists of hundreds of small companies generating relatively little revenue (less than \$50,000 annually) and a handful of big dominating corporations. Three percent of the firms in the survey take in 36% of the total revenue and the top 21% of the firms generate 80% of the revenue. These large companies consist of mainly local phone companies, long distance phone companies, cable companies, and large providers from the U.S. Profits are on the whole positive, but most small firms are losing money. Therefore we expect to see more market consolidation in future years. Recently it appears that broadband Internet connections such as cable modem and digital subscriber lines are surging in popularity. These high speed connections offer dedicated Internet access at an unprecedented speed up to 100 times faster than a regular dial-up service.⁵ One estimate predicts that by 2002 more than half of the households that have Internet access will have broadband connection.⁶

On the consumer side, Statistics Canada has carried out the Household Internet Use Survey since 1997 on a yearly basis.⁷ In 1999, 41.8% of Canadian households are regular users of the Internet compared to 35.9% in 1998 and 29.4% in 1997. The penetration rate for those with access from home has increased substantially from 16% in 1997 to 28.7% in 1999. Other modes of connection include, in 1999, workplace (21.9%), schools (14.9%), and public library (4.5%). Of those respondents using their computers from home, 91.7% use them for e-mails, while 84.7% use them for general browsing. Other usage includes playing

⁴See Hillary and Baldwin (1999) and Khanna (1999).

⁵Staihr (2000).

⁶See Brethour (2000)

⁷See Dickinson and Ellison (2000).

games (42.7%), engaging in chat groups (26.2%), electronic banking (27.7%), purchasing (19%), downloading music (27.1%), and education/training (32%). Higher income households (top 25%) are nearly four times more likely than lower income households (bottom 25%) to be connected. Usage also depends on education level, age, children at home, and provinces (Alberta, BC and Ontario have the highest usage rates while Québec and Newfoundland have the lowest).

3 Hedonic Method: Theoretical Foundation

In this section we examine the relation of quality change with the concept of a constant utility price index, commonly known as a cost of living index. The matched model is in fact what Pollak (1983) calls the goods approach, in which each variety is treated as an separate good. For the hedonic method there are at least three separate classes of economic theories available as background for quality change. We shall discuss each class briefly in this section.⁸ Our ultimate goal is to justify the use of a hedonic index as a proxy for a true cost of living index.

3.1 Household Production

In this theory the consumer is supposed to use market goods $x = (x_1, \dots, x_m)$ to ‘produce’ household commodities $z = (z_1, \dots, z_n)$ for consumption. The household production function $F(x, z) = 0$ is assumed to be neoclassical. Household utility u depends on z only, $u = U(z)$. The household faces the budget constraint

$$y = \sum p_i x_i, \quad i = 1, \dots, m.$$

The essence of the model is a two-stage optimization process. In the first stage, the household minimizes costs given the qualities of z to be produced:

$$C(p, z) = \min\{\sum p_i x_i : F(x, z) = 0\},$$

where p is the price vector for x and $C(p, z)$ is the cost function dual to the production function F . In the second stage the household maximizes utility subject to the budget constraint:

$$V(p, y) = \max\{U(z) : y \geq C(p, z)\},$$

where y is the total consumption expenditure or income and $V(p, y)$ is the indirect utility function. The expenditure function dual to V is

$$E(p, u) = \min\{C(p, z) : U(z) \geq u\}.$$

⁸For details see Pollak(1983) and Triplett (1987).

Denote $u_0 = V(p^0, y_0)$ as the base period utility level, the constant utility price index in period t is defined as

$$P_t = \frac{E(p^t, u_0)}{E(p^0, u_0)}.$$

Lancaster (1971) assumes that F is a Leontief (linear) production function and calls the feasible set of all z the characteristics space.

Suppose there is a quality change from period 0 to period t for market good x_i . This implies that the production function F_t changes in period t . As a consequence the cost function C_t , the indirect utility function V_t , and the expenditure function E_t in period t also change. A Laspeyres-type price index is a local linear approximation in the form (Muellbauer, 1974)

$$P_t^L = \frac{\sum_{i=1}^n \pi_{it} z_{i0}}{\sum_{i=1}^n \pi_{i0} z_{i0}}$$

where $\pi_{it} = \partial C_t / \partial z_{it}$ is the shadow marginal cost of z_{it} in period t .⁹ Similarly a Paasche-type price index formula can be obtained.

3.2 Simple Repackaging

In the simple repackaging case the utility function of a representative consumer is

$$u = U(x_1, \dots, x_m, b) = U(x, b)$$

where b is a quality index of x_1 .¹⁰ In the base period, $b = 1$. The expenditure function can be written as

$$E(p, u, b) = \min\left\{\sum p_i x_i : U(x, b) \geq u\right\}.$$

We want to seek a p_1^* such that

$$E(p_1^*, p_2, \dots, p_m, u, 1) = E(p_1, \dots, p_m, u, b). \quad (1)$$

Note that if $b = 1$ then $p_1^* = p_1$, i.e., no quality change implies no price adjustment. Simple repackaging means that $\partial p_1^* / \partial b$ is independent of x_2, \dots, x_m . This is true if and only if

$$U(x, b) = F(g(x_1, b), x_2, \dots, x_m), \quad (2)$$

⁹Muellbauer (1974, p.988) also shows that the semilog hedonic model is not compatible with the household production framework. The proof relies on the incompatibility of a function being in additive form and at the same time in multiplicative form. But since $\sum_{i=1}^n \pi_{it} z_{i0}$ is only a linear approximation, we cannot conclude that it is incompatible with a multiplicative form.

¹⁰See Fisher and Shell (1972).

that is, b is weakly separable from x_2, \dots, x_m . If this condition is not satisfied, then it may be appropriate to adjust the prices of other goods. For instance, if improvement in the quality of new refrigerators will make ice cream more enjoyable, then it is appropriate to adjust the price of ice cream instead of the price of refrigerators. Most empirical studies on hedonic price indices implicitly assume the above separability condition so that a subindex can be constructed with the product's own characteristics. Muellbauer (1974) argues that since b is a parameter in the utility function, it should be stable over the medium term at least. This is in contrast to yearly regression in the hedonic method,¹¹ where parameters change discretely from year to year. In high-tech sectors like computers and Internet services, however, the technology is developing so fast that consumers' preferences and expectations also change rapidly. Therefore it may be necessary to use yearly regression in these cases.

Assume that the consumer buys only one unit of good 1 ($x_1 = 1$) and let $b = f(z)$ where $z = (z_1, \dots, z_k)$ is a vector of characteristics of good 1. This means that b is an overall quality index of the product. Also, we group all the other goods x_2, \dots, x_m into one aggregate good X with price p_X . Then the utility function in (2) can be redefined as

$$u = U(q(z), X) \tag{3}$$

where $q(z) = g(1, f(z))$. Now (1) can be written as

$$E(p_1^*, p_X, u, 1) = E(1, p_X, u, f(z)) \tag{4}$$

where we have normalized $p_1 = 1$. The *hedonic price* of good 1 is implicitly defined in (4) as

$$p_1^* = \Phi(z, p_X, u). \tag{5}$$

Using a linear approximation for the marginal rate of substitution for $q(z)$ and X , Diewert (2001) derives a hedonic price equation similar to (5), which is independent of u and separable in p_X and z :

$$p_1^* = ap_X f(z)$$

where a is a positive constant. The hedonic equation in (5) only considers the consumer side. In the next section we will look at the producer side as well. The hedonic price function is then a result of market equilibrium.

3.3 Implicit Markets

Rosen (1974) uses spatial models to interpret hedonic regression in a market equilibrium framework. Products with different characteristics are treated as

¹¹Yearly regression will be discussed in Section 5.3.

product differentiation in pure competition. The regression equation is the locus of equilibrium points where the demand curves of different consumers intersect with the supply curves of different suppliers. On the consumer side, the utility function is in the form of (3), i.e., $u = U(q(z), X)$, assumed to be strictly concave, increasing, and differentiable. The price of the aggregate good X is set to unity and the price of the hedonic good is assumed to depend on its characteristics, $p = h(z)$, where h is usually called a hedonic function. The budget constraint faced by a consumer with income y is then $X + h(z) = y$. Define a value function $\Theta(z, u, y) = \theta$ implicitly according to

$$U(q(z), y - \theta) = u.$$

Conditional on income and a utility level, θ is the consumer's willingness to pay for the hedonic good with characteristic z . In fact in the z_i - θ space the graphs of Θ represent a family of indifference curve between the characteristic i and money or the amount of foregone aggregate good X . The partial derivative $\partial\Theta/\partial z_i$ can be interpreted as the inverse demand function for the amount of characteristic i . This implies that Θ is increasing and concave in z_i . For a given income, a consumer's utility is maximized when the demand curve for each characteristic crosses its derived market price and the overall willingness to pay matches the market package price of the hedonic good. That is,

$$\frac{\partial\Theta(z^*, u^*, y)}{\partial z_i} = \frac{\partial h(z^*)}{\partial z_i} \quad (6)$$

for $i = 1, \dots, k$ and

$$\Theta(z^*, u^*, y) = h(z^*) \quad (7)$$

where z^* and u^* are the optimum values. This means that utility are maximized at the point where $\Theta(z, u, y)$ and $h(z)$ are tangent to each other.

On the producer side, assume that each firm produces M number of units of a particular model with characteristic bundle z and has its own input factor prices and technology characterized by a parameter β . The cost function can be written as $C(M, z, \beta)$, which is increasing in M and z . Each firm is a price taker and given the market price $h(z)$ it maximizes profit

$$\pi = Mh(z) - C(M, z, \beta) \quad (8)$$

by choosing M and z . Similar to the value functions of the consumers we can define an offer function $\phi = \Phi(z, \pi, \beta)$ of a firm implicitly from

$$\pi = M\phi - C(M, z, \beta).$$

Notice that M has been eliminated from Φ by the first order condition with respect to M in maximizing π in (8), i.e.,

$$\phi = \partial C(M, z, \beta) / \partial M.$$

Conditional on π , ϕ is the minimum price the firm is willing to accept for producing the hedonic good with characteristic bundle z . Again the partial derivative $\partial\phi/\partial z_i$ can be interpreted as the inverse supply function of characteristic i . Profit is maximized at the point where the offer curve (or surface) is tangent to the hedonic function, i.e.,

$$\frac{\partial\Phi(z^*, \pi^*, \beta)}{\partial z_i} = \frac{\partial h(z^*)}{\partial z_i} \quad (9)$$

for $i = 1, \dots, k$ and

$$\Phi(z^*, \pi^*, \beta) = h(z^*) \quad (10)$$

where z^* and π^* denote the optimal values.

The market is populated by heterogeneous consumers with different taste and incomes on the demand side and by firms with different technologies and input endowments. Each firm finds its own niche by producing a model with a particular characteristic bundle z . Market equilibrium is achieved by equating the optimal conditions (6) with (9) and (7) with (10). The resulting equilibrium path is the market hedonic function $h(z)$.

Since $h(z)$ is the envelope of both the value functions and the offer functions, there is no restriction on its functional form. Its determination is entirely an empirical matter. Nevertheless, by adding identical and independently distributed random variables to the otherwise similar utility functions of different consumers, Feenstra (1995) shows the existence of the aggregate expected demand function and social welfare function. Consequently an exact hedonic price index can be calculated.¹² He also shows that under some assumption for the relationship between the marginal costs and product characteristics, the hedonic regression should be of linear form. A hedonic price index using the log-linear form will be upward biased. Once the hedonic function $h^t(z)$ in period t is determined a theoretical elementary price index or subindex for the hedonic good can be defined as

$$P = \frac{e^1(z^*)}{e^0(z^*)}$$

where

$$e^t(z^*) = \min_z \{h^t(z) : q(z) \geq q(z^*)\}$$

is the period t sub-expenditure function with reference characteristic bundle z^* , usually taken to be the period 0 or period 1 characteristics.

The Internet service market is probably closer to the implicit market model than the other models. Findings from the Household Internet Use Survey suggest that consumers have different needs on their Internet usage. For example, one of the characteristics in the packages is the number of hour, which requires time input from the users, and different users have different time allocation for accessing the Internet. Also, households have different needs for the number of

¹²See Diewert (1976) for a definition of exact index numbers.

e-mails included in their account. On the producer side, the Industry Canada survey indicates that providers differs in size, organization structure, location, and technology (e.g. cable versus dial-up connections). Therefore both consumers and firms are heterogeneous and each is trying to find its own niche in the market.

4 Data Source and Descriptive Statistics

Lists of Internet service providers (ISP) and their web site addresses in 1999 and 2000 are obtained from three sources:

- The membership list of Canadian Association of Internet Providers (CAIP), www.caip.ca/m-list.htm.
- The membership list of Responsible Internet Service Companies (RISC), www.risc.ca/membersa-g.htm.
- The List web site, thelist.internet.com/canada.html.

The prices of the packages offered, together with information on the various features were collected on-line by accessing the web sites of individual companies. In addition to prices, 11 characteristics are identified and used in the hedonic regression.

The prices and other information for the years 1993 to 1995 are obtained from Carroll and Broadhead (1994, 1995, 1996). These are handbooks and directories published at the beginning of the year and therefore the information on the prices and other features are from the previous year. Starting with the 1997 edition, this information is no longer available from this source. After contacting the authors, they admitted that the market had become too volatile, thus making it difficult to properly keep track of the rapidly changing information. Prices and the limited amount of information applicable to our study are however available from Boardwatch Magazine (1996, 1997, 1999) for the year 1996, 1997, and 1998 (hereafter called the Boardwatch data.) Of the information available, two characteristics are identical to the variables used in other years.

Since most ISPs allow customers to register on-line, the listed prices reflect the true transaction prices exclusive of taxes. Special offers targeted to particular groups such as students are not included in the sample. Moreover, we do not include the services offered by so-called free-nets for several reasons. First, free-nets in the early years offered only a text-based system instead of a full graphic environment to the world wide web. Second, they have a very high user-to-line ratio of 1800/1 instead of the usual 15/1 to 30/1 offered by most commercial ISPs. Third, each dial-up session is usually limited to one hour.¹³ For these reasons the nature and quality of services provided by free-nets is very different from

¹³See Teleconsult Limited (1994).

Table 1: Sample Information

Year	No. of Companies	No. of Packages	Mean	St. Dev.	Coef. of Variation
1993	15	27	1.668	1.577	0.946
1994	44	136	1.061	1.073	1.011
1995	95	306	0.927	0.805	0.868
1996	n.a.	257	0.766	0.753	0.982
1997	n.a.	378	0.660	0.683	1.035
1998	n.a.	522	0.573	0.575	1.002
1999	96	559	0.383	0.390	1.020
2000	109	590	0.306	0.295	0.996

that of the commercial ISPs. In recent years some providers have started to offer free access with graphic capacity. But the trade-off is users will be exposed to an advertisement panel on the computer screen which cannot be turned off. Moreover, free internet services do not show up in the expenditure survey. Truly free Internet services, however, do exist in some European countries, and sometimes the services are subsidized by local phone companies.¹⁴ Unlike North American local telephone services, which are flat-fee based, many European countries have metered rates for local calls. Therefore an economic incentive exists for European telephone companies to offer subsidies to providers or even to provide free services of their own.¹⁵

Table 1 gives the number of suppliers and the total number of packages included in each of the sampled years. The price and 11 features of each package are collected. The Boardwatch Magazine data are grouped by area codes of phone numbers so the numbers of companies are unavailable. For comparison purposes, the prices are unit prices expressed as charge per hour of connection time. In Table 1 we also list the means, standard deviations, and the coefficients of variation of the unit price per hour in the sample years. We see that both the means and the standard deviations decrease with time. One possible explanation is that the increasing number of ISPs has made the market more competitive so that prices are decreasing. But the coefficient of variation is fairly stable, indicating that the price variability has remained unchanged.

Next we shall look at some descriptive statistics for the characteristics of various packages. But first we list the names and definitions of the variables as follows.

Dependent variable:

PRICE - the unit price per hour of connection time for the various packages for each ISP.

¹⁴See OECD (2000) for a comparison of Internet service price structure among OECD countries.

¹⁵See Haan (2000) for a theoretical analysis of free Internet access in European countries.

Independent variables (real or integer):

MONTH - number of the prepaid or committed months for the unit price. (0 if bulk package)

BHOUR - number of bulk hours prepaid. Unlike monthly packages, where the unused hours cannot be transferred to the next month, bulk hours have no time limit. (0 for monthly package)

HOUR - number of hours per month in the package. Bulk hours are also included in this variable, making BHOUR a slope dummy in the regression.

SPEED - maximum modem speed supported (kilobit per second).

EMAIL - number of free e-mail accounts included in the package.

WEB - amount of free web page storage space included (MB).

SETUP - amount of set-up fee required (\$).

Dummy independent variables:

ROAM - free roaming hours. (0 = not included, 1 = included)

DEDIC - dedicated connection, i.e., high speed phone or cable connection can free up the phone line and there is no time limit for each login session. (0 = dial-up, 1 = dedicated)

TECH - availability of technical support service. (0 = normal hours, 1 = 24/7)

FNBH - free non-busy hours access. (0 = no, 1 = yes)

BULK - (0 = monthly package, 1 = bulk hours package) generated by the variable BHOUR.

Table 2 lists the means of the characteristics from 1993 to 2000. Notice that the maximum speed of modem (*SPEED*) and the memory provided for personal web sites (*WEB*) are increasing. In fact most of the dial-up services in 1993 have the maximum modem speed of 14.4 kilobit per second (kbps), while in 2000 broadband connections with phone lines and TV cable wires offer maximum speeds exceeding 1000 kbps. Also, average setup fees (*SETUP*) are decreasing. The provision of 24 hours a day, seven days a week technical support services (*TECH*) are also increasing, probably because as the customer base gets larger, more ISPs find this service more cost effective to run, and at the same time there is increasing demand for the service. The provision of free non-business hour connection time (*FNBH*) first appeared in 1995 but dropped to lower levels in 1999 and 2000. It is probably because more and more unlimited time packages are available so that this bonus is no longer needed to attract new customers.

Table 2: Mean values of characteristics and prices of ISPs in Canada, 1993-2000

YEAR	1993	1994	1995	1996	1997	1998	1999	2000
N	27	136	306	257	378	522	559	590
PRICE	1.67	1.06	0.93	0.77	0.66	0.57	0.38	0.31
MONTH	2.74	3.71	4.03	1.00	1.00	1.00	4.43	4.42
BHOUR	1.26	3.46	13.9	0	0	0	12.7	10.0
HOUR	160	116	87	48	59	74	170	221
SPEED	17.5	23.2	29.3	28.8	31.0	35.8	98.1	220
EMAIL	1.11	0.93	1.10	n.a.	n.a.	n.a.	1.64	1.97
WEB	0	0.04	2.14	n.a.	n.a.	n.a.	3.89	5.61
SETUP	47.2	32.7	15.2	n.a.	n.a.	n.a.	12.5	16.0
ROAM	0	0	0.07	n.a.	n.a.	n.a.	0.16	0.08
DEDIC	0.19	0.13	0.03	n.a.	n.a.	n.a.	0.02	0.12
TECH	0	0.01	0.03	n.a.	n.a.	n.a.	0.14	0.18
FNBH	0	0	0.08	n.a.	n.a.	n.a.	0.03	0.04
BULK	0.19	0.07	0.09	0	0	0	0.06	0.06

The Boardwatch data set contains only monthly packages with 28.8 and 56 kbps modem speed and all the other information is missing. Therefore the regressions are run with two independent variable, namely *HOUR* and a dummy for speed for the year 1996 to 1998. For all packages having an unlimited access time, the monthly hour is assumed to be 480, which is based on 16 hours per day times 30 days a month.

Next we compute the chain indices (annual percentage change in parentheses) and the average annual indices (AAI) for the unadjusted geometric mean (UGM) of the unit prices and the matched model in Table 3. In matching the packages we choose the same packages from the same companies so that there is no ‘brand name’ effect in the sample.¹⁶ Matches are not done on the Boardwatch data so packages in 1995 are matched with those in 1999. The average annual change for the matched index is -8.3% compared to -21.2% for the UGM index. This is probably due to the increasing number of unlimited monthly packages, which pulls down the average price from year to year. It is for this reason that we have to adjust the prices of the packages for quality changes. Otherwise the price index will be biased. Also, comparing the UGM index with the average price in Table 1 we see that the latter is slightly upward biased. On the other hand, due to the small percentage of matched samples, the matched model index can be unreliable as well. For these reasons we now turn to the construction of a hedonic price index.

¹⁶We also use the Jevon index (geometric mean of the price ratios) in the matched model.

Table 3: Matched Model Price Index and Unadjusted Geometric Mean Index for ISPs in Canada

Year	1993	1994	1995	1996	1997	1998	1999	2000	AAI
N	27	136	306	257	378	522	559	590	
No. Matched		6	17	-	-	-	38	176	
% Matched		22.2	12.5	-	-	-	12.4	31.5	
Matched Index	1.000	1.063	0.901	-	-	-	0.558	0.547	0.917
		(6.3%)	(-15.2%)	-	-	-	(-11.3%)	(-2.0%)	(-8.3%)
UGM Index	1.000	0.684	0.605	0.527	0.448	0.373	0.238	0.189	0.788
		(-31.6%)	(-11.6%)	(-12.9%)	(-15.0%)	(-16.6%)	(-36.2%)	(-20.7%)	(-21.2%)

5 Hedonic Price Index

In Section 3 we see that by using different theories with different assumptions as the underpinning of the hedonic approach we can justify different functional forms in the regression. Practically there is no *a priori* structural restriction on the choice of functional forms. A lot of studies done on durable goods using the hedonic method are based on the semilog and the log-linear models. These log models have the advantage or convenience that when pooled or adjacent-year regressions¹⁷ are carried out the time dummies can be interpreted as a ‘fixed-weight index’.¹⁸ On the other hand, a simple log model may not be the correct specification due to the possible nonlinear relationship between price and characteristics and interactions between characteristics. For these reasons we use a battery of nested models to find the right model with superior goodness-of-fit but consideration will also be given to the ease of use in routine production of an elementary price index for the CPI. In this section we first discuss an issue in choosing the unit of measurement of prices. Next we define the various functional forms used in the regression. Then we carry out specification tests to determine whether pooled data or yearly data should be used in the regression. The performance of various functional forms is then compared, followed by an introduction to a straight forward computation of the three commonly used elementary price indices. The often used Laspeyres-type and Paasche-type indices will also be computed for comparison.¹⁹

5.1 Unit of Measurement in Prices

Monthly packages are quoted with various numbers of hours of connection time. Sometimes packages may vary in the number of months. For example, a typical monthly package is given a discount if the customer commits for a period of six months or one year. Therefore we have to decide on the unit of price in the regression. There are three choices: the whole package price, the monthly price, or the hourly rate.²⁰ In this study we first use the hourly rate as the independent variable. As the following discussion shows, however, that using the hourly rate may cause a bias in the regression. Therefore we also use the monthly package price as the dependent variable in a number of models. The resulting price indices from the two approaches will be compared.

¹⁷Pooled and adjacent-year regression will be discussed in Section 5.3 below.

¹⁸This term is used by Feenstra (1995).

¹⁹See, example, Berndt, Griliches, and Rappaport (1995) in their study of computer price indices.

²⁰There is some ambiguity between the whole package price and the monthly price. For example, if a package of \$20 per month is given a 10% discount on the condition of a twelve month commitment, we can either treat the package as \$18 per month for twelve months or \$216 for twelve months.

In hedonic studies we try to adjust the price of a commodity or service for its quality, not quantity. *MONTH* and *HOUR* are quantity variables per se. We treat them as characteristics here because of non-linear pricing. That is, we expect the hourly rate to go down as *MONTH* and *HOUR* increase. Therefore if we use the package price in the regression, the linear method will be mis-specified and non-linear terms must be added to the above two variables. There are two consequences of this specification. First, the high R^2 resulting from the highly correlated price and quantity will make model selection less clear. Second, the flexible functional forms (to be described in Section 5.2) will become less efficient because the second order terms now have to accommodate non-linearity in pricing and not non-linearity in characteristics. Also, in collecting prices for elementary price indices, we require the prices to be in the same unit of measurement. This is because if the Dutot index formula (ratio of arithmetic means) is employed, the index will be biased otherwise.²¹ In this study we will also compare the price indices based on the Dutot formula with others. Therefore it seems logical to use the same unit of measurement in price.

There is, however, a potential problem in using the hourly rate as the dependent variable. In calculating the hourly rates for packages with unlimited access, we take the number of hours to be 480 per month, based on the assumption of 16 hours per day times 30 days per month. This number is nevertheless arbitrary and it is possible that the regression result will be different if another number is chosen. Moreover, dividing the monthly package price with such a large number results in a small hourly rate for the unlimited packages. As a consequence we may introduce a negative bias in the estimated coefficient for the variable *HOUR*. For these reasons we also use the monthly package price as the dependent variable in the linear, logarithmic, and Box-Cox models. In order to accommodate nonlinear pricing we employ the spline technique for variable *HOUR* as recommended by Diewert (2001). Instead of estimating one slope coefficient for the variable involved, it assumes a piecewise linear relationship. For example, if we break the variable x into three linear segments, with break points at x_1 and x_2 , the regression equation becomes

$$Y = \begin{cases} \alpha_0 + \alpha_1 x + \sum \beta_i z_i + u & \text{if } 0 \leq x \leq x_1 \\ \alpha_0 + \alpha_1 x_1 + \alpha_2(x - x_1) + \sum \beta_i z_i + u & \text{if } x_1 < x \leq x_2 \\ \alpha_0 + \alpha_1 x_1 + \alpha_2(x_2 - x_1) + \alpha_3(x - x_2) + \sum \beta_i z_i + u & \text{if } x_2 < x \end{cases} \quad (11)$$

where Y is the monthly package price, the z 's are other dependent variables in the regression, the α 's and β 's are the coefficients to be estimated, and u is the disturbance term. Depending on the model in question, x can be the original monthly hour, its log value, or that of the Box-Cox transformation. The last break point is at the maximum number of hour, x_{max} , for all the limited packages in the sample. In this way the last segment of the spline function will

²¹See Diewert (1995), p.15.

be effectively a dummy variable for the unlimited packages. The value that we choose to represent these packages is immaterial as long as it is greater than x_{max} .

5.2 Functional Forms

The functional forms used in this study are listed as follows.

1. Linear Model:

$$Y = \beta_0 + \sum_{i=1}^K \beta_i X_i + u.$$

2. Semilog Model:

$$\log Y = \beta_0 + \sum_{i=1}^K \beta_i X_i + u.$$

3. Log-linear Model:

$$\log Y = \beta_0 + \sum_{i=1}^K \beta_i \log X_i + u.$$

4. Box-Cox (BC) Model:

$$Y^{(\lambda)} = \beta_0 + \sum_{i=1}^K \beta_i X_i + u.$$

where the Box-Cox transformation is defined as

$$Y^{(\lambda)} = \begin{cases} \frac{Y^\lambda - 1}{\lambda} & \text{if } \lambda \neq 0 \\ \log Y & \text{if } \lambda = 0 \end{cases}$$

5. Extended Box-Cox Model:

$$Y^{(\lambda)} = \beta_0 + \sum_{i=1}^K \beta_i X_i^{(\lambda)} + u.$$

6. Restricted Box-Cox-Tidwell (BCT) Model:

$$Y^{(\lambda)} = \beta_0 + \sum_{i=1}^K \beta_i \log X_i + u.$$

7. Generic Flexible Functional Form (Taylor) Model:

$$Y = \beta_0 + \sum_{i=1}^K \beta_i X_i + \frac{1}{2} \sum_{i=1}^K \sum_{j=1}^K \beta_{ij} X_i X_j + u.$$

8. Translog Model:

$$\log Y = \beta_0 + \sum_{i=1}^K \beta_i \log X_i + \frac{1}{2} \sum_{i=1}^K \sum_{j=1}^K \beta_{ij} \log X_i \log X_j + u.$$

9. Restricted Quadratic Box-Cox (RQBC) Model:

$$Y^{(\lambda)} = \beta_0 + \sum_{i=1}^K \beta_i \log X_i + \frac{1}{2} \sum_{i=1}^K \sum_{j=1}^K \beta_{ij} \log X_i \log X_j + u.$$

10. Quadratic Box-Cox Model:

$$Y^{(\lambda)} = \beta_0 + \sum_{i=1}^K \beta_i X_i^{(\alpha)} + \frac{1}{2} \sum_{i=1}^K \sum_{j=1}^K \beta_{ij} X_i^{(\alpha)} X_j^{(\alpha)} + u.$$

Model 7 is a second order Taylor expansion of the independent variables at the origin. In model 6 and 9, we restrict the independent variables so that $\lambda = 0$ in the Box-Cox transformation. This has the advantage of being easy to converge during iterations. Model 7 to 10 are flexible functional forms as defined by Diewert (1993, p.159). Model 10 is an unrestricted quadratic Box-Cox model suggested by Halvorsen and Pollakowski (1981). This is the most general flexible functional form since all the other nine models are nested in it.

5.3 Specification Tests for Structural Change

In general there are three ways to carry out the regression. First, we can estimate the coefficients for each individual year separately. Second, we can pool the observations of two adjacent years together and estimate a common set of coefficients. And finally, we can pool the observations of all years together. Pooled regression should be used in the absence of structural change for two reasons. First, as discussed in Section 3.2 above the coefficients of the regression are parameters in the consumer's utility function and should be stable in the medium term if we base our analysis on the simple repackaging model. Most hedonic studies implicitly assume this because they are concerned with obtaining an elementary price index based on their own characteristics. Second, in constructing a price index from the results of yearly regressions we are in effect performing

out of sample prediction, which can pose a problem if the variances of some estimated coefficients are large.

Chow (F) tests are commonly used to detect any structural change during the adjacent years. The test, however, is only valid for models with a linear specification. For the various Box-Cox models, the individual regressions have different values for λ and therefore different degrees of nonlinearity. The test statistic of the Chow test thus is not an F distribution.²² Also, the Chow test assumes that the variances for different periods are the same, otherwise the computed statistic is not exactly an F distribution. We therefore carry out a series of Goldfeld-Quandt tests on the adjacent-year regressions. Results using a 99% confidence level indicate that homoskedasticity is rejected for about half of the models for at least some periods. This means that the variances of the disturbance terms in those adjacent periods are not the same. Since we do not know the source of heteroskedasticity, it is unlikely that we can correct the problem in the pooled regression. For these reasons we also turn to the Wald test, which in essence tests the equality of the slope coefficients from two separate regressions. The test statistic is asymptotically a χ^2 distribution, even with the presence of heteroskedasticity. Table 4 reports the results from the Chow test and the Wald test for the three logarithmic models. The linear and Taylor models have low goodness-of-fit and therefore are not included. In the table, Y means the test is positive with a significance level of 99% except for 1993-94, where we use 95% due to the small sample size, and N means the test is negative. Conclusions drawn from the two tests are the same except for 1993-94, probably because the asymptotic assumption does not hold for the Wald test. We see that for the 42 tests in total, 22 are positive. Both tests are negative for 1997-98 for the three models.

For the Box-Cox model, we experiment with a two-stage test for structural change. In the first stage, a common value for λ is obtained by an adjacent-year regression. Separate yearly regressions are then run for each of the two years by restricting λ to this common value. The separate regressions are repeated without the restriction. A likelihood ratio test is used to test the equality of the restricted λ specification with the unrestricted one in each year. If equality is rejected in either of the two periods, we conclude that a structural change has occurred; otherwise we go on to the second stage. In the second stage the Chow test and the Wald test are performed with the common λ from the adjacent-year regressions. Our results show that in the first stage all but three adjacent years, namely 1993-94, 1994-95, and 1996-97, have unequal λ 's. In the second stage the period 1994-95 and 1996-97 results for the Chow and the Wald tests are positive. Therefore the estimated coefficients are different when we impose the same λ in the transformation. For the period 1993-94, the Chow test is positive

²²Preliminary results do show that some of the computed F statistics in the Box-Cox models are negative.

Table 4: Chow and Wald Tests for Structural Change

Model	93-94	94-95	95-96	96-97	97-98	98-99	99-2000
<i>Semilog</i>							
Chow	N	N	Y	Y	N	Y	N
Wald	N	N	Y	Y	N	Y	N
<i>Log-linear</i>							
Chow	Y	Y	N	Y	N	N	Y
Wald	N	Y	N	Y	N	N	Y
<i>Translog</i>							
Chow	Y	Y	Y	N	N	Y	Y
Wald	N	Y	Y	N	N	Y	Y

while the Wald test is negative. In what follows we shall base our analysis on the yearly regressions because that is the unrestricted specification. But we shall also calculate the indices from adjacent-year regressions.

For the spline regressions using monthly package prices, the variable *HOUR* is broken up into five segments, with the last one representing the unlimited packages. The break points are selected so that approximately the same number of observations fall within each segment. Wald tests are carried out for the linear, semilog, and log-linear models. Of the total of 21 tests, 16 are positive using a 95% critical value. This indicates that structural changes occur more often when a spline function is used for the monthly hour. For these reason only yearly regressions are used to calculate the price indices for the spline models. Besides the monthly packages, there are bulk packages in the sample years 1994, 1995, 1999, and 2000. Using dummy variables for the bulk packages on top of the spline function would make the regression equations too messy. Therefore the two types of packages are treated differently. At first we consider using seemingly unrelated regressions (SUR) with unequal number of observation to increase the efficiency of the estimation.²³ This technique, however, requires that the equal portions of the observations from the two data sets to be ‘matched in time’ in order to satisfy the assumptions in the structure of the variance-covariance matrix of the disturbance terms. Since the monthly packages and bulk packages in the ISP data are cross-sectional, the choice of the matching observations will certainly affect the estimation results. Therefore SUR is not used and instead separate regressions are run and the resulting indices are aggregated using sample sizes as weights. Due to small sample sizes, only three spline segments are used in the regressions for the bulk packages.

²³See Zellner (1962) and Schmidt (1977).

5.4 Performance of Functional Forms and Regression Results

A number of the models we test here are nested, i.e., the simpler models can be obtained by restricting the parameters to specific values in the more complex models. It is natural that more unrestricted models usually give a higher value for the adjusted coefficient of determination, \bar{R}^2 . In this regard, the flexible functional forms often perform better than the various linear and Box-Cox models, with the exception of the Taylor model. We also use the Akaike Information Criterion (AIC) to compare the various models. With this test, the flexible functional forms lose some of their apparent advantage due to the loss of degree of freedom with the additional second order terms. Table 5 lists the values of \bar{R}^2 , AIC, and the Schwartz Criterion (SC) for all the yearly regressions.²⁴ Using \bar{R}^2 , the various Box-Cox, the translog, and the RQBC models all give very good results. With the AIC, the translog and RQBC both perform very well in early years and in the years with the missing variables, because in those years the number of second order terms is not too high. For the years 1999 and 2000, where we use almost the full set of second order terms, the translog and RQBC models are penalized by the AIC and SC because of the extra terms.²⁵ Compared with the \bar{R}^2 criterion, AIC and SC are heavily biased towards the linear and the log models. Note that the commonly used semilog model is inferior to all the other models except the linear one. The log-linear model, on the other hand, performs satisfactorily compared with the translog model. In future analysis with less severe sample size restrictions we expect the log-linear, Box-Cox, the translog and the RQBC models will all have satisfactory goodness-of-fit.

As mentioned above, the quadratic Box-Cox regression fails to converge. Instead a Box-Cox-Tidwell regression (without the second order terms) is tried. In this model the transformation for the independent variable (*PRICE*) is different from the transformation for the dependent variables. In other words, the transformation parameter λ on the left hand side of the regression equation is different from α on the right hand side. A non-linear iterative maximum likelihood method is used for the yearly regression for the years with the full data set. The iterations converge except for the year 1995. Therefore we cannot compute the price indices but it is interesting to point out that in the 1999 regression the estimated λ and α are -0.776 and 0.512 with *t*-ratios equal to -3.35 and 3.074 respectively. This implies that the linear model with the values of λ and α restricted to (1, 1), the semilog model (0, 1), the log-linear model (0, 0), the Box-Cox model (λ , 1), the extended Box-Cox model (λ , λ), and the restricted Box-Cox-Tidwell Model (λ , 0) should all be rejected.

²⁴Both AIC and SC minimize a loss function of the sum of the square error, therefore the lower the number the better the goodness-of-fit.

²⁵See Greene (2000, p.306) for this point.

Table 5: Model Selection Statistics for the Hourly Price Models

Year	Criterion	Linear	Semilog	Log-linear	Box-Cox	Ex. BC	BCT	Translog	RQBC
1993	\bar{R}^2	0.07	0.11	0.32	0.11	0.32	0.33	0.40	0.41
	AIC	3.01	1.06	0.82	1.07	0.82	0.82	0.72	0.72
	SC	4.63	1.64	1.26	1.65	1.26	1.26	1.10	1.11
1994	\bar{R}^2	0.29	0.53	0.80	0.57	0.80	0.80	0.81	0.80
	AIC	0.88	0.33	0.14	0.35	0.14	0.14	0.13	0.14
	SC	1.12	0.41	0.18	0.44	0.18	0.18	0.17	0.18
1995	\bar{R}^2	0.26	0.59	0.74	0.67	0.74	0.74	0.76	0.76
	AIC	0.49	0.31	0.20	0.34	0.18	0.19	0.19	0.18
	SC	0.57	0.36	0.23	0.40	0.22	0.22	0.22	0.21
1996	\bar{R}^2	0.19	0.62	0.64	0.80	0.68	0.64	0.75	0.77
	AIC	0.46	0.17	0.16	0.19	0.13	0.16	0.12	0.13
	SC	0.48	0.18	0.17	0.20	0.14	0.17	0.12	0.14
1997	\bar{R}^2	0.14	0.62	0.77	0.89	0.78	0.77	0.80	0.81
	AIC	0.40	0.19	0.12	0.22	0.10	0.11	0.10	0.11
	SC	0.42	0.20	0.12	0.23	0.10	0.11	0.11	0.12
1998	\bar{R}^2	0.17	0.65	0.77	0.74	0.79	0.77	0.81	0.81
	AIC	0.27	0.22	0.15	0.29	0.10	0.11	0.12	0.12
	SC	0.28	0.23	0.15	0.29	0.11	0.12	0.13	0.13
1999	\bar{R}^2	0.37	0.82	0.89	0.91	0.90	0.90	0.91	0.91
	AIC	0.10	0.19	0.11	0.41	0.08	0.08	0.10	0.12
	SC	0.11	0.21	0.12	0.45	0.09	0.09	0.12	0.14
2000	\bar{R}^2	0.57	0.85	0.89	0.88	0.91	0.91	0.91	0.91
	AIC	0.04	0.16	0.12	0.31	0.06	0.05	0.10	0.05
	SC	0.04	0.18	0.13	0.34	0.07	0.05	0.12	0.06

To test for multicollinearity and sampling sensitivity, we randomly remove a small set of data and repeat the regression. The estimated coefficients are found to be stable for the logarithmic models and the Box-Cox models. Also, for the purpose of computing the fitted prices and the price indices we re-estimate the coefficients without the extremely insignificant variables (those with a t -ratio less than 1). Due to large number of second order cross terms in the translog and RQBC models, the regression results sometimes are sensitive to the choice of omitted cross terms. This factor should be considered in choosing these models for future regular production of the price index. Because of the additional effort required to fine tune the results, sometimes involving the subjective judgement of the analyst, quality control in production will be more difficult.

Table 6 reports the regression result of the Box-Cox model for 2000. All the significant coefficients have the expected signs. The negative coefficients for *MONTH* and *HOURLY* confirm that pricing is nonlinear. *BULK* and *BHOURLY* have positive coefficients, meaning that consumers pay a premium for the bulk packages and prices for those packages decline less rapidly with hours purchased than the monthly packages. *SPEED* is positive but insignificant as a continuous variable in this model but is significant in a number of other models such as log-linear and translog. Since high speed (broadband) Internet connections are gaining popularity in the last two years and 56 kbps is the standard for most dial-up connection services, we experiment with replacing *SPEED* with a dummy variable for the former.²⁶ The resulting coefficients for this dummy in 2000 are positive and significant in most models. The coefficients for *EMAIL* and *WEB* are insignificant, probably because these services can be obtained free of charge on the Internet. The negative signs for full technical support (*TECH*) and out of town roaming services (*ROAM*) probably reflect a scale economy for the industry. Large national companies can provide those services at low costs and at the same time charge low prices for the packages.

As for the spline models, Table 7 shows that in the early years the performance of all the models are quite uniform according to the adjusted R^2 , but the values drop for the linear model in latter years. The AIC and SC also indicate that the linear model is inferior. The semilog model perform well in this case compared with the log-linear and the Box-Cox models. Goodness-of-fit among the Box-Cox models are very close in all years.

5.5 Elementary Price Indices

In each of the above regression models, the fitted price of each observation can be calculated after the coefficients have been estimated, not only for the year that the observation belongs to, but also for another year using that year's estimated

²⁶Since all high speed connections are dedicated, we drop the dummy *DEDIC* to avoid multicollinearity.

Table 6: Regression Results from the Box-Cox Model, 2000

Variable	Estimated Coefficient	t-statistic 564 DF
MONTH	-3.07×10^{-2}	-5.93
BHOUR	5.84×10^{-3}	11.49
HOUR	-7.76×10^{-3}	-56.72
SPEED	7.36×10^{-5}	0.89
EMAIL	-2.61×10^{-2}	-1.62
WEB	3.60×10^{-3}	0.68
SETUP	-1.43×10^{-3}	-1.65
ROAM	-2.85×10^{-2}	-0.30
DEDIC	1.301	9.18
TECH	-0.112	-1.74
FNBH	-0.187	-1.52
BULK	0.444	3.30
Constant	-0.514	-9.55

Box-Cox regression for $\lambda = -0.41$
 $R^2 = 0.9085$
Adjusted $R^2 = 0.9065$

Table 7: Model Selection Statistics for the Spline Models

Year	Criterion	Linear	Semilog	Log-linear	Box-Cox	Ex. BC	BCT
1993	\bar{R}^2	0.90	0.82	0.86	0.89	0.89	0.90
	AIC	7.74	-0.87	-1.13	2.00	1.18	1.21
	SC	8.19	-0.42	-0.68	2.45	1.63	1.66
1994	\bar{R}^2	0.85	0.88	0.88	0.88	0.88	0.88
	AIC	7.79	-1.85	-1.87	-1.85	-1.87	-1.87
	SC	8.08	-1.56	-1.58	-1.56	-1.58	-1.58
1995	\bar{R}^2	0.78	0.65	0.67	0.62	0.64	0.65
	AIC	6.58	-1.71	-1.77	-2.58	-2.63	-2.57
	SC	6.75	-1.54	-1.60	-2.40	-2.46	-2.40
1996	\bar{R}^2	0.50	0.54	0.65	0.56	0.58	0.60
	AIC	3.80	-1.68	-1.96	1.25	0.82	0.34
	SC	3.89	-1.60	-1.87	1.34	0.91	0.42
1997	\bar{R}^2	0.58	0.64	0.65	0.63	0.63	0.63
	AIC	3.55	-2.37	-2.39	0.32	0.32	0.16
	SC	3.62	-2.30	-2.32	0.40	0.39	0.23
1998	\bar{R}^2	0.47	0.51	0.54	0.51	0.53	0.54
	AIC	3.66	-2.07	-2.15	1.21	0.83	1.09
	SC	3.71	-2.01	-2.09	1.27	0.88	1.15
1999	\bar{R}^2	0.63	0.75	0.74	0.75	0.74	0.74
	AIC	3.53	-2.91	-2.85	-2.91	-2.45	-2.91
	SC	3.64	-2.79	-2.74	-2.79	-2.33	-2.80
2000	\bar{R}^2	0.54	0.77	0.76	0.80	0.79	0.79
	AIC	5.07	-2.59	-2.55	-5.02	-5.02	-5.24
	SC	5.18	-2.48	-2.44	-4.91	-4.91	-5.13

coefficients. For example, the estimated coefficients for 1993, $\hat{\beta}^{93}$, in the yearly regression can be used to evaluate the prices, $\hat{p}_i^{93/94}$, of the observed characteristics in 1994 as if those packages were available in 1993. In other words, we calculate

$$\begin{aligned}\hat{p}_i^{93/93} &= f(X_i^{93}, \hat{\beta}^{93}), \\ \hat{p}_i^{93/94} &= f(X_i^{93}, \hat{\beta}^{94})\end{aligned}$$

where X_i^{93} is observation i of the vector of characteristics in 1993, and f is the functional form used in the regression. Similarly we can calculate $\hat{p}_j^{94/94}$ and $\hat{p}_j^{94/93}$ for all j in 1994. In this way we can come up with the fitted prices of a package in the adjacent years. Using these fitted prices, we can calculate the values of the three commonly used elementary indices, namely the Carli index (arithmetic mean of the price ratios), the Dutot index (ratio of the arithmetic means), and the Jevon index (geometric mean of the ratios²⁷). The performance of these indices can be compared and evaluated. In this way we actually infer the comparison period prices of the products available in the reference period, and vice versa. The symmetric treatment of the two periods makes the resulting index similar in nature to a Fisher-type index. Moreover, it avoids the controversy of the ‘dummy’ price index.²⁸ In passing, we should mention that in all the log models, the ordinary least square assumptions imply that the estimated values for Y is the conditional median, $M(Y|X)$, instead of the conditional mean, $E(Y|X)$ (Goldberger (1968)). The two estimators are related by $M(Y|X) = E(Y|X)e^{-\sigma^2/2}$, where σ^2 is the disturbance variance. Therefore the estimations from the log models are not unbiased. We ignore this factor here since the variances are small and the bias applies to the base years and the comparison years.

Table 8 lists the three elementary chain indices (annual percentage change in parentheses) with the average annual index (AAI) from the semilog, Box-Cox, log-linear, translog, and the RQBC models. For the extended BC and restricted BCT models some values of the non-transformed fitted price are negative and so an index cannot be computed. In Table 8 we see that the Carli index is upwardly biased with respect to the Jevon Indices, as predicted by theory.²⁹ The Dutot index, on the other hand, is unstable. It fluctuates above and below the Carli index. For example, in the Box-Cox model, the 1995-1996 Dutot index is 0.808, exceeding the Carli index at 0.761. In the 1996-1997 indices, however, the reverse is true (0.932 versus 0.953). The Dutot index is sensitive to departure from the homogeneity assumption and the unit of measurement of the product (Diewert, 1995, p.15). Although the unit prices here are quality adjusted they are far from homogeneous. Therefore in hedonic pricing, the Jevon index is the most suitable

²⁷Note that it is also the ratio of the geometric means

²⁸See Triplett (1990) on this point.

²⁹See, for example, Diewert (1995).

Table 8: Hedonic Price Indices for Internet Services in Canada

Model	1993	1994	1995	1996	1997	1998	1999	2000	AAI
<i>Carli Indices</i>									
Semilog	1.000	0.718 (-28.2%)	0.663 (-7.7%)	0.512 (-22.7%)	0.505 (-1.3%)	0.476 (-5.9%)	0.458 (-3.7%)	0.449 (-2.1%)	0.892 (-10.8%)
Log-linear	1.000	0.727 (-27.4%)	0.669 (-7.9%)	0.528 (-21.1%)	0.505 (-4.4%)	0.460 (-8.8%)	0.393 (-14.8%)	0.382 (-2.8%)	0.871 (-12.9%)
Box-Cox	1.000	0.680 (-32.1%)	0.638 (-6.1%)	0.486 (-23.9%)	0.463 (-4.7%)	0.436 (-5.9%)	0.397 (-8.9%)	0.390 (-1.7%)	0.874 (-12.6%)
Translog	1.000	1.038 (3.8%)	0.980 (-5.6%)	0.773 (-21.1%)	0.746 (-3.5%)	0.689 (-7.7%)	0.626 (-9.2%)	0.620 (-0.9%)	0.934 (-6.6%)
RQBC	1.000	1.018 (1.8%)	0.963 (-5.4%)	0.742 (-22.9%)	0.715 (-3.7%)	0.665 (-7.0%)	0.602 (-9.4%)	0.604 (0.2%)	0.930 (-7.0%)
<i>Dutot Indices</i>									
Semilog	1.000	0.677 (-32.3%)	0.603 (-11.0%)	0.495 (-17.8%)	0.449 (-9.5%)	0.404 (-10.0%)	0.339 (-16.1%)	0.316 (-6.6%)	0.848 (-15.2%)
Log-linear	1.000	0.756 (-24.4%)	0.662 (-12.4%)	0.498 (-24.7%)	0.507 (1.8%)	0.475 (-6.3%)	0.386 (-18.9%)	0.365 (-5.3%)	0.866 (-13.4%)
Box-Cox	1.000	0.639 (-36.1%)	0.585 (-8.5%)	0.472 (-19.2%)	0.440 (-6.8%)	0.392 (-11.0%)	0.324 (-17.2%)	0.304 (-6.2%)	0.844 (-15.6%)
Translog	1.000	0.770 (-23.0%)	0.626 (-18.8%)	0.476 (-24.0%)	0.458 (-3.7%)	0.425 (-7.2%)	0.370 (-13.1%)	0.346 (-6.4%)	0.859 (-14.1%)
RQBC	1.000	0.767 (-23.3%)	0.624 (-18.6%)	0.464 (-25.8%)	0.444 (-4.3%)	0.415 (-6.3%)	0.362 (-12.9%)	0.342 (-5.6%)	0.858 (-14.2%)
<i>Jevon Indices</i>									
Semilog	1.000	0.671 (-33.0%)	0.604 (-9.9%)	0.417 (-30.9%)	0.399 (-4.3%)	0.370 (-7.4%)	0.339 (-8.4%)	0.329 (-2.8%)	0.853 (-14.7%)
Log-linear	1.000	0.666 (-33.5%)	0.586 (-12.0%)	0.461 (-21.3%)	0.439 (-4.9%)	0.399 (-9.1%)	0.338 (-15.1%)	0.326 (-3.8%)	0.852 (-14.8%)
Box-Cox	1.000	0.634 (-36.6%)	0.575 (-9.3%)	0.411 (-28.5%)	0.391 (-4.9%)	0.364 (-6.9%)	0.322 (-11.6%)	0.315 (-2.3%)	0.848 (-15.2%)
Translog	1.000	0.735 (-26.5%)	0.641 (-12.8%)	0.493 (-23.1%)	0.474 (-3.7%)	0.437 (-7.9%)	0.388 (-11.3%)	0.378 (-2.6%)	0.870 (-13.0%)
RQBC	1.000	0.726 (-27.4%)	0.634 (-12.7%)	0.477 (-24.8%)	0.459 (-3.8%)	0.425 (-7.3%)	0.377 (-11.4%)	0.371 (-1.6%)	0.868 (-13.2%)

among the three. This result is in line with the choice of most statistical agencies, which employ the Jevon index to aggregate elementary prices.

We see in Table 8 that there is a common downward trend for the ISP price index computed with the different models. For example, if we look at the Jevon index from the log-linear model, we see that there is a rapid decline of prices from 1993 to 1996. The decline exhibits a sudden drop in 1997 of about 5% and picks up slowly in 1998 and 1999. In 2000, prices seem to have stabilized. The average annual price decrease from 1993 to 2000 is 14.8%. In other words, the average Internet access hourly rate in Canada in 2000 is about 33% of what the consumers paid in 1993. We can also compare the hedonic indices with the matched model index and the unadjusted geometric mean index in Table 3. With average annual decreases of 8.3% and 21.2% respectively, the matched model index is upward biased while the UGM index is downward biased relative to the hedonic indices. The bias in the matched model is probably due to the low percentage of matching, albeit we try to match the companies in 2000 with those in 1999. As a result, a lot of information for new discount packages is not included in the index. The UGM index is downward biased because the samples in adjacent periods are not matched, therefore an increase in the number of unlimited packages in the comparison period would induce a bigger price drop.

Since a number of variables are missing in the regression from 1996 to 1998, the resulting indices will be biased. To assess the seriousness of this bias, we compute the same indices using 1995 as the base year and 1999 as the reference year with a full set of variables. These bilateral indices are then compared with the chain indices computed with the missing variable. Table 9 reports the comparison. Since the quality of the service improves over time, we anticipate the chain indices with missing variables be downward biased. But surprisingly most values in Column 4 of Table 9 are positive, indicating an upward bias, although the differences are small. For example, the Jevon chain index and bilateral index from the log-linear model are 0.578 and 0.544 respectively, which differ by a 6% spread. This result can be explained by looking at the individual regressions more closely. Table 10 lists the yearly regression results for the log-linear model of the two years. We see that in 1999 the variables *HOURLY* and *MONTH* are highly significant, while other variables such as *SPEED*, *EMAIL*, *WEB*, *DEDIC*, and *TECH* either have the incorrect signs or are insignificant at the 95% level. The samples from 1996 to 1998 contain monthly packages only and so the variables *MONTH*, *BHOURLY*, and *BULK* are irrelevant.³⁰ And since the other variables besides *HOURLY* do not have a lot of explanatory power, the chain indices from the Boardwatch data with the missing data give satisfactory results.³¹ In economic sense, this implies that Internet service pricing, at least

³⁰For this reason only subsets of the data from 1995 and 1999 are used in computing the 1995-1996 and 1998-1999 price indices. The 1995-1999 bilateral indices, however, are computed using the full data set.

³¹The adjusted R^2 is 0.89 with the full data set, compared with 0.86 with the missing

Table 9: Comparison of the 1995/1999 Bilateral Indices with the Chain Indices

Model	Chain	Bilateral	Difference
<i>Semilog</i>			
Carli	0.691	0.579	0.113
Dutot	0.562	0.532	0.030
Jevon	0.561	0.554	0.007
<i>Box-Cox</i>			
Carli	0.621	0.627	-0.006
Dutot	0.555	0.590	-0.035
Jevon	0.560	0.609	-0.049
<i>Log-linear</i>			
Carli	0.587	0.592	-0.005
Dutot	0.583	0.572	0.011
Jevon	0.578	0.544	0.034
<i>Translog</i>			
Carli	0.638	0.616	0.022
Dutot	0.591	0.578	0.013
Jevon	0.605	0.576	0.029
<i>RQBC</i>			
Carli	0.625	0.610	0.015
Dutot	0.579	0.572	0.007
Jevon	0.594	0.568	0.027

for the years from 1995 to 1999, is largely determined by the packaging of the service. Most of the other characteristics are add-ons that do not affect pricing strategy to a large degree.

We also calculate the Laspeyres-type and Paasche-type indices using the sample (arithmetic) means \bar{X}^k where $k = 0$ for the reference year and $k = 1$ for the comparison year. The Laspeyres-type index is defined as

$$P_L = \frac{f(\bar{X}^0, \hat{\beta}^1)}{f(\bar{X}^0, \hat{\beta}^0)},$$

whereas the Paasche-type index is defined as

$$P_P = \frac{f(\bar{X}^1, \hat{\beta}^1)}{f(\bar{X}^1, \hat{\beta}^0)}.$$

Table 11 lists the Laspeyres and Paasche-type chain indices (annual percentage change in parenthesis) with the average annual indices for some selected models. We see that the indices are sensitive to the choice of the reference year.³² For example, the log-linear Laspeyres average annual change is -12.0%, but the corresponding Paasche index changes by -16.9% annually. The gaps for individual variables.

³²This is observed by Berndt, Griliches, and Rappaport (1995, p.265) in their study of computer prices as well.

Table 10: Regression Results from the Log-linear Model for the Years 1995 and 1999

Variable	1995		1999	
	Estimated Coefficient	t-statistic 293 DF	Estimated Coefficient	t-statistic 546 DF
MONTH	-0.149	-4.69	-0.115	-6.62
BHOUR	0.568	9.05	0.421	7.38
HOUR	-0.661	-25.20	-0.716	-62.11
SPEED	-0.349	-3.38	0.039	0.51
EMAIL	1.043	7.76	-0.018	-0.47
WEB	-0.076	-2.45	-0.025	-1.39
SETUP	-0.012	-0.76	0.012	1.27
ROAM	-0.040	-0.38	0.097	2.37
DEDIC	1.474	8.36	-0.082	-0.25
TECH	-0.104	-0.71	0.051	1.17
FNBH	-0.121	-1.25	0.323	3.95
BULK	-1.389	-5.07	-0.755	-2.63
Constant	2.624	6.96	1.683	5.43

Table 11: Laspeyres and Paasche-Type Hedonic Price Indices

Model	1993	1994	1995	1996	1997	1998	1999	2000	AAI
<i>Laspeyres-type Indices</i>									
Semilog	1.000	0.688 (-31.2%)	0.629 (-8.7%)	0.378 (-39.9%)	0.353 (-6.5%)	0.326 (-7.8%)	0.278 (-14.7%)	0.266 (-4.3%)	0.828 (-17.2%)
Log-linear	1.000	0.844 (-15.6%)	0.740 (-12.3%)	0.587 (-20.8%)	0.564 (-3.8%)	0.524 (-7.1%)	0.435 (-17.1%)	0.409 (-5.9%)	0.880 (-12.0%)
Box-Cox	1.000	0.619 (-38.1%)	0.588 (-5.0%)	0.304 (-48.2%)	0.288 (-5.4%)	0.292 (1.3%)	0.241 (-17.4%)	0.254 (5.3%)	0.822 (-17.8%)
Translog	1.000	1.067 (6.7%)	1.502 (40.8%)	1.123 (-25.3%)	1.079 (-3.9%)	1.012 (-6.2%)	0.846 (-16.4%)	0.804 (-4.9%)	0.969 (-3.1%)
<i>Paasche-type Indices</i>									
Semilog	1.000	0.663 (-33.7%)	0.202 (-69.6%)	0.153 (-24.3%)	0.148 (-2.8%)	0.138 (-7.1%)	0.144 (4.3%)	0.142 (-1.4%)	0.757 (-24.3%)
Log-linear	1.000	0.719 (-28.1%)	0.472 (-34.4%)	0.370 (-21.6%)	0.349 (-5.5%)	0.313 (-10.6%)	0.278 (-11.2%)	0.273 (-1.8%)	0.831 (-16.9%)
Box-Cox	1.000	0.611 (-38.9%)	0.389 (-36.3%)	0.264 (-32.2%)	0.252 (-4.4%)	0.260 (3.0%)	0.245 (-5.6%)	0.266 (8.6%)	0.828 (-17.2%)
Translog	1.000	0.893 (-10.7%)	0.868 (-2.9%)	0.680 (-21.7%)	0.655 (-3.6%)	0.596 (-9.0%)	0.591 (-0.9%)	0.519 (-12.1%)	0.911 (-8.9%)

Table 12: Hedonic Price Indices from Adjacent Year Regressions

Model	1993	1994	1995	1996	1997	1998	1999	2000	AAI
Semilog	1.000	0.671 (-32.9%)	0.641 (-4.6%)	0.467 (-27.1%)	0.441 (-5.6%)	0.406 (-7.9%)	0.401 (-1.3%)	0.386 (-3.7%)	0.878 (-12.7%)
Box-Cox	1.000	0.677 (-32.3%)	0.647 (-4.4%)	0.472 (-27.0%)	0.457 (-3.2%)	0.421 (-7.9%)	0.429 (1.8%)	0.414 (-3.5%)	0.886 (-11.8%)
Log-linear	1.000	0.668 (-33.2%)	0.645 (-3.5%)	0.509 (-21.1%)	0.485 (-4.5%)	0.447 (-8.0%)	0.394 (-11.9%)	0.373 (-5.3%)	0.875 (-13.1%)
Ex. Box-Cox	1.000	0.670 (-33.0%)	0.646 (-3.6%)	0.499 (-22.8%)	0.479 (-4.0%)	0.444 (-7.1%)	0.406 (-8.8%)	0.384 (-5.4%)	0.879 (-12.8%)
BCT	1.000	0.669 (-33.1%)	0.644 (-3.7%)	0.504 (-21.8%)	0.481 (-4.5%)	0.445 (-7.6%)	0.394 (-11.4%)	0.372 (-5.6%)	0.875 (-13.2%)
Translog	1.000	0.692 (-30.8%)	0.711 (2.7%)	0.549 (-22.7%)	0.528 (-3.9%)	0.492 (-6.8%)	0.458 (-6.8%)	0.436 (-5.0%)	0.895 (-11.2%)
RQBC	1.000	0.694 (-30.6%)	0.712 (2.6%)	0.551 (-22.7%)	0.529 (-4.0%)	0.492 (-6.8%)	0.460 (-6.5%)	0.436 (-5.2%)	0.895 (-11.2%)

years are sometimes even wider. There are three possible problems here. First, using the sample (arithmetic) means of the characteristics is equivalent to using a Carli index that may cause biases. Second, all the functional forms used here are nonlinear so that the fitted price of the sample mean is not equal to the mean of the fitted prices. Third, by using the sample means of the reference year only, the price index, in a statistical sense, is not sufficient. That is, it does not make use of all available information. The Jevon index above avoids all three problems here and should be used for future hedonic analysis.

As discussed in Section 5.3, the results from the Chow and Wald tests indicate that structural change is rejected in slightly less than half of the periods we studied. Therefore we also calculate the indices from the adjacent year regressions. Table 12 reports the results. A time dummy is used in each regression to compute the price index. In all the log models, the price indices are simply the anti-log values of the time dummies. For the various Box-Cox models, however, the price indices are functions of the time dummies, values of the characteristics and their estimated coefficients, and λ . Therefore we can in principle compute the three elementary price indices. But in light of the above discussion, we report the Jevon indices only. In Table 12 we see that the resulting indices from different models agree with each other more than those from the yearly regressions. The pooled regressions force the slope coefficients to be the same in the adjacent years and hence do not suffer the out of sample prediction problem discussed above. Therefore when producing a monthly or quarterly price index of ISP for the CPI, where structural change is unlikely to happen, pooled adjacent period regressions should be used. The average annual percentage change ranges from -11.2% for the translog and RQBC models to -13.2% from the BCT model. In the yearly regression, the Jevon index drops more with average annual change, from -13.0% to -15.2% (see Table 8).

Jevon indices computed from yearly regressions using the spline function for *HOUR* are shown in Table 13. These indices follow the similar trends as those of the Jevon indices hedonic indices using hourly prices. The average annual

Table 13: Hedonic Indices using Spline Functions

Model	1993	1994	1995	1996	1997	1998	1999	2000	AAI
Semilog	1.000	0.773 (-22.69)	0.584 (-24.5)	0.441 (-24.5)	0.427 (-3.3)	0.395 (-7.5)	0.356 (-9.9)	0.344 (-3.4)	0.858 (-14.2)
Log-linear	1.000	0.643 (-35.7)	0.477 (-25.8)	0.361 (-24.4)	0.346 (-4.2)	0.320 (-7.4)	0.287 (-10.2)	0.278 (-3.3)	0.833 (-16.7)
Box-Cox	1.000	0.709 (-29.1)	0.523 (-26.2)	0.425 (-18.7)	0.403 (-5.2)	0.376 (-6.6)	0.325 (-13.7)	0.311 (-4.3)	0.846 (-15.4)

changes for the semilog, log-linear, and Box-Cox models are -14.2%, -16.7%, and -15.4%, compared with -14.7%, -14.8%, and -15.2% respectively for the hourly prices. Therefore it seems that the resulting indices are robust with respect to the two different approaches.

5.6 Hedonic Index versus Cost-of-living Index

In view of the discussions in section 3 the hedonic price index can be interpreted as a cost-of-living index. There are, however, two qualifications of such an interpretation. First, in the theory of the cost-of-living index, the preference structure of the consumer is assumed to be the same in the base period and comparison period. But the frequent structural changes we observe here imply that consumer taste is not constant. This is not an unreasonable result given that technology in this sector is rapidly changing and there have been high increases in the growth of Internet users every years. Second, the hedonic analysis does not capture the increased utility from network effects. The increasing population of Internet users induces the creation of more content on the world wide web, and popularizes the use of electronic mails as an important means of communication. These extra benefits of Internet use are the result of positive externalities but are not reflected in the price index.

6 Recommendations and Conclusion

Based on the above findings, we recommend the following actions for the treatment of Internet services in the CPI:

- From the survey of household computer communication usage, the share of Internet services in total consumption has exceeded the recommended threshold of 0.1% in 1998. Therefore this expenditure should be incorporated in the CPI soon.
- The hedonic method for the price index can sometimes be time consuming and expensive due to the amount of information to be collected. The information for the Internet services, however, is available on-line and relatively easy to get. Also, the list prices mostly represent the transaction prices.

We therefore recommend the use of the hedonic method for constructing the elementary price index.

- The observed prices listed on-line are not volatile in the medium term (month-to-month) compared to the prices of other commodities such as grocery items or gasoline. We recommend the survey interval to be three months to six months.
- A detailed set of instructions should be developed regarding the procedure and methodology of the whole process from sampling to computing the price index. This will ensure consistency and the reputation of the CPI and enhance public understanding of the indices.
- Specification tests for structural change should be used to decide whether the regressions should be pooled or separately conducted in each period. Pooled regressions should be carried out whenever possible.
- The log-linear, Box-Cox, translog, and RQBC models give superior performance to the other models we tested. Apart from the difficulty in testing for structural change, the Box-Cox model runs the risk of negative fitted prices, which would pose a problem for index computation. The translog and RQBC have the best goodness-of-fit. In order to avoid the problem of multicollinearity, some second order cross terms in these models are excluded in the regressions. The resulting index, however, can be sensitive to the choice of the excluded cross terms. This can pose a quality control problem in the regular production of the price index. The log-linear model gives satisfactory results compared to the other three models. The simplicity of the computation procedure makes it an attractive choice in regular production.
- The Jevon index from yearly regressions of the log-linear model shows that prices of Internet service providers decreased from 1993 to 2000 at an average rate of 14.8%. The corresponding indices from the adjacent year regression and the matched model are 13.1% and 8.3%. Therefore the matched model has an upward bias on the average. Results using the spline function technique for the number of hours in the packages are very close to the corresponding indices using hourly rates as the dependent variable. The average annual decrease for the spline models is 15.9%.
- The Jevon index (geometric mean of price ratio) is the most stable index among the three most commonly used formulas and is therefore recommended. The Carli index is known to be upward biased and Dutot index is found to be unstable in our study. The Laspeyres-type and Paasche-type indices commonly used in hedonic studies are sensitive to the selected reference characteristics and therefore not recommended.

- Since the market is concentrated in a handful of large companies, it would be desirable to include market shares for weighting in the construction of the index. But consideration must be given to the fact that most other elementary price indices are not weighted and market share information is not always available. Also, to keep operating cost to a reasonable level, it may be desirable to collect information from the large firms (say the top 10%) only. They represent over half of the market share in Internet services.

References

- Bascher, Jérôme and Thierry Lacroix (1999) 'On the Use of Hedonic Methods in the CPI: an Application to Consumer Durables and Apparel', in Walter Lane, ed., *Proceedings of the Fourth Meeting of the International Working Group on Price Indices*, U.S. Department of Labor, Bureau of Labor Statistics.
- Berndt, Ernst, Zvi Griliches, and Neal J. Rappaport (1995) 'Econometric estimates of price indexes for personal computers in the 1990's,' *Journal of Econometrics* 68, 243-268.
- Boardwatch Magazine (1996, 1997, 1999) 'Directory of Internet Service Providers'.
- Brethour, Patrick (2000) 'Fast Web access to dominate by 2002', *The Globe and Mail*, October 2, 2000, B1.
- Carroll, Jim and Rick Broadhead (1994) *Canadian Internet Handbook*, 1994 Edition, Prentice Hall Canada Inc.
- Carroll, Jim and Rick Broadhead (1995) *Canadian Internet Handbook*, 1995 Edition, Prentice Hall Canada Inc.
- Carroll, Jim and Rick Broadhead (1996) *Canadian Internet Directory 1996*, Prentice Hall Canada Inc.
- Dickinson, Paul and Jonathan Ellison (2000) 'Plugging In: The Increase of Household Internet Use Continues into 1999', *Connectedness Series*, November 2000, Catalogue No. 56F0004MIE, No. 1, Statistics Canada.
- Diewert, Erwin (1976) 'Exact and Superlative Index Numbers', *Journal of Econometrics*, Vol. 4, No. 2, 115-145.
- Diewert, Erwin (1993) 'Duality Approaches to Microeconomic Theory', in Diewert, Erwin and Alice Nakamura, ed., *Essays in Index Number Theory*, Vol. 1, North-Holland.

- Diewert, Erwin (1995) 'Axiomatic and Economic Approaches to Elementary Price Indexes', Discussion Paper No.: 95-01, Department of Economics, The University of British Columbia.
- Diewert, W.E. (2001) *Hedonic Regressions: A Consumer Theory Approach*, Discussion Paper No. 01-12, Department of Economics, The University of British Columbia.
- Feenstra, Robert (1995) 'Exact Hedonic Price Indexes', *The Review of Economics and Statistics*, Vol. LXXVII, 634-653.
- Fisher, M. Franklin and Karl Shell (1972) *The Economic Theory of Price Indices*, Academic Press.
- Goldberger, Arthur S. (1968) 'The Interpretation and Estimation of Cobb-Douglas Functions', *Econometrica*, Vol. 35, No. 3-4, 464-472.
- Greene, William H. (2000) *Econometric Analysis*, Fourth Edition, Prentice Hall.
- Haan, Marco (2000) 'The Economics of Free Internet Access', Working Paper, Department of Economics, University of Groningen.
- Halvorsen, Robert and Henry O. Pollakowski (1981) 'Choice of Functional Form for Hedonic Price', *Journal of Urban Economics* 10, 37-49.
- Hillary, Norah and Gord Baldwin (1999) 'Challenges facing Canada's Internet service providers; A snapshot from a survey of ISPs', *Services Indicators*, 1st Quarter, 1999, 35-50, Statistics Canada.
- Khanna, Kris (1999) *Pulse of a Nascent Industry: Theme I of the 1997 Survey of Internet Service Providers: Industry Structure, Composition, and Performance*, Industry Canada.
- Lancaster, Kevin (1971) *Consumer Demand: A New Approach*, Columbia University Press.
- Muellbauer, John (1974) 'Household Production Theory, Quality, and the Hedonic Technique', *American Economic Review*, Vol. 64, No. 6, 977-994.
- OECD (2000) *Local Access Pricing and E-commerce*, DSTI/ICCP/TISP(2000)1/FINAL, OECD, Paris.
- Pollak, Robert (1983) 'The Treatment of Quality in the Cost of Living Index', *Journal of Public Economics*, 20, 25-53.
- Rosen, Sherwin (1974) 'Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition', *Journal of Political Economy*, 92, 34-55.

- Schmidt, Peter (1977) 'Estimation of Seemingly Unrelated Regressions with Unequal Numbers of Observations', *Journal of Econometrics*, 5, 365-377.
- Staihr, Brian (2000) 'The Broadband Quandary for Rural America', *The Main Street Economist*, Center for the Study of Rural America, Federal Reserve Bank of Kansas City, August 2000.
- Statistics Canada (1996) *Family Expenditure in Canada*, Catalogue No. 62-555-XPB.
- Statistics Canada (1997, 1998) *Spending Patterns in Canada*, Catalogue No. 62-202.
- Teleconsult Limited (1994) *Free-Net Strategic and Marketing Plan*, Victoria Free-Net Association, www.bccna.bc.ca/nonbc/sustain.html.
- Triplett, Jack E. (1987) 'Hedonic Functions and Hedonic Indexes', in J. Eatwell, M. Milgate and P. Newman (eds.), *The New Palgrave: A Dictionary of Economics*, The Macmillan Press, 630-634.
- Triplett, Jack E. (1990) 'Hedonic Methods in Statistical Agency Environments: An Intellectual Biopsy', in Ernst R. Berndt and Jack E. Triplett, (eds.), *Fifty Years of Economic Measurement*, The University of Chicago Press, 207-237.
- Zellner, Arnold (1962) 'An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias', *Journal of the American Statistical Association*, Vol. 57, 298, 348-368.