A hedonic model for Internet access service in the Consumer Price Index

A hedonic model is presented for use in making direct quality adjustments to prices for Internet access service collected for the Consumer Price Index; the Box–Cox methodology for functional form selection improves the specification of the model

Brendan Williams

Brendan Williams is an economist in the Transportation and Household Commodities Section, Office of Prices and Living Conditions, Bureau of Labor Statistics. E-mail: williams. brendan@bls.gov

The practice of making hedonic-based price adjustments to remove the effects of quality changes in goods and services that enter into the calculation of the U.S. Consumer Price Index (CPI) has to date focused primarily on indexes for consumer electronics, appliances, housing, and apparel. In an effort to expand the use of hedonic adjustments to a service-oriented area of the CPI, this article investigates the development and application of a hedonic regression model for making direct price adjustments for quality change in the index for Internet access services (known as "Internet services and electronic information providers," item index SEEE03). The analysis presented builds on past research in hedonics and makes use of a Box-Cox regression to select a functional form that allows for better estimation than that produced by standard functional forms. Experimental[1] price indexes are constructed with hedonic regression coefficients to make direct adjustments to CPI price quotes in order to account for changes in characteristics of Internet service access, such as improved bandwidth and length of service contract. These experimental indexes are compared with the official index for Internet access service to measure the impact of hedonic-based quality adjustments on the CPI index SEEE03.

The Internet access industry

The first commercial services allowing users to access content with their personal computers by connecting to interhousehold networks appeared in 1979 with the debut of CompuServe and The Source, an online service provider bought by Reader's Digest soon after the service was launched. The same year also marked the beginning of Usenet, a newsgroup and messaging network. Early online services proliferated during the 1980s, and each allowed users to access a limited network, but not the Internet.

The U.S. Government's ARPANET is commonly cited as the beginning of what we now know as the Internet. The project that developed ARPANET started in the 1960s and provided much of the technological and physical infrastructure for the early Internet. In 1990, ARPANET shut down, and a National Science Foundation network took over where it left off. Taking the final steps to create the Internet, the National Science Foundation expanded the network to commercial traffic and privatized the Internet backbone in the 1990s.

The early Internet lacked a convenient interface. In 1990, researchers at the European Organization for Nuclear Research (Conseil

Européen pour la Recherche Nucléaire, or CERN) developed the World Wide Web, a hypertext-based graphical interface. The World Wide Web provided an easy way to display and organize information that resided on the Internet. With the 1993 introduction of Mosaic, the first popular Web browser, the Internet went mainstream. Many online service providers began including Internet access with their services, and Americans rapidly signed on for such access, mostly through dial-up connections.

In the late 1990s, Internet service providers began to offer high-speed cable and digital subscriber line (DSL) Internet access to consumers. Cable had a significant market share advantage at first, but, according to a May 2006 report by Pew/Internet, DSL has become the broadband access method of choice, with about 50 percent of the broadband market, compared with 41 percent for cable.[2] The same report states that 73 percent of Americans have Internet access in their homes and 42 percent of Americans have broadband Internet access.

Prior hedonic studies of Internet access

Several researchers have developed hedonic models for Internet access. Generally, these models either were focused on dial-up access or were based on a data set that consisted largely of observations on dial-up access. Greg Stranger and Shane Greenstein showed that a hedonic price index for Internet access from November 1993 to January 1999 declines much more than an index that does not account for quality change.[3] Stranger and Greenstein constructed a model with dummy variables for time-limited monthly access, several different levels of hourly limits, different types of speed and forms of access, and each period. Following the time dummy hedonic index method, the coefficients on the time dummy variables are interpreted to represent the quality-adjusted price change. Stranger and Greenstein's hedonic price index covers a timeframe that is too early to include any of the usual forms of consumer broadband access, such as cable or DSL. The closest they come is 1 year of data on T1 access, a technology used predominantly by businesses. Stranger and Greenstein also have data on 64-kbs and 128-kbs Integrated Services Digital Network (ISDN) lines that, while faster than dial-up, do not qualify as broadband.

A paper by Kam Yu and Marc Prud'homme similar to Stranger and Greenstein's produced a hedonic index for Internet access in Canada.[4] The model included variables for speed, dedicated lines, hourly limits, 24-hour technical support, roaming hours, prepaid bulk hours, number of free off-peak hours, number of e-mail addresses, amount of Web storage, and installation fees. Yu' and Prud'homme's index pooled all available types of Internet access, but even in 2000, the last year of the sample, the index was composed primarily of observations for dial-up access. Although the authors utilized time dummy variables, they did not make a straight time dummy index; rather, they used the coefficients from these variables to adjust prices and then computed indexes with the use of the adjusted prices. Like Stranger and Greenstein, Yu and Prud'homme found that the hedonic index decreased faster than nonhedonic indexes; however, Stranger and Greenstein did not use a matched model, whereas Yu and Prud'homme constructed a matched model with few matches, which they acknowledged likely biased their index.[5] Despite the methodological differences between the two papers, both showed that quality-adjusted price indexes for Internet services exhibit larger price declines than those of unadjusted indexes.

Past recommendations for the BLS

The BLS added an elementary price index for Internet access to the CPI in 1997. The Bureau of Economic Analysis funded a 2002 report by Greenstein that made a number of recommendations for improving the Internet access price index.[6] The analysis that follows addresses several of the concerns raised in that report. Greenstein identified six areas in which Internet access issues should be addressed: speed, availability, contract features, reliability, network effects, and other features of users' experiences.[7] The subsequent analysis covers the use of hedonic methods to make direct quality adjustments to prices used in the calculation of the index and so specifically addresses issues within two of Greenstein's areas: speed and contract features. Greenstein also raises weighting, sampling, and other issues that cannot be addressed by a hedonic regression.

Greenstein identifies a number of issues that, though amenable to a hedonic regression, are nonetheless hard to assess. For instance, while consumers benefit from having a larger number of choices in accessing the Internet, there is little reliable data available on local or regional Internet penetration and availability of service. The Federal Communications Commission (FCC) releases data on the number of broadband service providers within a given zip code, but the methods it uses has many critics, including the General Accountability Office, which took issue with those methods in a May 2006 report. As Greenstein wrote, assessing exactly how much a consumer benefits from additional choices, even with good data on service availability, cannot be easily accomplished. Likewise, according to Greenstein, quality change related to service reliability, network effects, and features such as additional e-mail addresses, pop-up ad-blocking software, and in-

stant messaging cannot be reliably estimated. Moreover, many of the extra features that once came as part of a service agreement can now be obtained for free. For example, users can get e-mail accounts with large—even unlimited—data storage limits for free from companies like Google, AOL, and Yahoo. Services for instant messaging, online file storage, picture sharing, and antivirus software also can be had free of charge. With many services now offered free of charge, the aforementioned features do not play as large a role as price-determining characteristics as they once did.

Greenstein also recommended that the CPI use broadband as a comparable replacement item for dial-up once a quality adjustment is applied to account for the improved speed of broadband. Although this issue is amenable to a hedonic regression, making the necessary adjustments would involve creating a hedonic model that covers both broadband and dial-up, and such a model would estimate dial-up and broadband speed with the same continuous function. Past research suggests that dial-up and broadband Internet access can be considered different goods;[8] therefore, their components should not be treated equally.

Another of Greenstein's recommendations was that the CPI should do a better job of taking into account contract features. Greenstein focuses mainly on the issue of contracts with hourly limits; however, he notes that, although such limits were an important feature of Internet contracts in the 1990s, these sorts of agreements have become rare and are probably no longer relevant.[9] Moreover, while some dial-up agreements in the CPI sample from late 2006 still have hourly limits, none of the broadband agreements impose these restrictions.

Although hourly limits no longer play much of a role as a contract feature, broadband service plans often come with set contract lengths. Service agreements in the sample range from 1 to 15 months. Consumers benefit from the greater flexibility of shorter term agreements that do not lock them into one form of service and preclude other options. They also pay a premium for shorter term service agreements. Hedonic quality adjustments for changes in service contract lengths allow the index to reflect the changes in contract value from changes in term-length agreements.

As Greenstein acknowledged, there is no consensus on how to measure Internet access speed.[10] Most Internet users are familiar with bandwidth measures such as 56 kilobits per second or 5 megabits per second. These measures do not fully represent the speed of an Internet connection. Bandwidth indicates only a connection's throughput; it does not give any indication of the connection's latency.

Although throughput measures the amount of information that can be transferred, latency represents the actual speed at which information travels. A frequently used analogy compares Internet access to plumbing. A service with high throughput can be likened to a pipe with a large diameter. Such a pipe can move a large amount of water at once, but the rate of flow might be slow. In order to move a large amount of water quickly, the pipe must both be wide and have a high rate of flow. Similarly, in order to move information quickly, an Internet connection needs to have both high throughput (a larger pipe) and low latency (a fast rate of flow). While most consumers place their focus on throughput, having a low latency connection can be particularly important for certain applications, such as Voice over Internet Protocol (Internet telephony), remote computer access, and gaming, in which the quick relay of information is very important.

Despite the inadequacy of bandwidth as a measure of Internet access speed, no other measures can be readily obtained. For the models estimated and described in this article, bandwidth will serve as a proxy measure for speed. While technically questionable, bandwidth seems a reasonable proxy because Internet service providers generally use estimated upper bandwidth rates when advertising their services, and consumers make their decisions with bandwidth as their primary measure of Internet access speed.

Dial-up and broadband: comparable services?

Although Greenstein recommends that the CPI treat dial-up and broadband as equivalent services (in terms of the value of their bandwidth), a debate has grown over whether the two can be compared as substitutes for each other. Jerry A. Hausman, J. Gregory Sidak, and Hal J. Singer argued that, in the context of government market power regulation, dial-up and broadband are distinct goods that cannot be directly compared.[11] To support treating broadband and dial-up as distinct items, they estimated a regression with the logarithm of cable broadband price as the dependent variable and the logarithm of narrowband price as one of the independent variables. The regression failed to find any statistically significant impact of the price of narrowband on the price of cable broadband in the same area. The authors assert that this finding implies that the two types of Internet access are distinct goods.[12]

A 2002 report by Pew/Internet also concluded that broadband and dial-up users have different Internet usage patterns. Broadband users not only spend more time doing a variety of basic activities online, but are far more likely to use high-bandwidth features such as gaming and

streaming media.[13]

Treating the value of bandwidth as equivalent across dial-up and broadband would disregard the empirical and theoretical evidence indicating that the two Internet services are distinct. Users would be expected to value an increase in broadband bandwidth differently than they would an increase in dial-up bandwidth. Internet users also have different uses for different levels of bandwidth. While lower levels of bandwidth, like those available to dial-up users, may be sufficient for certain activities (such as e-mail, online banking, online shopping, and checking weather reports), users with broadband bandwidth can employ their higher speeds to access content (such as streaming audio-video and gaming) that dial-up users cannot access—at least not without prohibitively long waiting periods. Consumers can be expected to give different values to the different uses of high and low bandwidth. Estimating the value of bandwidth with the same continuous linear function across two distinct levels of bandwidth would likely provide a flawed estimate of bandwidth's value.

Another problem is that dial-up and broadband market structures differ. Tom Downes and Shane Greenstein found that 92 percent of people in the United States live in areas with competitive dial-up markets.[14] In contrast, the market for broadband tends towards a duopoly, with consumers facing the choice between one cable provider and one DSL provider.[15] Although competition among suppliers may not be classified as a consumer preference, such competition will at least affect the price data used in data analysis. Nestor M. Arguea, Cheng Hsiao, and Grant A. Taylor argued that arbitrage would create linear pricing in competitive markets, so a hedonic model can be expected to have a linear functional form.[16] Sherwin Rosen also noted that a hedonic model will be linear if arbitrage in the characteristics is possible.[17] Jack Triplett, by contrast, cautions against the assumption of linearity, because characteristics in hedonic models are rarely truly open to competitive arbitrage. Triplett uses the example of a car and its engine; hypothetically, the two could be bought separately, but such a purchase would be impractical and expensive.[18] Setting the specifics of these arguments aside, past research has shown that market structure relates to functional form in hedonic models. Attempting to fit price data produced in two different market structures with a regression that accommodates only one functional form will lead to misspecification.

In addition, combining dial-up and broadband Internet service into a single model does not make practical sense for the BLS. Setting aside theoretical arguments against quality adjusting for a change from dial-up to broadband service, a regression model covering both types of service would make such an adjustment technically possible; however, the opportunity to make this type of adjustment might never come. There were no cases of substitution between dial-up and broadband services in the 2 years of data examined for this study. Of course, such a result could be expected because the BLS computes the CPI with a "matched-model" method in which prices are collected for the same unique good or service from the same outlet on a repeated basis. Many dial-up providers have no broadband offering, and others offer broadband only within certain geographic areas. Given the tendency of Internet service providers to focus on either dial-up or broadband service, few changes in type of service would be expected within the CPI sample.

Given, then, the differences in market structure of broadband and dial-up (with broadband in a duopolistic market and dial-up in a relatively competitive one), as well as the differences in the way consumers use the two services, combining them into a single model would be theoretically problematic. A combined dial-up and broadband model would have a weaker theoretical foundation and offer little, if any, practical benefit. For these reasons, dial-up and broadband are treated as entirely distinct services in this article, with all analysis focusing on broadband services.

Functional form and the Box-Cox transformation

The theory behind hedonic regression has offered little guidance in selecting the functional form for hedonic models. As mentioned in the previous section, a competitive market implies a linear model if arbitrage is not hindered by bundling, but few markets are truly competitive. Without standards derived from theory, the BLS has generally employed a semilog functional form in the hedonic models it uses to directly adjust prices in the CPI. Other researchers have used goodness of fit as the standard for selecting functional form in hedonic models.[19] In hedonics research, Box-Cox regression has been a particularly popular method of finding an appropriate functional form based on goodness of fit.

Various Box-Cox transformations have been recommended as the preferred functional form for hedonic regressions, in part because they allow for some flexibility. For $Y^{(\lambda)}$, a basic Box-Cox transformation on a single variable, the transformation is defined as

$$Y^{(\lambda)} = \frac{Y^\lambda - 1}{\lambda} \text{ for } \lambda \neq 0 \text{ or} \tag{1}$$
$$Y^{(\lambda)} = \ln Y \text{ for } \lambda = 0.$$

A more complex version transforms both sides of the equation with different parameters. In this article, λ denotes the Box-Cox transformation parameter on the dependent variable while θ denotes the Box-Cox transformation parameter on independent variables. Such a transformation for nonzero values, with logarithms providing the transformation when λ is zero, can be represented as[20]

$$\frac{Y^\lambda - 1}{\lambda} = \alpha + \sum_{i=1}^{K} \beta_i \frac{X_i^\theta - 1}{\theta} + \sum_{s=1}^{J} \gamma_s D_s + \varepsilon \text{ for } \lambda \text{ and } \theta \neq 0. \quad (2)$$

Equation (2) will be referred to as an *unrestricted Box-Cox (uBC) model*, to distinguish it from three other transformations. A *restricted Box-Cox (rBC) model* requires that both sides of the equation, excluding dummy variables, be transformed by the same parameter (that is, rBC = uBC with the restriction that $\lambda = \theta$):

$$\frac{Y^\lambda - 1}{\lambda} = \alpha + \sum_{i=1}^{K} \beta_i \frac{X_i^\lambda - 1}{\lambda} + \sum_{s=1}^{J} \gamma_s D_s + \varepsilon \text{ for } \lambda \neq 0 \text{ or}$$

$$\ln Y = \alpha + \sum_{i=1}^{K} \beta_i \ln X_i + \sum_{s=1}^{J} \gamma_s D_s + \varepsilon \text{ for } \lambda = 0. \quad (3)$$

A *left-hand Box-Cox (lhBC) model* transforms only the dependent variable and leaves the independent variables unaltered:

$$\frac{Y^\lambda - 1}{\lambda} = \alpha + \sum_{i=1}^{K} \beta_i X_i + \sum_{s=1}^{J} \gamma_s D_s + \varepsilon \text{ for } \lambda \neq 0 \text{ or}$$

$$\ln Y = \alpha + \sum_{i=1}^{K} \beta_i X_i + \sum_{s=1}^{J} \gamma_s D_s + \varepsilon \text{ for } \lambda = 0. \quad (4)$$

A *right-hand Box-Cox (rhBC) model* transforms only the continuous independent variables:

$$Y = \alpha + \sum_{i=1}^{K} \beta_i \frac{X_i^\theta - 1}{\theta} + \sum_{s=1}^{J} \gamma_s D_s + \varepsilon \text{ for } \theta \neq 0 \text{ or}$$

$$Y = \alpha + \sum_{i=1}^{K} \beta_i \ln X_i + \sum_{s=1}^{J} \gamma_s D_s + \varepsilon \text{ for } \theta = 0. \quad (5)$$

In each of these models, the statistical software uses an iterative process to select the Box-Cox parameter values with the best fit, based on maximum likelihood. The Box-Cox form accommodates data in multiple functional forms, and certain Box-Cox parameter values are associated with basic functional forms, including the linear, log-log, and semilog forms. An rBC model represents a linear model when the transformation parameter equals 1 ($\lambda = 1$); an rBC model is equivalent to a log-log equation when the transformation parameter equals 0 ($\lambda = 0$). An lhBC model is equivalent to a left-side semilog model when $\lambda = 0$; an lhBC model represents a linear form when $\lambda = 1$. An rhBC model represents a linear form when $\theta = 1$; an rhBC

model is equivalent to a right-side semilog model when $\theta = 0$. An rhBC represents a reciprocal functional form when $\theta = -1$. A uBC model, the most general Box-Cox form used here, can represent any model represented by a uBC, an lhBC, or an rhBC model. As mentioned earlier, a uBC model is an rBC model when it has the restriction that λ must be equal to θ. A uBC model represents an lhBC model when $\theta = 1$; a uBC model represents an rhBC model when $\lambda = 1$.

Box-Cox regression can be used both as a test of functional form and as a form in itself. Because the Box-Cox regression can represent the standard functional forms, it can find whether any of these forms are appropriate and, if so, the one that works the best. For instance, if the Box-Cox regression returns values of 0 for both λ and θ, then a log-log model is indicated. In his handbook on hedonic price indexes, Triplett offers further discussion of the Box-Cox regression as a test of functional form in hedonic models.[21]

If the Box-Cox regression rejects all the parameter values associated with the standard functional forms, the parameter values it returns can still be used to represent alternative forms. The use of Box-Cox transformations as the functional form of choice (and not just a test) in hedonic regression generally receives strong support in the literature. The 1988 work by Maureen L. Cropper, Leland B. Deck, and Kenneth E. McConnell has often been cited for its recommendation of a Box-Cox transformation in hedonic models. In this work, the authors found that a linear Box-Cox function performs better than linear, semilog, double-log, quadratic, and quadratic Box-Cox functions. They also found that a linear Box-Cox function performs well in estimating marginal attribute prices, even in the case of specification error. In contrast, the quadratic Box-Cox form has similar goodness of fit, but provides biased results in the presence of specification error.[22] Cropper and her colleagues attempted only one form of the linear Box-Cox transformation, the uBC,[23] and therefore do not offer any insight into whether the uBC, rBC, or some similar form is the best linear Box-Cox transformation. Without a clear, preferred Box-Cox form defined in the literature, the study described herein uses best-fit criteria to determine the appropriate functional form.

Data

Data for this study were extracted from the official CPI database during November 2006. Data from that month and bimonthly sampled quotes from October were combined into a preliminary data set for the index category

"Internet services and electronic information providers" (formerly known as "other information services"). These data were then pared down into a data set of 139 broadband price quotes covering three types of Internet access. Cable Internet access, with 94 quotes, accounted for 67.6 percent of the data. DSL followed with 41 quotes, or 29.5 percent, and the remaining 4 quotes were for satellite Internet access. In comparison, when the Pew Internet Project first surveyed relative cable and DSL Internet usage in March 2003, it found that 28 percent of broadband subscribers used DSL and 67 percent used cable. In March 2006, the same survey found that DSL's market share had increased to 50 percent while cable's share had fallen to 41 percent.[24] These numbers suggest that the CPI data may be a bit out of step with current trends, but quite representative of the market a few years ago. The close relationship between the CPI sample and the market several years ago should be expected, because the CPI sample rotates continually over a 4-year cycle, so some quotes may be based on expenditure data from several years earlier. Also, the time needed to complete expenditure surveys and incorporate their results into the sample extends this lag.

The four satellite Internet service quotes were dropped from the data set because satellite service does not seem to compete directly with the other forms of broadband. Satellite Internet is more expensive and slower than both DSL and cable broadband. Its market is generally limited to rural areas that lack access to other methods of fast Internet service. Given the differences in market and market structure, the satellite Internet quotes were dropped from the sample used for hedonic regression, leaving 135 quotes in the final data set slated for regression modeling.

The data included several variables in addition to each service plan's price, which in turn included additional fees for services such as modem rental and installation. Each quote had information on a number of service plan characteristics: connection speed, length of the contract, promotional pricing, whether the plan came as part of a bundled package that included cable television and/or telephone service, and more. If information on any of these characteristics was missing or suspicious—such as listing an extremely slow or fast connection speed—the information was verified by going to the service provider's Web page and collecting the proper data value.

The variable "bandwidth" is a continuous measure of the reported download bandwidth in kilobytes per second. In the sample, reported bandwidth ranged from 256 kbps for low-level DSL plans to 10 mbps for the fastest cable connections. Although cable tends to be faster than DSL, it is not always so. The fastest DSL observation was 5 mbps, while the slowest cable observation was 300 kbps.

Many broadband providers offer Internet service in packages bundled with various combinations of television, landline telephone, and mobile telephone services. Observations in the sample were considered to be bundled if the price listed for Internet service was a component of an explicit package offer or if the price was listed at a discount for customers who subscribed to another service. The sample contained no observations bundled with mobile telephone service. Of the paired-service packages, whenever Internet service was bundled with either telephone or subscription television services, all of the observations bundled with television services were from cable broadband providers and all of the observations with telephone service bundling were from DSL providers. Only two "triple-play" packages (packages with Internet, television, and telephone services in a single bundle) were in the sample, and both were from cable companies.

A dummy variable represented television bundling in the regression models. No variable for telephone bundling was used. Preliminary models showed that bundling an Internet service with telephone service did not have a significant impact on the listed price of the Internet service. This finding may be explained in part by the fact that, in order to get DSL service, customers must also pay for a telephone line with their DSL provider. At the time this article was written, very few companies offered stand-alone DSL, known as "naked DSL," and there were no such packages in the sample. Even when not explicitly sold as part of a bundle, DSL service essentially came in tandem with telephone service. Thus, even limiting a dummy variable to representing the telephone service in the triple-play packages did not produce statistically significant results, so only the dummy variable representing bundling with television service was used in the regressions that were carried out for this study.

Most of the observations in the sample represented Internet service from either cable television companies or large telephone companies. A few companies lease communications infrastructure from major broadband providers and sell their own Internet service. The dummy variable "other ISP" indicates an observation with service from one of these providers.

Several different semilog models were specified, and the results from these models are presented in table 1. First, Model 1, consisting of only the theoretical model variables, was estimated. Second, control variables for Census Bureau region and city size, wherever the data were collected, were added to Model 1 to produce Model 2. Finally, after the results of Model 2 were reviewed, Model 3 was specified, using the theoretical model variables and the only significant control variable: the dummy variable for

the Western region.

Four different forms of the Box-Cox transformation were attempted with the variables from Model 3: a transformation on the dependent variable alone (lhBC); a transformation on the continuous, independent variables alone (rhBC); transformations using the same value on both sides of the equation (rBC); and transformations using different values on both sides of the equation (uBC). The results of these transformations are presented in table 2.

The statistical software tests null hypotheses that the Box-Cox parameter(s) for an estimated model is/are equal to −1, 0, or 1. The results from these hypothesis tests can act as tests for functional form. The rBC and uBC results

rejected Box-Cox transformation parameters of −1, 0, and 1. Because a parameter value of 1 represents a linear model and a parameter value of 0 represents a log-log model, the rBC and uBC regression results indicate that the linear and log-log transformations would not be appropriate here. The tests for the lhBC model also rejected λ values of −1, 0, and 1. Because a λ value of 0 represents a semilog model, such a model also can be eliminated as an appropriate functional form. The significance tests for the rhBC transformation model failed to reject any of the parameter values, so that model provided no useful tests of functional form.

As tests of functional form, these Box-Cox regressions eliminated the standard linear, log-log, and semilog forms. While Box-Cox regressions can be used to test functional form, they also can be used as functional forms themselves. Standard functional forms are usually preferred for the sake of parsimony, but the simpler forms were all rejected. Though more complex, the estimated Box-Cox models provide transformations that fit the data best. To help select the appropriate Box-Cox model from the four discussed earlier, the Akaike information criterion (AIC) and the Bayesian information criterion (BIC) were used.[25] As shown in the following tabulation, the rBC had the lowest AIC and BIC values, suggesting that it provides the best transformation:

Model	AIC result	BIC result
Right-hand (rhBC)	877.95950	880.86478
Left-hand (lhBC)	870.44479	873.35006
Restricted (rBC)	868.84819	871.75346
Unrestricted (uBC)	869.89652	875.70707

However, these information criteria are sensitive to differing functional forms, so comparing the values across models is not entirely accurate. The rBC found a significant value for a parameter that transformed both sides of the equation, but the uBC value for the right-hand parameter was not significant. Thus, the rBC seems preferable because it transforms both sides of the equation and does not have an insignificant transformation parameter, as the uBC does.

Table 1. Regression results: semilog models

Variable	Model 1	Model 2	Model 3
Constant	[1]3.837417	[1]3.816153	[1]3.820113
	(80.27)	(53.76)	(84.06)
Bandwidth	[2].000017	[2].0000188	[1].0000185
	(2.10)	(2.14)	(2.42)
Promotional price	[1]−.3865237	[1]−.4407197	[1]−.4366383
	(−9.77)	(−11.08)	(−11.06)
Bundled television	[1]−.1637662	[1]−.1638882	[1]−.1677243
	(−4.05)	(−4.03)	(−4.39)
Contract months	[1]−.0147753	[1]−.0126181	[1]−.0132862
	(−2.98)	(−2.58)	(−2.82)
DSL	[1]−.3636489	[1]−.4271105	[1]−.4137092
	(−7.61)	(−8.62)	(−8.82)
Other ISP	[1]−.2381208	[1]−.234327	[1]−.2100399
	(−3.13)	(−3.09)	(−2.90)
West	—	[1].1847002	[1].1676677
		(3.19)	(3.99)
Midwest	—	.0082532	—
		(.17)	
South	—	.0656907	—
		(1.12)	
Bsize	—	−.0741243	—
		(−1.57)	
Csize	—	−.0316996	—
		(−.48)	
R-squared	.7073	.7456	.74
Adjusted *R*-squared	.6936	.7229	.7256
F-statistic	51.56	32.78	51.63

[1] Significant at the 1-percent level (two-tailed test for control variables, one-tailed test for others).

[2] Significant at the 5-percent level (two-tailed test for control variables, one-tailed test for others).

NOTE: *t*-statistics are in parentheses. West, Midwest, and South are census regions. Dash indicates variable not used in model.

Table 2. Hypothesis tests for Box-Cox transformations

Transformation	λ	θ	H_0 equation	Chi^2 statistic for rejecting H_0 when $X =$			Standard functional forms rejected
				1	0	−1	
lhBC	[1]0.4610551	…	$\lambda = X$	[1]8.72	[1]7.55	[1]81.41	Semilog and linear
rhBC	…	−1.724741	$\theta = X$	1.21	.05	.07	
rBC	[2].401735	[2].401735	$\lambda = \theta = X$	[1]10.32	[2]5.73	[1]78.82	Log-log and linear
uBC	[2].4210553	−.3620293	$\lambda = \theta = X$	[1]11.27	[2]6.68	[1]79.77	Log-log and linear

[1] Significant at the 1-percent level.   [2] Significant at the 5-percent level.

As noted in table 2, the rBC selected 0.401735 as the value of λ that produced the best transformation. The Box-Cox procedure also produced probability values for the coefficients on the basis of chi-square tests, because using ordinary least squares estimates of coefficient variances produces inaccurate measures of significance.[26] The results of this regression are presented in the following tabulation (superscript 1 indicates significance at the 1-percent level, superscript 2 at the 5-percent level):

Variable	Regression result, final model
Constant	8.575593
Bandwidth	[1].0117482
Promotional price	[1]–1.7730443
Bundeled television	[1]–.7251095
Contract months	[2]–.1189097
DSL	[1]–1.675438
Other ISP	[1]–.8505007
West	[1].6512617
λ	[2].401735
p-value for λ	.022

No probability test was run on the constant, but all coefficient values were significant at the 1-percent level except for the coefficient for contract months, which was significant at the 5-percent level.

This estimated rBC model can be used to find implicit prices for the characteristics of an Internet service plan. The price of a characteristic is estimated with the implicit price derived from a hedonic equation. Let

$$\frac{Y^\lambda - 1}{\lambda} = \alpha + \sum_{z=1}^{K} \beta_z \frac{X_z^\theta - 1}{\theta} + \sum_{s=1}^{J} \gamma_s D_s + \varepsilon \text{ for } \theta \text{ and } \lambda \neq 0 \quad (6)$$

be an equation for a uBC. Then the implicit price for a continuous characteristic X_Z is calculated by taking the partial derivative of the price Y with respect to X_Z:

$$\frac{\partial Y}{\partial X_z} = \beta_z X_z^{\theta-1} Y^{1-\lambda}. \quad (7)$$

Or, similarly, for partial derivatives with respect to dummy variable characteristics,

$$\frac{\partial Y}{\partial D_s} = \gamma_s Y^{1-\lambda}. \quad (8)$$

These formulas can be applied to an rBC model by invoking the restriction $\lambda = \theta$. Based on the partial-derivative formula for a continuous variable, the marginal price of bandwidth is

$$\frac{\partial Y}{\partial X_z} = 0.011748 \, X^{0.401735-1} Y^{1-0.401735}. \quad (9)$$

This formula incorporates the original item's price and bandwidth. One can visualize the formula by plotting the marginal price curve of bandwidth (the cost of an increase of 1 kilobit per second) and observing how the resulting curve varies with changes in initial price and bandwidth in a two-dimensional representation. Chart 1 illustrates how the marginal price of bandwidth in this rBC model depends on both the initial price and the initial bandwidth. In the model, the marginal price of bandwidth is higher at lower initial bandwidths and higher at higher initial prices. In contrast, chart 2 illustrates how marginal price in a semilog model (with a logged dependent variable) is dependent upon the initial price only and does not vary with the initial speed. Together, the two charts highlight how the estimated rBC model accommodates the diminishing marginal price of bandwidth while the semilog model does not.

Experimental price index estimation

The theoretical literature on hedonic regression and price indexes presents a variety of methods for incorporating hedonic methods into price indexes. Some of these methods involve creating an entire price index through a hedonic regression, but the BLS uses hedonic regressions to make direct adjustments to prices only when an item (or, in this article, a service) is replaced by a new item (or service).

Price indexes generally use a price relative—the ratio of the current-period price $(P_{a,t})$ for an item a to its price $(P_{a,t-1})$ in the previous period—to measure the change in the price of the item. If item a is phased out and replaced in the current period by an item b, the price of b must be adjusted for the difference in the value of features between a and b. For example, if b is identical to a, except that it includes an improved characteristic Z, then the unadjusted price relative, $P_{b,t}/P_{a,t-1}$, would not take the improvement in b into account. To account for the difference in characteristics, a hedonic model is used to estimate what the price of a in the previous period would have been had a included characteristic Z. This model allows prices from the two periods to be compared as if the same item were being priced in both periods. The adjusted price relative is the ratio of the current-period price of item b to the previous-period price of item a, adjusted by the imputed value, $P_{z,t-1}$, an estimate of the value of characteristic Z. This new price relative can be represented as $P_{b,t} /(P_{a,t-1} +$

Chart 1. Marginal price of bandwidth for a given initial price, Box-Cox model

Price of 1 kbps (dollars)

Price of 1 kbps (dollars)

Legend:
— $75
— $65
— $55
— $45
— $35
— $25
— $15
— $5

Initial bandwidth (kbps)
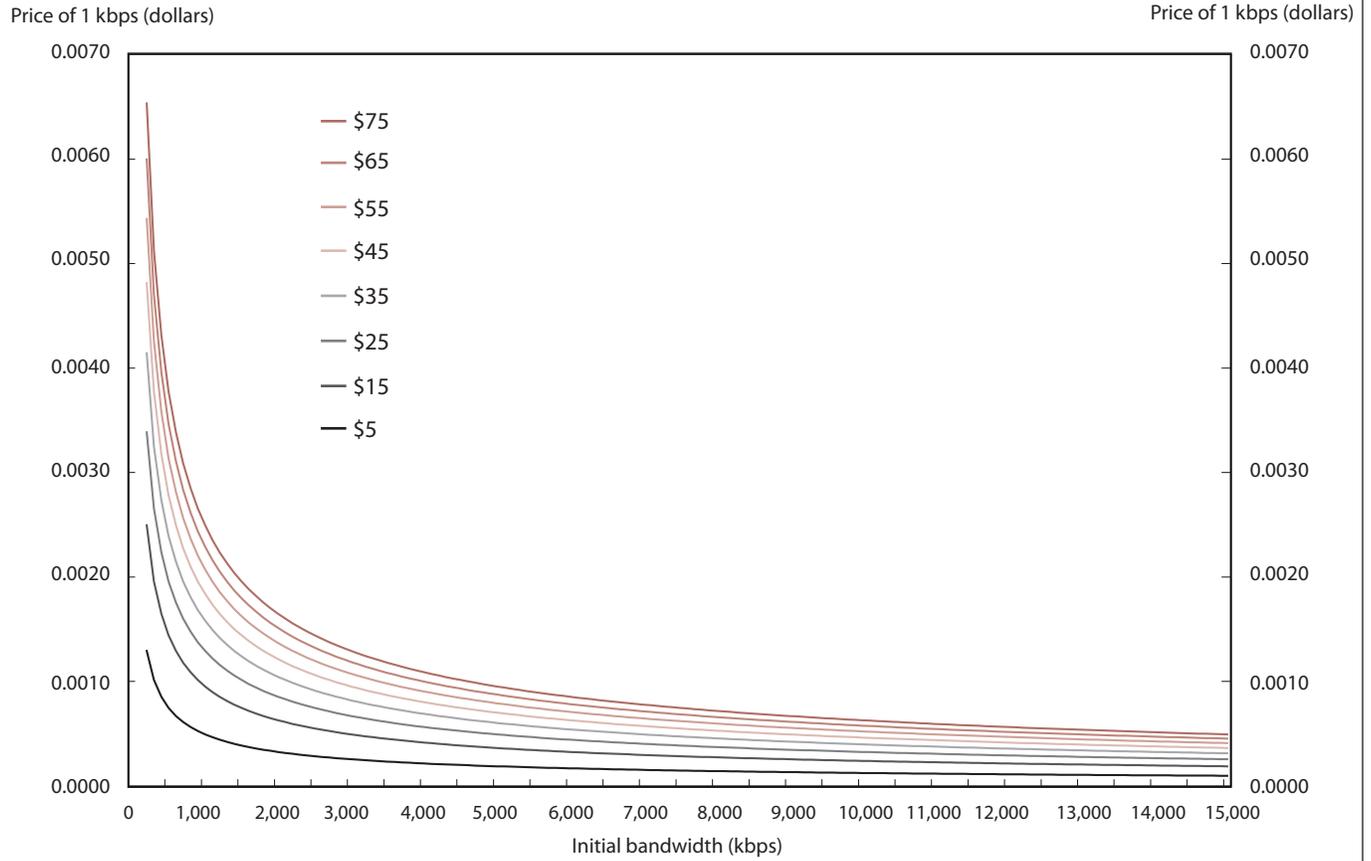
Chart 2. Marginal price of bandwidth for a given initial price, semilog model

Legend: - - - - $5   — — — $15   — — — $25   — $35   — $45   — $55   — $65   — $75

Price of 1 kbps (dollars)

Price of 1 kbps (dollars)

Initial bandwidth (kbps)
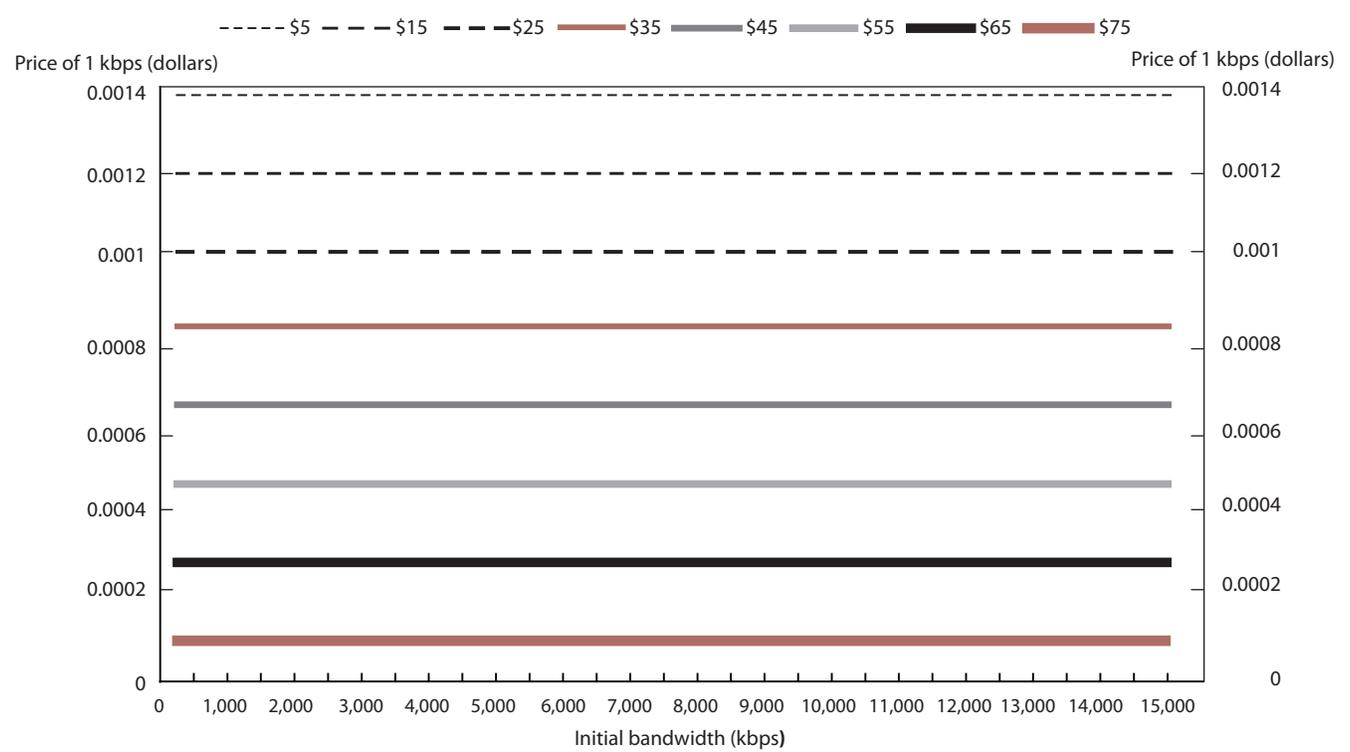
$P_{z,t-1}$). In order to calculate an adjusted price relative, $P_{z,t-1}$, the previous-period value for the new characteristic must be calculated.

The regression coefficient for a variable can be interpreted by taking the partial derivative of the dependent variable with respect to a given independent variable. In a hedonic model, the partial derivative of a characteristic can be used to find an implicit price for a characteristic. One method of incorporating quality adjustments involves using such implicit prices. For dummy variables, the quality adjustment for the addition of a characteristic would simply be the value of the partial derivative (equation 8). For continuous variables, the implicit price is found by calculating the partial derivative (equation 7) and multiplying it by the change in value of a characteristic between an old and a new item:

$$P_{z,t-1} = \frac{\partial y_{t-1}}{\partial x_{z,t-1}}(x_{z,t} - x_{z,t-1}). \tag{10}$$

The total quality adjustment is calculated by adding the quality adjustments for each characteristic:

$$\sum P_{z,t-1} = \sum_{z=1}^{K} \frac{\partial Y_{t-1}}{\partial x_{z,t-1}}(x_{z,t} - x_{z,t-1}) + \sum_{s=1}^{J} \frac{\partial Y_{t-1}}{\partial D_{s,t-1}}(D_{s,t} - D_{s,t-1}). \tag{11}$$

An experimental index was created with this method, with the implicit prices derived from the estimated rBC model presented in the tabulation on page 00. This index will be referred to as the *marginal Box-Cox index*.

A second experimental price index, referred to as the *semilog index*, was created on the basis of the predicted price from Model 3 of table 1. The BLS usually calculates an adjusted price (P_{adjusted}) by taking the item's previous-period price (P_{previous}) and multiplying it by the mathematical constant e to the power of the difference of the sum of the product of the replacement item's characteristics ($X_{z,t}$) and their respective coefficients and the sum of the product of the previous item's characteristics ($X_{z,t-1}$) and their respective coefficients:

$$P_{\text{adjusted}} = P_{\text{previous}} e^{\sum_{z=1}^{k}\beta_z X_{z,t} - \sum_{z=1}^{k}\beta_z X_{z,t-1}}. \tag{12}$$

Equation (12) is derived by dividing the model equation for the predicted price of the replacement item, $P_{\text{replacement}} = e^{\sum \beta z Xz,t+\alpha+\varepsilon}$, by the model equation of the previous price, $P_{\text{previous}} = e^{\sum \beta z Xz,t-1+\alpha+\varepsilon}$. The result is an estimated value for the price of the replacement item, based on the previous price. The process can be viewed as effectively adjusting

the previous-period price for the changes in characteristics. The quality adjustment, which is the sum of the individual values for the changes in characteristics, can be found by subtracting the price of the previous item from the adjusted price, which is the same as the predicted price of the replacement item:

$$\sum P_{z,t-1} = P_{\text{adjusted}} - P_{\text{previous}}. \tag{13}$$

The formula for the semilog index can be used only when the dependent variable (the price in a hedonic regression model) is transformed by a natural logarithm.

A third experimental index, referred to as the *predicted-price Box-Cox index*, was created by developing a formula, similar to equation (12), that relates the previous-period price of an item to the predicted price from a Box-Cox model (note that (θ) denotes a Box-Cox transformation by the parameter θ, while λ is simply the value of the parameter λ):

$$P_{\text{adjusted}} = \left[\lambda(\sum_{z=1}^{K}\beta_z X_{z,t}^{(\theta)} - \sum_{z=1}^{K}\beta_z X_{z,t-1}^{(\theta)}) + P_{\text{previous}}^{\lambda}\right]^{\frac{1}{\lambda}} \text{ for } \lambda \neq 0. \tag{14}$$

Equation (14) was derived by taking the model equation for the replacement item, $P_{\text{replacement}}^{(\lambda)} = \sum_{z=1}^{K}\beta_z X_{z,t}^{(\theta)} + \alpha + \varepsilon$, and subtracting the model equation for the previous-period price, $P_{\text{previous}}^{(\lambda)} = \sum_{z=1}^{K}\beta_z X_{z,t-1}^{(\theta)} + \alpha + \varepsilon$. With the observed previous-period price and the characteristic information for both items substituted into the formula, the formula predicts a price, denoted P_{adjusted}, that represents the previous-period price had the item included the replacement item's characteristics.

The predicted-price method of calculating adjustments provides a more accurate estimate of quality-adjusted prices than does the marginal-price method. The latter calculates the value of a characteristic at an initial point and assumes that the value remains the same. For example, in the rBC model, the value of an additional 1 kbps for a $30/month service plan that already offers 1 mbps (1,000 kbps) can be estimated with equation (9). Substituting 30 for the value of the initial price Y and 1,000 for the value of the initial bandwidth X results in an estimate of $0.001441738 for the marginal value of the bandwidth. If the same plan were increased by 1,000 kbps instead of 1 kbps, the estimated quality adjustment for the increased speed would be 1,000 times $0.001441738, or $1.441738. This calculation assumes that the one-thousandth additional kbps is valued the same as the first additional one. However, the model

predicts that the value of an additional kbps added to a $30/month service with a speed of 1,999 kbps would be $0.000952622, about a third less than the value assumed under a marginal price adjustment.

The Box-Cox predicted-price method (equation 14) avoids the problem of dynamic marginal values, because it is based on undifferentiated Box-Cox models instead of the differentiated version (equation 11) used to calculate marginal prices. These adjustments could be made by taking the model equation and substituting the characteristics of the new item into each variable to find the predicted price of the new item, doing the same to find the predicted value of the old item, and then determining the quality adjustment by taking the difference of the two predicted values. By combining the formulas for the predicted prices of the old and new items, the calculations can be simplified so that only the variables for characteristics that change between the old and new items need to be entered into the price adjustment formula.

Although a predicted-price formula is used to calculate the quality adjustments on the basis of the semilog model, the adjustments will not reflect changes in the value of characteristics, because the semilog model itself assumes that the value of one unit of a characteristic will remain constant no matter the value of a characteristic variable. Going back to the earlier example and using semilog Model 3 indicates that a 1-kbps increase in a $30/month service will be valued at $0.000555 (that is, 0.0000185 × 30), but, unlike the Box-Cox model adjustments, the value for 1 kbps will be the same whether it is added to a 100-kbps service or a 5,000-kbps service.

All item replacements within the item index category "Internet services and electronic information providers" between December 2004 and January 2007 were revaluated in light of the findings of the hedonic models. Forty-four item replacements qualified for adjustment. The coefficients from the Box-Cox (see tabulation on page 40) and semilog (table 1, Model 3) models were utilized to calculate quality-adjusted prices. The results of these adjustments were then used to calculate three experimental indexes corresponding to the three methods of adjustment discussed here: the marginal Box-Cox, predicted-price semilog, and predicted-price Box-Cox adjustments.

The difference between the experimental indexes and the official CPI for this index category is interpreted as a measure of the impact of adjusting for quality change. Table 3 summarizes the three experimental indexes by the type of regression model and the method used for quality adjustment.

The overall impact of these changes was small. The official CPI for the category "Internet services and electronic information providers" fell 24.451 percent between December 2004 and January 2007. In comparison, the marginal Box-Cox, the semilog predicted price, and the predicted-price Box-Cox indexes fell 24.594, 24.612, and 24.575 percent, respectively, over the same period.

The difference between the percent change of the experimental indexes and the percent change of the official index is referred to as a *discrepancy*. The discrepancies produced by the three experimental indexes are listed in table 3. Compared with the official index, the semilog index displayed the largest absolute difference, a downward discrepancy of 0.1613 percentage point over the 2-year period. The marginal Box-Cox index produced a slightly smaller downward discrepancy of 0.1429 percentage point, while the predicted-price Box-Cox index had a slightly smaller discrepancy with the official index, falling 0.1239 percentage point more than the published number.

The experimental indexes decreased more than the official index because they took account of quality change that the official index missed. Of the 44 item replacements that were selected for reevaluation, 40 were originally deemed comparable to the official index. In such cases of comparable replacements, the price change from the old to the new item is treated as if the old item had not been replaced. No quality adjustment was made for these replacements, and the price relative was calculated under the assumption that none of the price change was attributable to quality change. Twenty-nine of the comparable replacements had improvements in bandwidth. In these cases, the price relatives, and thus the official price index, exhibited an upward bias because they did not take into account quality improvements in bandwidth.

Three of the four noncomparable replacements had price relatives imputed by cell-relative imputation, meaning that they were essentially dropped from index calcu-

Table 3.	Summary of experimental indexes		
Experimental index	Model for quality adjustment	Quality adjustment method	Discrepancy with official index over 2 years
(1) BCmarg..	Box-Cox	Marginal price	−0.1429
(2) Semilog...	Semilog	Predicted price	−0.1613
(3) BCpred..	Box-Cox	Predicted price	−0.1239

lations for one period. When a price change is dropped from an index, the price change is basically imputed from the price change in similar items that either were not replaced or had comparable replacements. The remaining replacement had a price change imputed through the class-mean method, an imputation method that uses the price changes from comparable or quality-adjusted replacements to estimate a noncomparable replacement's price change.[27] In his handbook on price indexes, Triplett notes that both class-mean and cell-relative imputation can lead to bias, although the direction of the bias may not be clear and depends on the particular circumstances.[28] Thus, even though the preceding replacements were not treated as comparable, they may still have contributed bias to the official index.

The item replacements in the sample generally show a trend of improvements in service quality in the form of increased bandwidth rates. The official index missed most of this trend because faster service was often treated as comparable to slower service. Using the hedonic adjustments to reevaluate these replacements produces an index that decreases faster than the official index by alleviating at least some of the upward bias created by ignoring the improving quality of Internet service.

Comparing the three experimental indexes reveals that the semilog index, falling more than the other indexes, produces the largest downward discrepancy with the official index. The semilog regression does not accommodate the diminishing marginal price of bandwidth, so the semilog model will produce price estimates that are too low at slow bandwidth rates and too high at high bandwidth rates. Under this model, adjustments are made without regard to the initial amount of bandwidth. For example, given the same initial price, the quality adjustment for increasing a 1-mbps service to 2 mbps will be the same dollar value as the adjustment for increasing a 14-mbps service to 15 mbps. Adjustments to faster services appear to be overestimated, and the semilog index falls too fast as a result.

Similarly, the marginal Box-Cox method seems to be biased downward. Although it does allow for the marginal price to vary with the initial bandwidth rate, it does not account for changes in marginal price in going from one bandwidth rate to another. When there is diminishing marginal price, which is suggested by the model for the bandwidth of interest here, the marginal Box-Cox method will overestimate the price change associated with increased bandwidth.

The predicted-price Box-Cox index decreases faster than the official index because it incorporates many of the quality improvements missed in the official index. However, it decreases less rapidly than the other experimental indexes because it accommodates the diminishing marginal price of bandwidth, whereas the semilog index does not, and the marginal Box-Cox index accommodates diminishing marginal price only in the initial bandwidth. By fully accommodating changes in marginal value, the predicted-price Box-Cox index avoids the downward bias of the other two experimental indexes.

Chart 3 shows the running discrepancies between the experimental indexes and the official index. The discrepancies are given by the percentage-point change in the official index from December 2004 to the given month, subtracted from the percentage-point change in the experimental index over the same period. After several months of consistent downward discrepancies compared with the official index, the experimental indexes began to move higher, closer to the official index. Adjustments made in these months demonstrate why hedonic adjustments will not always push an index downward.

In December 2006, all the experimental indexes increased relative to the official index. The December 2006 change is due entirely to a single replacement wherein the estimated value of increasing a $35-per-month plan's connection speed to 3 mbps from 384 kbps was imputed as $12.27 by the marginal Box-Cox adjustment, $8.48 by the predicted-price Box-Cox adjustment, and $6.69 by the semilog model. The marginal Box-Cox adjustment was the largest because it uses the estimated marginal value of bandwidth at 384 kbps as the estimated value for each 1-kbps increase. The predicted-price Box-Cox adjustment is less than the marginal Box-Cox adjustment because the value of bandwidth is estimated as the estimated difference between bandwidths at 3 mbps and 384 kbps. The semilog adjustment gives the lowest estimated value because it holds the value of bandwidth fixed and does not account for the fact that the value of increased bandwidth added to a very low connection speed will be relatively high. However, none of the models attributes all of the real-world price difference between the two services to the value of greater bandwidth. The faster service was $15 more than the original service it was replacing. Although this replacement was deemed noncomparable in the official index, and its price change was imputed, in the experimental indexes the foregoing estimated values were subtracted from the $15 increase and the remaining price differences were shown as price increases.

In the next month, January 2007, the experimental indexes had another large increase relative to the official index. The increase came from a single replacement in which

Chart 3. Running discrepancy: difference between experimental and official index, cumulative percent changes

Percent discrepancy   Percent discrepancy
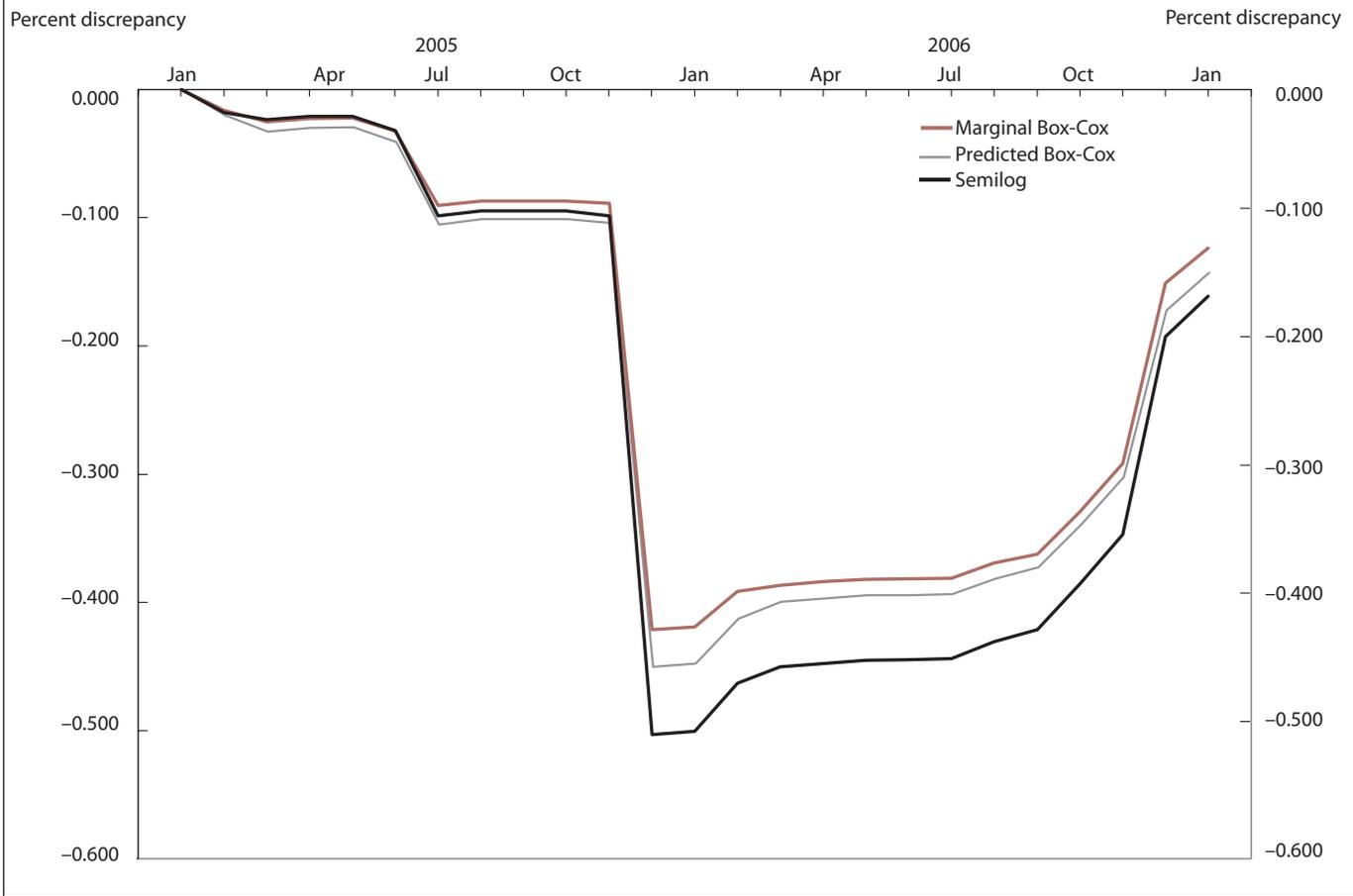
an Internet service package bundled with cable television replaced an à la carte offering. In the official index, this change was considered comparable, so the $17 price decline from the à la carte service to the cheaper, bundled service was reflected in the index. In the experimental indexes, the regression models were used to offset some of this price decline by estimating the expected price difference between Internet service sold à la carte and Internet service bundled with television service. The marginal Box-Cox, semilog, and predicted-price Box-Cox models respectively estimated $8.39, $9.26, and $8.05 price declines. In each case, the associated experimental indexes reflected price decreases by the portion of the $17 decline not offset by these estimates. The official index showed the entire $17 as a price decline, so the hedonic adjustments effectively pushed the experimental indexes upward relative to the official one.

Depending on the circumstances, hedonic adjustments can move an index in either direction. The adjustments used to create the experimental indexes generally showed more downward price movement than the methods used to create the official index, but there were also cases in

which adjustment moved the indexes upward compared with the movement of the official index. A look at the data used to compute the indexes shows that a large number of item replacements with quality improvements were treated as comparable in the official index, so the official index effectively ignored these improvements. The downward movement from incorporating them more than offset the upward adjustments, resulting in all three experimental indexes having downward discrepancies with the official index.

A trio of *Monthly Labor Review* articles compared indexes calculated with and without hedonic adjustments. In one, Paul R. Liegey and Nicole Shepler investigated the effects of hedonic adjustments on indexes for VCR prices from December 1996 to December 1997.[29] They found that the quality-adjusted index fell 8.0 percent over this period, while an unadjusted index fell 8.1 percent, meaning that the quality adjustment actually produced a 0.1-percent upward discrepancy. In another article, Craig Brown and Anya Stockburger looked at the impact of quality adjustments on the CPI apparel indexes. Comparing the official index, which uses direct hedonic-based adjustments,

with an experimental index that lacked these adjustments, they found that the unadjusted experimental index had an upward discrepancy of about 0.2 percent annually.[30]

In a third article, David S. Johnson, Stephen B. Reed, and Kenneth J. Stewart presented a table of the estimated yearly impacts from hedonic models in 10 categories to which the BLS had applied hedonic adjustment since 1998. Instead of using discrepancies, these authors used the percent difference between the hedonic and nonhedonic index levels.[31] The effects of hedonic adjustment ranged from −3.81 percent for computers to 1.89 percent for VCR's, but 6 of the 10 categories had differences between −1.0 percent and 1.0 percent: televisions (−0.11 percent), camcorders (0.15 percent), refrigerators (0.02 percent), clothes washers (−0.78 percent), dryers (0.06 percent), and microwave ovens (−0.17 percent).[32] In comparison, hedonic adjustment for Internet access had an annual effect of approximately −0.06 percent to −0.08 percent (depending upon which model was used), about as much of an absolute effect as that from adjusting dryers.

The adjusted Internet access index changed so little, in part because broadband makes up only a portion of the index. As of November 2006, broadband quotes accounted for about 36 percent of the quotes used to calculate this index. Broadband quotes make up only a portion of the sample used in the adjusted Internet access index, so the effects of broadband quality adjustments are dampened.

Another factor that could be contributing to the absence of any major differences between the quality-adjusted experimental and official indexes is that the quality adjustments are based on a hedonic model developed with data from the end of the period used to create the experimental indexes. The pricing structure of broadband access in November 2006, represented by the model, probably differed significantly from the pricing structure in December 2004. Bandwidth was more expensive in earlier periods and probably had a higher marginal price. If so, using a model based on more recent data underestimated the marginal price of bandwidth and gave low estimates of quality change.

Future developments

The technology behind Internet access has been in constant change since users first signed onto the service in the early 1990s, and this trend will likely continue for the near future. Specifically, two growing forms of Internet access—fiber optics and wireless broadband—will probably radically alter the state of the Internet access market. Optical fiber has long been used in the Internet backbone, but consumers could connect to these high-speed lines only through their slow, household connections. Some service providers have begun running fiber directly to the consumer—a service known as *fiber to the home* (FTTH). Fiber connections offer speeds much faster than those available through cable or DSL.

Whereas fiber offers speed, wireless offers flexibility. Wireless Internet access has been available for several years, but emerging technologies, such as WiMAX, may enable wireless to be competitive as a mainstream form of Internet access. WiMAX cuts the binds of wired Internet by providing a wireless broadband network spread over a large area. WiMAX technology includes both mobile and fixed wireless technologies. Some providers have focused on stationary applications, in which the user would have a stationary connection to a WiMAX router. Stationary WiMAX could be particularly useful to those in rural areas who do not have the wired infrastructure for broadband. Some communication companies have explored the possibilities of mobile WiMAX and have begun deploying WiMAX by installing routers on cell phone towers to create a broadband network with coverage comparable to that afforded by cell phone networks. WiMAX is also only one of several emerging wide-area, wireless broadband technologies. WiMAX has received more attention than the other technologies, but its dominance is not guaranteed.

The impact of new technologies such as FTTH and wireless broadband remains unclear. Depending on pricing and the reliability of service, wireless broadband could compete directly with DSL and cable, or it may be relegated to certain niche markets. Wireless broadband may also reshape the market structure for broadband Internet. Instead of choosing between one cable provider and one DSL provider, consumers may have the added choice of one or more wireless broadband providers. If wireless broadband can compete with current broadband technologies, another hedonic regression model will have to be developed to address the benefit of mobility and the changing marketplace. The expansion of FTTH could also alter the validity of the hedonic model presented in this article. FTTH probably will alter the pricing structure for bandwidth and allow access to higher levels of bandwidth than are currently available to most consumers. The model will then have to be revisited to account for these and other changes in the Internet access market.

BUILDING OFF OF PAST RESEARCH on hedonic regression modeling, this article has developed a model to explain the monthly price of Internet access as a composite of several factors. Coefficients from the model can be used to make

direct price adjustments for changes in quality. Making such adjustments will help account for improvements and other changes to the services in the sample. Given the rapid changes in the Internet access industry, the model will need to be updated periodically, especially as new technology changes the way the Internet is accessed and used.

Past research has indicated that Box-Cox regression provides a better estimation of hedonic models than do more restrictive functional forms. The Box-Cox method offers a relatively easy way to find a suitable transformation for data without having to run many regressions to find the best way to specify the functional form of the model. Of the various Box-Cox forms, a restricted Box-Cox model was found to provide the best fit in this particular case. Estimates from the restricted Box-Cox model were used to create two experimental price indexes utilizing two different price adjustment methods, one based on the change in predicted price with a change in Internet service characteristics and another based on derived implicit prices. A third experimental index was calculated with the current BLS methodology that favors using semilog prices with predicted price adjustments. This article recommends that the BLS adopt, of the experimental methods presented, price adjustments using the predicted-price method based on the Box-Cox model. This model provides the best estimation of a hedonic model for Internet service, and the predicted-price adjustment method is preferable to the alternative methods because it does not assume a fixed marginal price. The Box-Cox model produces more accurate estimates than the semilog model, and adjustments based on the predicted-price method allow the marginal price of a characteristic to vary, unlike adjustments made in accordance with the marginal-price adjustment method, which assumes that the marginal price of a characteristic remains fixed.

The experimental indexes initially showed large downward discrepancies compared with the official index. The experimental indexes accounted for quality improvements that had not been accounted for in the official index, which treated improved, faster Internet service as if it were comparable to slower service. Later observations happened to push the experimental indexes higher. Over the long run, given improving quality, a hedonically adjusted index should decline more than an index that does not account for these quality improvements. It is recommended that hedonic adjustments be made to the official index for Internet service in order to help account for improving quality. Also, the Box-Cox functional form should be adopted in other CPI hedonic regressions, along with predicted price adjustments based on estimated Box-Cox models. □

Notes

[1] The BLS uses the term "experimental" to denote statistics produced outside the regular production systems used for "official" statistics. The experimental indexes are not considered to be of the same quality as the official indexes.

[2] John B. Horrigan, "Home Broadband Adoption 2006," *Pew Internet and American Life Project*, May 28, 2006, on the Internet at **www.pewinternet.org/pdfs/PIP_Broadband_trends2006.pdf**.

[3] Greg Stranger and Shane Greenstein, "Pricing at the On-ramp to the Internet: Price Indexes for ISP's during the 1990s," on the Internet at **www.nber.org/~confer/2003/CRIWf03/greenstein.pdf**.

[4] Kam Yu and Marc Prud'homme, "Econometric Issues in Hedonic Price Indices: The Case of Internet Service Providers," Oct. 12, 2007, on the Internet at **flash.lakeheadu.ca/%7Ekyu/Papers/ISP.pdf** (visited July 15, 2008).

[5] *Ibid*.

[6] Shane Greenstein, "Is the Price Right? The CPI for Internet Access," Report for the Bureau of Economic Analysis, Dec. 20, 2002, on the Internet at **www.kellogg.northwestern.edu/faculty/greenstein/images/htm/Research/WP/Is_the_price_right.pdf**.

[7] *Ibid*., p. 9.

[8] See, for example, Jerry A. Hausman, J. Gregory Sidak, and Hal J. Singer, "Cable Modems and DSL: Broadband Internet Access for Residential Customers," Papers and Proceedings of the Hundred Thirteenth Annual Meeting of the American Economic Association, *American Economic Review*, May 2001, pp. 302–07.

[9] Greenstein, "Is the Price Right?" p. 20.

[10] *Ibid*., p. 14.

[11] Hausman, Sidak, and Singer, "Cable Modems and DSL."

[12] *Ibid*., p. 340.

[13] John B. Horrigan and Lee Rainie, "The Broadband Difference: How online Americans' behavior changes with high-speed Internet connection at home," *Pew Internet and American Life Project*, on the Internet at **www.pewinternet.org/pdfs/PIP_Broadband_Report.pdf**.

[14] Tom Downes and Shane Greenstein, "Universal Access and Local Internet Markets in the U.S.," *Research Policy*, September 2002, pp. 1035–52.

[15] "Does the Residential Broadband Market Need Fixing?" (Congressional Budget Office, December 2003), on the Internet at **www.cbo.gov/ftpdocs/48xx/doc4868/12–03-Broadband.pdf**.

[16] Nestor M. Arguea, Cheng Hsiao, and Grant A. Taylor, "Estimating Consumer Preferences Using Market Data—an Application to U.S. Automobile Demand," *Journal of Applied Econometrics*, January–March 1994, pp. 1–18.

[17] Sherwin Rosen, "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition," *Journal of Political Economy*, January–February 1974, pp. 34–55.

[18] Jack Triplett, *Handbook on Hedonic Indexes and Quality Adjustments in Price Indexes: Special Application to Information Technology Products*, STI Working Paper 2004/09, Oct. 8, 2004, on the Internet at **www.oecd.org/**

dataoecd/37/31/33789552.pdf; see especially pp. 185–86.

[19] Maureen L. Cropper, Leland B. Deck, and Kenneth E. McConnell, "On the Choice of Functional Form for Hedonic Price Functions," *Review of Economics and Statistics*, November 1988), pp. 668–75.

[20] In cases where a transformation parameter equals 0, the logarithmic transformation is used instead of the usual Box-Cox transformation by that parameter.

[21] Triplett, *Handbook on Hedonic Indexes*, p. 182.

[22] Cropper, Deck, and McConnell, "On the Choice of Functional Form," p. 668.

[23] *Ibid.*, p. 671.

[24] Horrigan, "Home Broadband," p. ii.

[25] Yu and Prudhomme, "Econometric Issues in Hedonic Price Indices," also used these two criteria to help select functional form.

[26] John J. Spitzer, "A Fast and Efficient Algorithm for the Estimation of Parameters in Models with the Box-and-Cox Transformation," *Journal of the American Statistical Association*, December 1982, pp. 760–66; see especially p. 760.

[27] For more on these imputation methods, see BLS *Handbook of Methods*, June 2007, chapter 17, "The Consumer Price Index," on the Internet at **www.bls.gov/opub/hom/pdf/homch17.pdf**.

[28] Triplett, *Handbook on Hedonic Indexes*, p. 26.

[29] Paul R. Liegey and Nicole Shepler, "Adjusting VCR prices for quality change: a study using hedonic methods," *Monthly Labor Review*, September 1999, pp. 22–37; on the Internet at **www.bls.gov/opub/mlr/1999/09/art3full.pdf**.

[30] Craig Brown and Anya Stockburger, "Item replacement and quality change in apparel price indexes," *Monthly Labor Review*, December 2006, pp. 35–45; on the Internet at **www.bls.gov/opub/mlr/2006/12/art3full.pdf**.

[31] The formula for the discrepancy between the experimental (adjusted) index and the official (unadjusted) index is $(AdjustedIndex_t - UnadjustedIndex_{t-1})$. In comparison, David S. Johnson, Stephen B. Reed, and Kenneth J. Stewart, "Price measurement in the United States: a decade after the Boskin Report, *Monthly Labor Review*, May 2006, pp 10–19 (on the Internet at **www.bls.gov/opub/mlr/2006/05/art2full.pdf**), compared indexes with the formula $(AdjustedIndex_t - UnadjustedIndex_t)/UnadjustedIndex_t$.

[32] *Ibid.*