

Bias and Inefficiency in Quality-Adjusted Hedonic Regression Analysis

Auer, Ludwig von; Brennan, John E.

Postprint / Postprint

Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:

www.peerproject.eu

Empfohlene Zitierung / Suggested Citation:

Auer, L. v., & Brennan, J. E. (2006). Bias and Inefficiency in Quality-Adjusted Hedonic Regression Analysis. *Applied Economics*, 39(1), 95-107. <https://doi.org/10.1080/00036840500427841>

Nutzungsbedingungen:

Dieser Text wird unter dem "PEER Licence Agreement zur Verfügung" gestellt. Nähere Auskünfte zum PEER-Projekt finden Sie hier: <http://www.peerproject.eu> Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

gesis
Leibniz-Institut
für Sozialwissenschaften

Terms of use:

This document is made available under the "PEER Licence Agreement". For more information regarding the PEER-project see: <http://www.peerproject.eu> This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

Mitglied der

Leibniz-Gemeinschaft



Bias and Inefficiency in Quality-Adjusted Hedonic Regression Analysis

Journal:	<i>Applied Economics</i>
Manuscript ID:	APE-05-0166.R1
Journal Selection:	Applied Economics
Date Submitted by the Author:	29-Sep-2005
JEL Code:	C29 - Other < , C43 - Index Numbers and Aggregation < , L63 - Microelectronics Computers Communications Equipment <
Keywords:	hedonics, price index, measurement bias, technical progress, laser printer

powered by ScholarOne
Manuscript Central™

Bias and Inefficiency in Quality-Adjusted Hedonic Regression Analysis

March 2005

Abstract

Numerous quality-adjusted hedonic price-trend studies based on computer prices have provided support to widely held suspicions that officially released price indices are not accurately measuring the price declines occurring in many information technology (IT) products. If provable, then general price inflation is being overestimated and, consequently, real GDP is being underestimated. Existing evidence, however, is inconclusive. First, empirical findings for IT-products other than computers are essentially nonexistent and, secondly, estimation bias is inherent in the hedonic regression technique most commonly employed. This paper presents an unbiased method together with an estimated quality-adjusted price trend for laser printers (1993–2004).

Keywords: Hedonics, Price Index, Measurement Bias, Technical Progress,
Laser Printer

JEL Classification: C29; C43; L63

1 Introduction

For many years concerns have been articulated that the officially released price indices are not being correctly computed in numerous countries. Empirical evidence is mounting that those national statistical agencies that rely on conventional statistical methods are publishing price trends that indicate smaller price declines for products in the information and communication technology sector than do the corresponding price trends estimated by quality-adjusted hedonic regression techniques (see e.g., Triplett, 2004, p. 8). If substantiated, these findings indicate that these officially released price trends contain an upward bias. This would imply that overall price inflation is being overestimated and, consequently, real GDP growth is being underestimated. If this were the case, the consequences for trustworthy economic policy-making in these countries would be serious indeed.

As a result, the implementation of satisfactory procedures for the accurate measurement of long-term price trends is necessary. For scores of products, however, this precise assessment presents a tricky problem. In many of today's high-tech industries, for example, the most successful manufactures are those adept companies that are capable of continuously incorporating state-of-the-art technologies into their product lines. As a consequence, these suppliers deliver an uninterrupted stream of technologically enhanced product models to the marketplace. In spite of the fact that these new models repeatedly render their forerunners lackluster in comparison, they are, nevertheless, frequently sold at lower prices. For this reason, it has long been recognized that quality-adjusted price measurement techniques are a prerequisite when attempting to accurately evaluate the long-term price trends of such products.

Products belonging to the category identified as "computers" have unquestionably experienced this phenomenon. During the previous half-century, technical progress has transformed computer models from massive, obscure devices, affordable and programmable by only a handful of prominent organizations, into the convenient, ever-present personal helpers of today. Apart from for their name, the evermore reasonably priced, increasingly powerful, and smaller computer models of today bear slight resemblance to their slower, less capable, oversized, and exceedingly expensive ancestors of yesteryear.

Many other products have experienced a similar evolution over time. In the United States, for example, at the dawn of the twentieth century the ownership of an automobile was a luxury that only a privileged few could justify. Nowadays,

1
2
3 however, just about every household has at least one car and they consider it to
4 be a prerequisite for everyday life. Early in the 1930's, the automotive industry
5 was beginning to feel its economic power and government regulators in the U.S.
6 began reacting by expressing concern that the General Motors Corporation might
7 be exploiting its dominant market position by implementing monopolistic pricing
8 practices. In order to defend its domestic pricing policies, General Motors commis-
9 sioned T. C. Court to estimate a quality-adjusted price trend for automobiles over
10 the period 1920 to 1939. In the course of Court's (1939) pioneering investigation,
11 he originated an innovative research methodology that subsequently became known
12 as the Adjacent Year Regression (AYR) method, (Griliches, 1971). To the present
13 day, when data is available only in limited quantities, the AYR method remains the
14 most frequently used quality-adjusted hedonic price measurement technique.
15
16
17
18
19
20
21
22

23 The AYR method, however, due to its reliance on pooled data from overlapping
24 time periods, has some serious inherent theoretical shortcomings. It is straightfor-
25 ward to show that the AYR method produces either biased estimates or at best
26 inefficiently estimated coefficients. To the best of our knowledge, these obvious
27 shortcomings have never been adequately elucidated. This is particularly surprising,
28 since the estimation deficiencies present in the AYR method can easily be prevented
29 by a simple procedural modification called the NOAYR method, where the letters
30 "NO" indicate the "non-overlapping" feature of this improved regression technique.
31
32
33
34
35

36 Up till now, a considerable amount of research has been carried out in order
37 to estimate the quality-adjusted price trends of computers. In comparison, rela-
38 tively little empirical evidence exists regarding other IT products (e.g., computer
39 peripherals). If the published price trends for computers are taken as a proxy for
40 all of the products in the IT sector, it is highly likely that erroneous conclusions
41 could be drawn. This paper, therefore, concerns itself with an important IT prod-
42 uct that, heretofore, has received negligible attention. Utilizing both the AYR and
43 the NOAYR methods, a quality-adjusted long-term price trend for laser printers was
44 estimated. Certainly the product model enhancements that have occurred in laser
45 printers have had much less to do with advances in chip technology than have com-
46 puters. Therefore, the link between laser printer prices and technical chip-progress
47 should be less well defined and, as a result, a reduced amount of upward measure-
48 ment bias should be present in the published statistics. This study attempts to
49 discover whether or not this seemingly reasonable hypothesis can be substantiated
50 by actual price data.
51
52
53
54
55
56
57
58
59
60

1
2
3 This paper is organized into six sections. Section 2 presents the theoretical
4 framework for quality-adjusted hedonic price measurement. It begins by describing
5 the basic model. Next, the AYR method and its inherent theoretical shortcomings
6 are discussed. It is pointed out that a flexible variant of the AYR method removes
7 only part of the shortcomings. Subsequently, the NOAYR method is presented
8 as a solution to the estimation problems intrinsic to the AYR method. Section 3
9 introduces a simple approach for transforming the estimated coefficients into long-
10 term price trends. A description of the data set employed in this study is contained
11 in Section 4. The empirical specification used and the empirical results are presented
12 in Section 5. Therein, a long-term price trend for laser printers is provided using
13 the AYR and the NOAYR methods. These estimated results are compared with an
14 officially published price trend for laser printers. Section 6 contains the concluding
15 remarks.
16
17
18
19
20
21
22
23
24
25
26

27 **2 The Theoretical Framework for Quality-Adjusted** 28 **Hedonic Price Measurement**

29 **2.1 The Basic Model**

30
31
32 Assume that during some time period t ($t = 1, 2, \dots, T$), the various manufactures
33 of an IT product (e.g., laser printers) are offering N^t different models to the mar-
34 ketplace, each of these models identified by a subscript i ($i = 1, 2, \dots, N^t$), and they
35 are being sold for different prices, p_i^t . Waugh (1928), who was interested in studying
36 price differences in the Boston vegetable market, was the first to point out that
37 these diverse contemporaneous market prices could be formally related to variations
38 in product quality. In order to accomplish this, however, certain distinctive product
39 quality characteristics need to be identified. For modeling purposes these charac-
40 teristics must be measurable, time invariant, and appropriate for all of the various
41 product models considered.
42
43
44
45
46
47
48
49
50
51

52 Consider a trivariate random vector (P, i, t) from a mixed population consisting
53 of the random variables product price, P , the product model identifier, i , and the
54 time period marker, t . Furthermore, assume that the univariate conditional random
55 variable $(P|i, t)$ has a Conditional Expectation Function (CEF) that is linearly re-
56 lated to K product quality characteristics, q_{ki} ($k = 1, 2, \dots, K$). According to this
57
58
59
60

assumption, the CEF (also known as the Population Regression Function) is:

$$E(P|i, t) = \alpha_i^t + \sum_{k=1}^K \beta_{ki}^t q_{ki}, \quad t = 0, 1, 2, \dots, T \quad \text{and} \quad i = 1, 2, \dots, N^t.$$

The CEF is a single valued function that assigns to each product model a conditional mean price in each time period. The population intercept term, α_i^t , can be interpreted as capturing those factors in the marketplace that influence a product model's conditional mean price other than its own inherent product quality characteristics. The population slope coefficients, β_{ki}^t , on the other hand, can be interpreted as providing the implicit price contribution of each of the K product quality characteristics.

In order to enable sample estimation of the population coefficients in the CEF, the assumption is made that the intercept and slope coefficients remain constant for all of the models considered during time period t ($\alpha_i^t = \alpha^t$ and $\beta_{ki}^t = \beta_k^t$). According to this assumption the CEF is:

$$E(P|i, t) = \alpha^t + \sum_{k=1}^K \beta_k^t q_{ki}.$$

Utilizing cross-sectional data, the estimated Sample Regression Function for model i in time period t is:

$$\hat{p}_i^t = \hat{\alpha}^t + \sum_{k=1}^K \hat{\beta}_k^t q_{ki}, \quad t = 0, 1, 2, \dots, T. \quad (1)$$

Under the usual assumptions, unbiased estimates $\hat{\alpha}^t$ and $\hat{\beta}_k^t$, of the $K+1$ population coefficients are obtained by performing $T+1$ Ordinary Least Squares (OLS) regressions utilizing the observed prices, p_i^t , and product quality characteristics, q_{ki} . The estimate of the conditional mean price of model i in time period t , \hat{p}_i^t , obtained from the sample is the best minimum Mean Squared Error (MSE) prediction of the market price.

Those factors present in the marketplace that exert an influence upon the prices of all models are particularly prone to variations over time. For that reason, the value of the intercept term is likely to vary. In contrast, the population slope coefficients, β_k^t , relating the model's price to its product quality characteristics, could conceivably conform to three plausible alternative assumptions.

Assumption 1 *The population slope coefficients vary throughout the $T+1$ time periods considered.*

In this case, provided an adequate quantity of data is available, Sample Regression Function (1) should be estimated separately for each time period. Alternatively, estimation could be accomplished by pooling the data from all $T+1$ time periods and using a single Sample Regression Function containing time dummy variables, d_i^s , that are defined according to the time period marker s ($s = 1, 2, \dots, T$):

$$\hat{p}_i^t = \hat{\alpha}^t + \sum_{s=1}^T \hat{\gamma}^s d_i^s + \sum_{k=1}^K \hat{\beta}_k^t q_{ki} + \sum_{k=1}^K \sum_{s=1}^T \hat{\delta}_k^s d_i^s q_{ki}, \quad (2)$$

where

$$d_i^s = \begin{cases} 1 & \text{if model } i \text{ is observed in time period } s \\ 0 & \text{if model } i \text{ is not observed in time period } s. \end{cases}$$

In this case, the estimated coefficients $\hat{\delta}_k^s = (\hat{\beta}_k^t - \hat{\beta}_k^0)$ provide unbiased estimates of the differences between the values of the population slope coefficients in time period $t=s$ with those of the initial time period, $t=0$. Sample Regression Function (1) and this single Sample Regression Function (2) will both produce the same unbiased estimates of the population slope coefficients.

Assumption 2 *The population slope coefficients remain constant throughout the $T+1$ time periods considered.*

Under this assumption the CEF is:

$$E(P|i, t) = \alpha^t + \sum_{k=1}^K \beta_k q_{ki}.$$

In this case, unbiased estimates of the population coefficients can be obtained once again by pooling all of the data from the $T+1$ time periods and estimating a simplified version of single Sample Regression Function (2):

$$\hat{p}_i^t = \hat{\alpha}^t + \sum_{s=1}^T \hat{\gamma}^s d_i^s + \sum_{k=1}^K \hat{\beta}_k^t q_{ki}, \quad (3)$$

where d_i^s is defined as in Sample Regression Function (2).

Assumption 3 *The population slope coefficients vary over some of the time periods considered while remaining constant over the others.*

This assumption is simply a combination of Assumptions 1 and 2. In this case, estimation of the population coefficients over the complete $T+1$ time periods is accomplished by selectively applying either single Sample Regression Function (2) or (3). The selection of these regression functions is made depending upon whether the population slope coefficients are assumed to vary or to remain constant over the range of time periods under consideration.

2.2 The AYR Method

In order to estimate Sample Regression Function (2), a sufficient amount of cross-sectional data is required in each of the time periods considered. Unfortunately, an ample quantity of data is not always available. A number of studies have found a solution for this problem by using scanner data (e.g., see Silver, 1999 or Silver and Heravi, 2001). Scanner data, however, is often prohibitively expensive. Furthermore, when attempting to estimate longer-term price trends retrospectively, scanner data is seldom available. These considerations necessitate the search for alternative sources of publicly available data. Product catalogues and the product test reports published in trade magazines often provide the answer. Unfortunately, however, the problems associated with a limited amount of data more often than not remain.

The AYR method was proposed as a pragmatic “solution” to this data shortage problem. The method enables estimation by pooling observations from a selected number of adjacent time periods. Estimation is conducted under the assumption that the population intercept term varies from period to period. In contrast, however, the slope coefficients are assumed to remain constant over the entire pooled time span.

For expository purposes, consider the situation where the AYR method is accomplished using pooled data from simply two adjacent time periods, t and $t-1$ for $t \geq 1$. In order to identify each of the T possible regressions that could be estimated in this manner, an index τ is introduced that takes a value equal to t . The $(\tau = 1)^{\text{th}}$ regression is estimated by pooling data from time periods $t = 1$ and 0. The $(\tau = 2)^{\text{th}}$ regression uses the time periods $t = 2$ and 1, and in like manner, the $(\tau = T)^{\text{th}}$ regression pools data from the last available pair of time periods $t = T$ and $T-1$. A Time Dummy Variable (TDV), d_i^τ , is included in the τ^{th} regression in order to allow the intercept term to vary from period to period and, therefore, the two-period Sample Regression Function (hereafter referred to as a TDV regression) is:

$$\hat{p}_i^\tau = \hat{\alpha}^\tau + \hat{\gamma}^\tau d_i^\tau + \sum_{k=1}^K \hat{\beta}_k^\tau q_{ki}, \quad \tau = 1, 2, \dots, T, \quad (4)$$

where

$$d_i^\tau = \begin{cases} 1 & \text{if model } i \text{ is observed in time period } t \\ 0 & \text{if model } i \text{ is not observed in time period } t-1. \end{cases}$$

A salient feature of the AYR method is its reliance on these overlapping TDV regressions that are conducted using pooled data from two adjacent time periods.

The τ^{th} TDV regression is estimated under the assumption that the population slope coefficients in the two pooled time periods are equal, $\beta_k^t = \beta_k^{t-1}$ ($k = 1, 2, \dots, K$). If this assumption is correct, the estimated coefficients, $\hat{\beta}_k^\tau$, provide unbiased estimates of the true population values. The estimated intercept coefficient $\hat{\alpha}^\tau$ is an unbiased estimate of the population intercept term α^{t-1} , and, due to the presence of the time dummy variable, $(\hat{\alpha}^\tau + \hat{\gamma}^\tau)$ is an unbiased estimate of the population intercept term α^t .

This regularly used overlapping estimation procedure, however, has some rather serious inherent theoretical shortcomings. Court (1939) and subsequently other authors have referred to some of the AYR method's various problems. Largely ignored heretofore, however, is the fact that the AYR method produces either biased or at best inefficient estimation results. This can be demonstrated by considering the three alternative assumptions outlined hereinbefore in Section 2 (Assumptions 1, 2, and 3) pertaining to the population slope coefficients, β_k^t .

If Assumption 1 were true, the population slope coefficients would vary throughout the $T+1$ time periods considered: $\beta_k^t \neq \beta_k^{t-1}$ for $1 \leq t \leq T$. Unbiased estimates of the population coefficients of periods t and $t-1$ could only be obtained by utilizing a two-period version of the single Sample Regression Function (2),

$$\hat{p}_i^\tau = \hat{\alpha}^\tau + \hat{\gamma}^\tau d_i^\tau + \sum_{k=1}^K \hat{\beta}_k^\tau q_{ki} + \sum_{k=1}^K \hat{\delta}_k^\tau d_i^\tau q_{ki}, \quad \tau = 1, 2, \dots, T, \quad (5)$$

where d_i^τ is defined as in TDV regression (4). Therefore, the AYR method's TDV regression (4) is improperly specified. This is due to the well-known omission of relevant variables problem, specifically in this case, the K interaction terms, $d_i^\tau q_{ki}$. Consequently, every one of the estimated coefficients ($\hat{\alpha}^\tau$, $\hat{\gamma}^\tau$, and $\hat{\beta}_k^\tau$) from the T two-period pooled TDV regressions (4) is a biased estimate of its corresponding population coefficient, except for those coefficients relating to variables that are uncorrelated with the omitted variables $d_i^\tau q_{ki}$.

Result 1 *In the case of Assumption 1, the AYR method produces estimated coefficients, $\hat{\alpha}^\tau$, $\hat{\gamma}^\tau$, and $\hat{\beta}_k^\tau$, that are biased estimates of the corresponding population coefficients in all of the $T+1$ time periods.*

If Assumption 2 were true, the population slope coefficients would remain constant throughout the T time periods considered. The AYR method relies on overlapping TDV regressions that are conducted under the assumption that the population

slope coefficients are only equal over the two time periods pooled. In this case, efficient estimates of the population slope coefficients can be obtained by pooling all of the data from the $T+1$ time periods and estimating the single Sample Regression Function (3). Consequently, the AYR method's use of TDV regression (4) is an inefficient estimation procedure.

Result 2 *In the case of Assumption 2, the AYR method produces estimated coefficients, $\hat{\alpha}^\tau$, $\hat{\gamma}^\tau$, and $\hat{\beta}_k^\tau$, that are unbiased but inefficient estimates of the corresponding population coefficients in all of the $T+1$ time periods.*

If Assumption 3 were true, then the population slope coefficients would vary over some of the time periods considered while remaining constant over others. As indicated heretofore, the AYR method's use of TDV regression (4) produces either biased or inefficient estimates of the population slope coefficients. Throughout those adjacent time periods where the population slope coefficients are varying, the AYR method's overlapping TDV regressions would produce biased estimated coefficients. On the other hand, over those adjacent time periods where the population slope coefficients remain constant, the AYR method's estimated coefficients are unbiased, but they are inefficiently estimated.

Due to the AYR method's overlapping TDV regressions, two different sets of estimated slope coefficients are generated for each time period. Consider the case where the population slope coefficients have remained constant for several time periods prior to and including time period t , but then they change so that in time period $t+1$ they are different. In such a case, the τ^{th} TDV regression (covering periods t and $t-1$) would produce inefficient estimates of the population slope coefficients in time period t . The $(\tau+1)^{\text{th}}$ TDV regression (covering periods t and $t+1$), however, would produce coefficient estimates for time period t that are biased.

Result 3 *In the case of Assumption 3, the AYR method produces estimated coefficients, $\hat{\alpha}^\tau$, $\hat{\gamma}^\tau$, and $\hat{\beta}_k^\tau$, that are biased estimates of the corresponding population coefficients in those time periods where the population slope coefficients are varying, and inefficient estimates in those time periods where they remain constant for more than two periods.*

2.3 The Flexible AYR Method

To the best of our knowledge, the theoretical shortcomings inherent in the AYR method, even though they are rather obvious, have never been explicitly brought

out into the open or adequately elucidated. Even in published studies utilizing a more flexible AYR approach, supported by hypothesis testing relating to parameter stability, these deficiencies have not been adequately addressed.

A well-known study by Berndt and Rappaport (2001) provides a good case in point. Investigating the price trend of personal computers (both desktop and laptop models) over the time period from 1976 to 1999, these authors first tested each pair of adjacent time periods in order to determine whether or not the slope coefficients did in fact remain constant before estimating the two-period TDV regression (4). In order to accomplish this, they first estimated each of the T TDV regressions (5). For each of these regressions, F -Tests ($H_0: \delta_1^\tau = \delta_2^\tau = \dots = \delta_K^\tau = 0$) were conducted for the purpose of determining whether or not the K estimated coefficients, $\hat{\delta}_k^\tau$, were in fact close enough to zero to conclude that the population slope coefficients were actually constant over the two time periods considered, $\beta_k^t = \beta_k^{t-1}$. If this proved to be the case, then TDV regression (4) was used for estimation. If not, the two time periods were estimated separately using Regression Function (1) that produces the same estimates of the population coefficients as TDV regression (5). When utilizing Regression Function (1), provided that sufficient data is available, a separate set of K slope coefficients are estimated for each of the time periods, $\hat{\beta}_k^t$.

Berndt and Rappaport used TDV regression (5) on pooled data from the time periods 1989-1990 and 1990-1991. In both of these regressions only laptops were considered and the null hypothesis could not be rejected. Therefore, the population slope coefficients were deemed to be equal over the time periods pooled: $\beta_k^{1989} = \beta_k^{1990}$ and $\beta_k^{1990} = \beta_k^{1991}$. As a consequence, they estimated two consecutive TDV regressions, $\tau = 1990$ and $\tau = 1991$, using TDV regression (4). If in fact it is true that $\beta_k^{1989} = \beta_k^{1990}$ and $\beta_k^{1990} = \beta_k^{1991}$, then necessarily it is true that $\beta_k^{1989} = \beta_k^{1990} = \beta_k^{1991}$.

By conducting an F -Test on a three period variant of TDV regression (5), one could statistically ascertain whether this latter equality is likely to be true. If equality could be (correctly) accepted, then estimating two separate TDV regressions instead of a single TDV regression covering the complete time span 1989-1991 would be inefficient. On the other hand, if equality could be (correctly) rejected, then at least one of the two TDV regressions performed by Berndt and Rappaport would suffer from omitted-variable bias. In other words, the Flexible AYR method does not resolve the inefficient estimation problem, but it avoids some of the situations where estimation bias could occur. Inefficiency, however, can be avoided by employ-

ing a minor procedural modification. This modification produces a procedure called the NOAYR method.

2.4 The NOAYR Method

Consider a population where Assumption 3 holds true. Under this set of circumstances, the population slope coefficients would vary over some of the time periods considered while they would remain constant over the others. The estimation problems inherent in the AYR and Flexible AYR methods arise from the use of the overlapping time periods in the TDV regressions that are conducted. By simply avoiding these time period overlaps these problems can be eliminated. The improved method presented here is called the NOAYR method (“NO” for “non-overlapping”).

Estimation utilizing the NOAYR method is accomplished in the following manner. Consider the simple case where only one product quality characteristic, q_i , is considered and thus the two-period TDV regression (4) simplifies to:

$$\hat{p}_i^\tau = \hat{\alpha}^\tau + \hat{\gamma}^\tau d_i^\tau + \hat{\beta}^\tau q_i, \quad \tau = 1, 2, \dots, T. \quad (6)$$

Furthermore, assume that the population slope coefficients vary only between time periods $t = t'-1$ and $t = t'$, such that $\beta^{t'} > \beta^{t'-1}$ and $1 < t' < T-1$. The population slope coefficients remain constant in every other time period considered. The time-series of the population slope coefficient values is indicated in Figure 1.

Here Figure 1:

The AYR, Flexible AYR, and NOAYR Methods.

In Figure 1, the brackets indicate the time periods pooled in the respective TDV regressions. For example, if estimation were accomplished utilizing the AYR method, a sequence of overlapping pooled time periods would be used. Using the TDV regression (6), the $(\tau = t')$ th TDV regression would pool data from the periods $t = t'$ and $t'-1$. As heretofore stated, however, this TDV regression would produce biased estimates of the population coefficients.

If testing for parameter stability correctly indicates that $\beta^{t'-1} \neq \beta^{t'}$, the Flexible AYR method would avoid this bias because it would not estimate the $(\tau = t')$ th TDV regression (see Figure 1). This method, however, does not incorporate the information that the population slope coefficients remain constant over the time

span $t = 0$ to $t'-1$ and from $t = t'$ to T . As a consequence, the Flexible AYR method is not efficient.

In order to remove this inefficiency in this simple case, the NOAYR method estimates only two TDV regressions. The first TDV regression is estimated using pooled data from all of the time periods between $t = 0$ and $t'-1$. The second TDV regression pools data from the time periods between $t = t'$ and T (see Figure 1). Unbiased, as well as efficient, estimates of the population coefficients are obtained for all of the $T+1$ time periods by simply performing these two TDV regressions. The modification necessary to turn the AYR method into the NOAYR method is so trivial that one might rightly ask, Why is the AYR methodology of pair wise overlapping TDV regression still so popular? A possible answer is provided in Section 3.

For all practical purposes, the appropriate guidelines for applying the NOAYR method are as follows. First one should test whether the slope coefficients' values change over time. If these tests indicate that the slope coefficients remain constant over two or more adjacent years, then all of these years should be pooled into a single TDV regression using additional time dummy variables. In those years where tests indicate that the coefficients change, separate single year regressions utilizing Sample Regression Function (1) should be estimated, provided sufficient data is available.

If no period exists in which the slope coefficients remain unchanged (as in Assumption 1), then all of the estimated coefficients of the AYR method's overlapping TDV regressions are biased. In this case, however, the NOAYR method unfortunately does not provide a solution either. The NOAYR method also relies on some (although fewer) TDV regressions and therefore upon the assumption of constant slope coefficients. If at least some periods exist in which the slope coefficients remain constant (as in Assumption 3), then the NOAYR method could represent a useful and reliable method for long-term price measurement. In contrast, the AYR method with its overlapping TDV regressions produces biased estimated coefficients and therefore should be avoided.

3 Estimation of Long-Term Price Trends

An estimate of the individual price relative of model i in time period t , \hat{r}_i^t , is defined as the ratio of the estimated conditional mean price in time period t over its value

in time period $t-1$:

$$\hat{r}_i^t = \hat{p}_i^t / \hat{p}_i^{t-1} \quad , \quad t = 1, 2, \dots, T \quad \text{and} \quad i = 1, 2, \dots, N^t .$$

More often than not, however, an estimate of a comprehensive price relative, \hat{r}^t , is desired. This estimate, capable of estimating the price change for a whole product category in general (e.g., laser printers), can be derived in a variety of ways. Usually, however, this price relative is calculated as some function of the individual price relatives of the N^t models,

$$\hat{r}^t = h(\hat{r}_1^t, \hat{r}_2^t, \dots, \hat{r}_{N^t}^t) .$$

When a semi-logarithmic specification of the two-period TDV regression (4) is used, a particularly useful expression results. In this case, the dependent variable becomes: $\ln \hat{p}_i^t$. Consequently, the estimated coefficient on the time dummy variable equals,

$$\hat{\gamma}^\tau = \ln \hat{p}_i^t - \ln \hat{p}_i^{t-1} = \ln \hat{r}^t ,$$

and therefore

$$\exp(\hat{\gamma}^\tau) = \hat{r}^t . \tag{7}$$

According to this commonly used specification, all models now exhibit exactly the same price relative, \hat{r}^t . It has been pointed out in various studies (e.g., Goldberger, 1968), however, that least squares estimation of this logarithmic specification leads to bias in the conditional mean prices, \hat{p}_i^t . Fortunately, in hedonic regression analysis this bias is usually very small and therefore customarily neglected. This present study is no exception. Dalen and Bode (2004) provide a careful discussion of this type of bias.

Equation (7) implies that in the AYR method the estimated comprehensive price relatives for all time periods, \hat{r}^t , can easily be obtained directly from the estimated coefficient on the time dummy variable, $\hat{\gamma}^\tau$. The simplicity of this approach is one plausible explanation for the AYR method's widespread popularity. When the Flexible AYR or the NOAYR method is employed, however, there could be pairs of adjacent time periods where no TDV regressions are conducted. Consequently, there will be no estimated time dummy variable coefficients, $\hat{\gamma}^\tau$, for these time periods. In this case, the estimation of the price relatives, \hat{r}^t , will require some alternative approach. The following method was used in the empirical section of this paper.

Suppose that during the time periods $t-2$ and $t-1$, the population slope coefficients remain constant but they change between time periods $t-1$ and t . They

change again between the time periods t and $t+1$. Consequently, a TDV regression covering time periods $t-2$ and $t-1$ can be conducted accompanied by a separate regression for time period t . In this case, however, how does one obtain the estimated price relative, \hat{r}^t ?

Assume that during time periods t and $t-1$, price data has been collected for J different product models ($j = 1, 2, \dots, J$) and that some of these models have only been observed in one of these two time periods. Furthermore, suppose that a semi-logarithmic specification was used. From the TDV regression covering the time periods $t-1$ and $t-2$, the $(\tau-1)^{\text{th}}$ TDV regression, an estimate of the conditional mean price, \hat{p}_j^{t-1} , can be obtained for each of the J product models:

$$\hat{p}_j^{t-1} = \exp \left(\hat{\alpha}^{\tau-1} + \hat{\gamma}^{\tau-1} + \sum_{k=1}^K \hat{\beta}_k^{\tau-1} q_{kj} \right). \quad (8)$$

In a similar manner, from the separate regression covering time period t , an estimated conditional mean price, \hat{p}_j^t , can be obtained in like manner for each of the J product models:

$$\hat{p}_j^t = \exp \left(\hat{\alpha}^t + \sum_{k=1}^K \hat{\beta}_k^t q_{kj} \right). \quad (9)$$

The ratio of these two estimated conditional mean prices, $\hat{r}_j^t = \hat{p}_j^t / \hat{p}_j^{t-1}$, is the individual price relative for product model j in time period t . For the calculation of the comprehensive price relative, \hat{r}^t , a Jevons index can be used:

$$\hat{r}^t = \exp \left[(1/J) \sum_{j=1}^J \ln \hat{r}_j^t \right]. \quad (10)$$

Difficulties with this approach only arise when the range of product models observed in time period t differs strongly from those observed in time period $t-1$. In this case, a matched model Jevons index could be computed. A matched model Jevons index is one where only those product models that have observed price data in both time periods are considered in the calculation.

In the empirical section of this paper, the approach utilizing Equations (8) to (10) was applied. There are, however, various other conceivable methods. For example, Berndt and Rappaport (2001) follow an alternative approach that was proposed by Dhrymes (1971) and further developed by Cowling and Cubbin (1972).

4 Data Description

Currently, even the introductory macroeconomic textbooks (e.g., Blanchard, 2003, p. 256) are expressing misgivings concerning the accuracy of the officially published inflation statistics issued by national statistical agencies. These suspicions are in accord with the published empirical results from numerous hedonic regression studies that have been conducted using data on computer prices. Consequently, evidence is mounting that the officially released price trends tend to underestimate the true quality-adjusted ones. A fair and comprehensive assessment of this suspicion is to be found in ILO *et al.* (2004).

Apart from computers, however, evidence of price trend measurement bias for other products in the IT sector is for the most part lacking, Cole *et al.* (1986) being a notable exception. For example, there are no comparable studies using price data for computer monitors, printers, or for any other such related product. This omission is somewhat surprising since computer peripherals represent a sizable part of the total expenditure on IT products. In an effort to alleviate this oversight, this study provides a look at the quality-adjusted long-term price trend of laser printers in the Federal Republic of Germany from 1993 to 2004. The estimated price trend obtained has been compared with the officially published one released by the *Statistisches Bundesamt* (Federal Statistical Office of Germany). The *Statistisches Bundesamt* has computed and published a separate quality-adjusted price index for laser printers only since 2000. Prior to 2000 the published price trend for ink jet printers has been used as proxy for these missing values and prior to 1995 computer prices have been used for this purpose.

The information utilized in this study was obtained from a sequence of test reports that were published in a bi-weekly German language computer trade magazine entitled: *c't - magazin für computertechnik*. Many experts regard these test results to be among the most reliable and sophisticated in the field. For that reason, test results appearing in other computer magazines were deliberately barred from consideration in the study. This policy ensured an unusually coherent data set of 385 observations. Twelve time-invariant qualitative characteristics were utilized, q_{ki} ($k = 1$ to 12). Six of these were measured as continuous variables, while dichotomous variables represented the others. The continuous variables measure such characteristics as: printing speed, standard memory size, added memory size, maintenance cost per page, paper capacity, and printing resolution. The dichotomous variables repre-

1
2
3 sented the features: equipped with postscript, upgraded with postscript, equipped
4 with network connectivity, upgraded with network connectivity, equipped with extra
5 interface (provided that the printer has no network connectivity), and GDI-printer
6 (Graphical Device Interface). A complete description of the data set used in this
7 study, in conjunction with the data collection process, can be obtained on request
8 from the authors.
9

10
11
12 Some selected summary statistics are presented in Table 1. The second column
13 shows the number of observations per year. The third column gives the average
14 price in EUR for the printers. The table demonstrates that in the years 1996 and
15 1997, primarily “low quality” printers were tested, whereas in the years 1998 and
16 1999, many of the “top-of-the-line” models received consideration. By and large,
17 a clear downward trend in model prices was observed. The last column displays
18 the average printing speed measured in pages per minute. These numbers vividly
19 indicate that a significant amount of technical progress has been occurring during
20 this time period.
21
22

23
24
25
26
27
28
29
30 *Here Table 1:*

31
32 *Summary Statistics from the Laser Printer Data.*
33

34 5 Empirical Specification and Results

35
36 Before estimation began, certain diagnostic tests were conducted. These tests
37 suggested a semi-logarithmic functional form for the Sample Regression Function.
38 Based on this functional form, two different long-term price trends were computed.
39 One is based on the AYR method and the other on the NOAYR method.
40
41

42
43 As described in Section 2.2, the applied AYR method was based upon OLS esti-
44 mation of a sequence of overlapping TDV regressions, each of which covers exactly
45 two consecutive years. The estimated coefficient standard errors were corrected for
46 heteroscedasticity.
47
48

49
50 The TDV regression that was employed is:
51

$$52 \ln \hat{p}_i^\tau = \hat{\alpha}^\tau + \hat{\gamma}^\tau d_i^\tau + \sum_{k=1}^K \hat{\beta}_k^\tau q_{ki}, \quad \tau = 1994, 1995, \dots, 2004. \quad (11)$$

53
54 where

$$55 d_i^\tau = \begin{cases} 1 & \text{if model } i \text{ is observed in time period } t \\ 0 & \text{if model } i \text{ is not observed in time period } t. \end{cases}$$

The dependent variable in TDV regression (11), $\ln \hat{p}_i^\tau$, is in a natural logarithmic form. As a consequence, the estimated slope coefficients on the product quality variables, $\hat{\beta}_k^\tau$, are no longer indicating the implicit price contribution of these respective characteristics. Instead, they are estimating the semi-elasticities of each printer's conditional mean price with respect to these qualitative characteristics.

The AYR method, based on TDV regression (11), directly produces a time series of estimated coefficient values for each time period, $\hat{\gamma}^\tau$. Utilizing Equation (7), these coefficients can be used to produce the required price relatives, \hat{r}^t . Therefore, an AYR Price Index, I^t , for use in measuring the quality-adjusted long-term price trend can be defined in terms of these price relatives \hat{r}^t . The value of this price index in time period t (define $I^{1993} = 100$) is:

$$I^t = 100 \cdot \prod_{s=1994}^t \hat{r}^s, \quad t = 1994, 1995, \dots, 2004.$$

The estimated AYR quality-adjusted long-term price trend for laser printers is depicted in Figure 2. According to this price trend, the prices of laser printers fell by 76.6 percent.

Here Figure 2:

Comparison of Laser Printer Price Trends.

The calculation of the AYR price trend was based upon estimates obtained from the TDV regressions that were conducted. Each of these TDV regressions utilized pooled data from two consecutive years. Hypothesis testing was conducted under the null hypothesis that the slope coefficients were equal over the two-pooled years. These tests revealed that only in the case of the $(\tau = 2000)^{\text{th}}$ and the $(\tau = 2004)^{\text{th}}$ TDV regression could the null hypotheses be accepted at a 5 percent significance level. In fact, for all other pairs of adjacent years the null hypothesis that the slope coefficients were equal was rejected at a 1 percent significance level. For that reason, a NOAYR price trend was computed utilizing the NOAYR method outlined in Section 2.4 and Section 3. In particular, TDV regressions were conducted using pooled data from the years 1999-2000 and 2003-04. Separate regressions were conducted for the years 1993, 1994, 1995, 1996, 1997, 1998, 2001, and 2002. The resulting price trend is also depicted in Figure 2. There is a divergence from the AYR price trend, but the deviation is not large.

1
2
3 The similarity of these two price trends should not be misconstrued as an excuse
4 for utilizing the AYR method in situations where the NOAYR method is feasible. In
5 such situations the AYR method produces biased (or at best inefficient) estimates
6 and these shortcomings should be avoided by applying the NOAYR method. An
7 interesting conclusion can be drawn, if one compares the AYR and the NOAYR price
8 trends of the years 1993 to 1998. For the computation of the NOAYR price trend for
9 each of these years separate regressions were conducted and from those the NOAYR
10 price trend was calculated. The results were similar to those obtained from the
11 overlapping TDV regressions conducted for computing the AYR price trend. This
12 result lends support to the conjecture that in situations where a lack of data renders
13 separate regressions for each time period impossible and where the population slope
14 coefficients vary throughout the time periods considered (Assumption 1 in Section
15 2.1), the measurement bias arising from the utilization of the AYR method might
16 not be very large.
17

18
19 The official price trend published by the *Statistisches Bundesamt* is also dis-
20 played in Figure 2. Using the two quality-adjusted price trends (AYR and NOAYR)
21 as a point of reference, the official price trend does not appear to have reacted to
22 the considerable price decline that occurred between 1993 and 2004. The official
23 price trend falls by 52.7 percent over the entire time period considered. The quality-
24 adjusted hedonic price trends measure a price decline of more than 76 percent. This
25 comparison provides some additional evidence for the (heretofore unconfirmed) sus-
26 picion that not only computers are experiencing impressive downward price trends
27 but also other IT products as well. Furthermore, it underlines the fact that fre-
28 quently the officially published estimates are not accurately measuring these price
29 trend declines.
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46

47 **6 Conclusions**

48
49 For those public agencies responsible for the accurate measurement of price trend
50 developments, adjusting for quality changes in many of today's technologically ad-
51 vanced products poses a serious challenge. In many cases, hedonic price measure-
52 ment techniques do offer a solution to this problem and can produce the quality-
53 adjusted price trend measurements required. This type of analysis, however, is labor
54 intensive and can frequently absorb a considerable amount of the agency's scarce
55 resources. For that reason, hedonic price measurement should be limited to those
56
57
58
59
60

1
2
3 selected product groups that are particularly prone to this type of measurement
4 bias. In order to identify these product categories, long-term price trends can be
5 estimated and compared to the officially reported price trends. For this purpose,
6 the AYR method has been the hedonic measurement technique most often utilized.
7 The AYR method is essentially a sequence of overlapping TDV regressions. This
8 paper demonstrates, however, that these overlapping regressions should be avoided.
9 A minor modification to the AYR method is introduced that avoids the theoretical
10 shortcomings inherent in the AYR method. This modification to the AYR method
11 was called the NOAYR method.
12
13
14
15
16
17

18 In the empirical section of this paper, both the AYR and the NOAYR method
19 were applied to data derived from publicly available test reports concerning laser
20 printers. These reports were published in a respected bi-weekly German computer
21 trade magazine. For each of these two hedonic regression methods, the price trend
22 of laser printers was computed for the time span 1993 to 2004. According to these
23 quality-adjusted price trends, the market prices of laser printers in Germany declined
24 by more than 76 percent. Over this identical time period, however, the official
25 price trend reported by the *Statistisches Bundesamt* (Federal Statistical Office of
26 Germany) declined by only 52.7 percent providing evidence of upward bias in their
27 price trend measurement.
28
29
30
31
32
33
34

35 Deviations between hedonic price measurement and the published official price
36 trends have been repeatedly reported in connection with computers. In response,
37 various national statistical agencies have introduced hedonic techniques for meas-
38 uring the price change of computers. The earliest example is the U.S. Bureau of
39 Economic Analysis (see Cole *et al.*, 1986). Many other national statistical agen-
40 cies have followed their lead. Recently, Linz and Eckert (2002) reported that the
41 *Statistisches Bundesamt* has also taken serious steps towards a routinely computed
42 hedonic price index for computers. The results presented in this study, however, in-
43 dicate that it is prudent for national statistical agencies to also consider computing
44 a hedonic price index for laser printers. Furthermore, additional studies need to be
45 conducted in order to get a better picture concerning other major IT products.
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

References

- Berndt, Ernst R. and Neal J. Rappaport (2001) Price and Quality of Desktop and Mobile Personal Computers: A Quarter-Century Historical Overview, *American Economic Review*, 91(2), 268-73.
- Blanchard, Olivier (2003) *Macroeconomics*, Prentice Hall, Upper Saddle River (New Jersey).
- Cole, Rosanne, Y. C. Chen, Joan A. Barquin-Stolleman, Ellen Dulberger, Nurhan Helvacian, and James H. Hodge (1986) Quality Adjusted Price Indexes for Computer Processors and Selected Peripheral Equipment, *Survey of Current Business*, 66, January, 41-50.
- Court, Andrew T. (1939) Hedonic Price Indexes with Automotive Examples, in *The Dynamics of Automobile Demand*, General Motors Corporation, New York, 99-117.
- Cowling, Keith and John Cubbin (1972) Hedonic Price Indexes for United Kingdom Cars, *Economic Journal*, 82, 963-78.
- Dalen, Jan van, Ben Bode (2004) *Estimation Biases in Quality-Adjusted Hedonic Price Indices*, paper presented at the SSHRC-Conference, Vancouver 2004.
- Dhrymes, Phoebus J. (1971) Price and Quality Changes in Consumer Capital Goods: Analysis of Quality Change, in *Price Indexes and Quality Change* (Ed.) Z. Griliches, Harvard University Press, Cambridge (Mass.), 88-149.
- Goldberger, Arthur S. (1968) The Interpretation and Estimation of Cobb-Douglas Functions, *Econometrica*, 36(3/4), 464-72.
- Griliches, Zvi (1971) Hedonic Price Indexes for Automobiles: An Econometric Analysis of Quality Change, in *Price Indexes and Quality Change*, (Ed.) Z. Griliches, Harvard University Press, Cambridge (Mass.), 55-87.
- ILO, IMF, OECD, UNECE, Eurostat, The World Bank (2004) *Consumer Price Index Manual: Theory and Practice*, International Labour Office, Geneva.
- Linz, Stefan and Gudrun Eckert (2002) Zur Einführung hedonischer Methoden in die Preisstatistik, *Wirtschaft und Statistik*, 10, 857-63.

1
2
3 Silver, Mick (1999) An Evaluation of the Use of Hedonic Regressions for Basic
4 Components of Consumer Price Indices, *Review of Income and Wealth*, 45(1),
5 41-56.
6
7

8
9 Silver, Mick, Saeed Heravi (2001) Scanner Data and the Measurement of Inflation,
10 *Economic Journal*, 111, 384-405.
11

12
13 Triplett, Jack E. (2004) *Handbook on Hedonic Indexes and Quality Adjustments in*
14 *Price Indexes: Special Application to Information Technology Products*, STI
15 Working Paper 2004/9, OECD, Paris.
16
17

18
19 Waugh, Frederick V. (1928) Quality Factors Influencing Vegetable Prices, *Journal*
20 *of Farm Economics*, 10(2), 185-96.
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

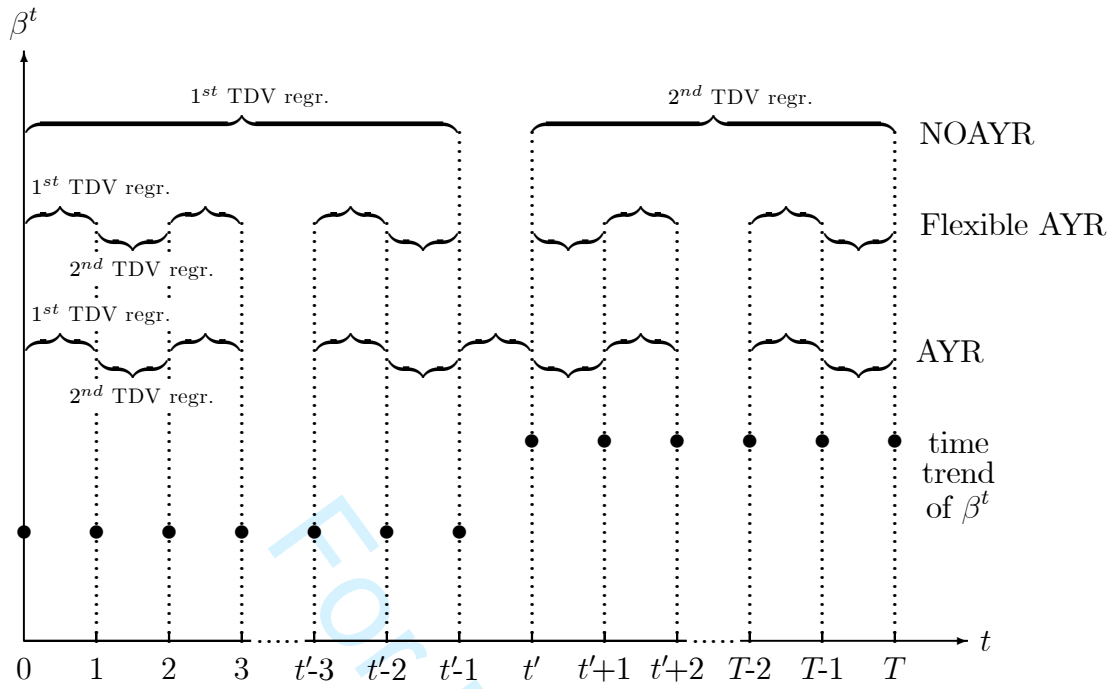


Figure 1: The AYR, Flexible AYR, and NOAYR Methods.

Table 1: Summary Statistics from the Laser Printer Data.

Year	Observations	Price	Speed
1993	32	1072	5.4
1994	21	1009	6.0
1995	22	1003	6.4
1996	28	584	6.4
1997	18	504	6.1
1998	26	861	10.2
1999	66	768	11.1
2000	52	590	10.9
2001	53	416	10.6
2002	25	336	11.8
2003	22	286	14.8
2004	20	253	15.9
Total	385		

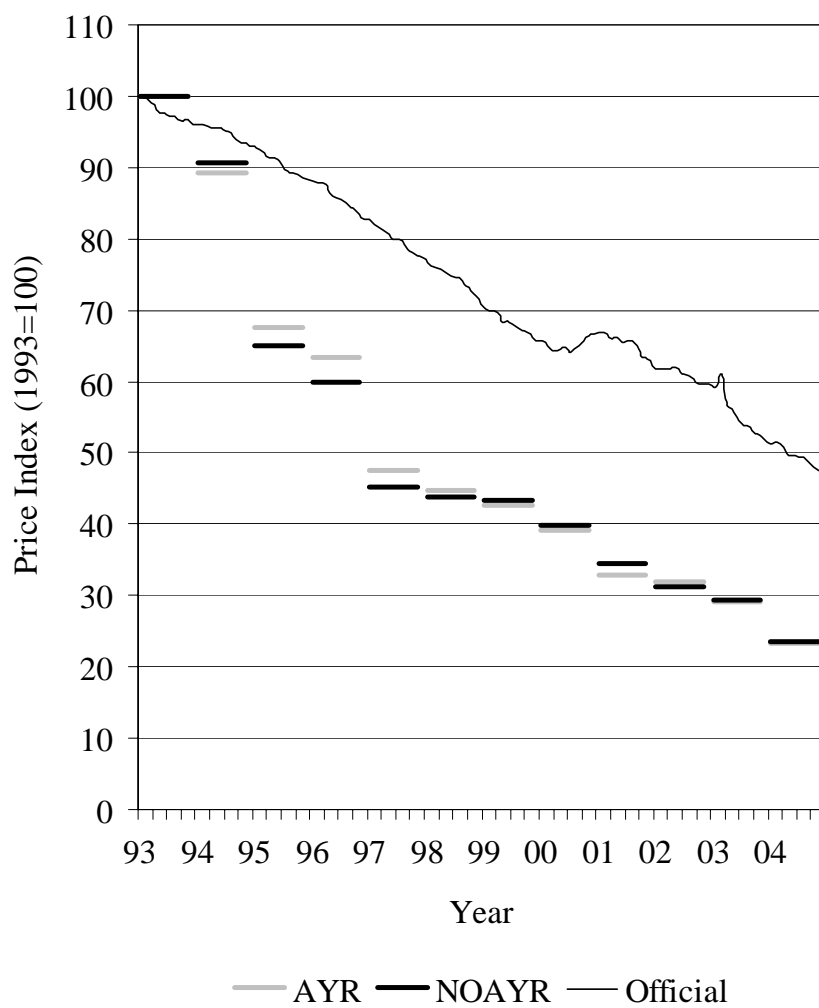


Figure 2: Comparison of Laser Printer Price Trends.



Bias and Inefficiency in Quality-Adjusted Hedonic Regression Analysis

by

Ludwig von Auer and John E. Brennan

Faculty of Economics and Management

Otto-von-Guericke-Universität Magdeburg, Germany

October 2005

Abstract

Numerous quality-adjusted hedonic price-trend studies based on computer prices have provided support to widely held suspicions that officially released price indices are not accurately measuring the price declines occurring in many information technology (IT) products. If verifiable, then general price inflation is being overestimated and, consequently, real GDP is being underestimated. Existing evidence, however, is inconclusive. First, empirical findings for IT-products other than computers are extremely rare and, secondly, estimation bias is inherent in the hedonic regression technique most commonly employed. This paper presents an unbiased method together with an estimated quality-adjusted price trend for laser printers (1993–2004).

Keywords: Hedonics, Price Index, Measurement Bias, Technical Progress,
Laser Printer

JEL Classification: C29; C43; L63

1 Introduction

For many years concerns have been articulated that the officially released price indices are not being correctly computed in numerous countries. Empirical evidence is mounting that those national statistical agencies that rely on conventional statistical methods are publishing price trends that indicate smaller price declines for products in the information and communication technology sector than do the corresponding price trends estimated by quality-adjusted hedonic regression techniques (see e.g., Triplett, 2004, p. 8). If substantiated, these findings indicate that these officially released price trends contain an upward bias. This would imply that overall price inflation is being overestimated and, consequently, real GDP growth is being underestimated. If this were the case, the consequences for trustworthy economic policy-making in these countries would be serious indeed.

As a result, the implementation of satisfactory procedures for the accurate measurement of long-term price trends is necessary. For scores of products, however, this precise assessment presents a tricky problem. In many of today's high-tech industries, for example, the most successful manufactures are those adept companies that are capable of continuously incorporating state-of-the-art technologies into their product lines. As a consequence, these suppliers deliver an uninterrupted stream of technologically enhanced product models to the marketplace. In spite of the fact that these new models repeatedly render their forerunners lackluster in comparison, they are, nevertheless, frequently sold at lower prices. For this reason, it has long been recognized that quality-adjusted price measurement techniques are a prerequisite when attempting to accurately evaluate the long-term price trends of such products.

Products belonging to the category identified as "computers" have unquestionably experienced this phenomenon. During the previous half-century, technical progress has transformed computer models from massive, obscure devices, affordable and programmable by only a handful of prominent organizations, into the convenient, ever-present personal helpers of today. Apart from for their name, the evermore reasonably priced, increasingly powerful, and smaller computer models of today bear slight resemblance to their slower, less capable, oversized, and exceedingly expensive ancestors of yesteryear.

Many other products have experienced a similar evolution over time. In the United States, for example, at the dawn of the twentieth century the ownership of an automobile was a luxury that only a privileged few could justify. Nowadays,

1
2
3 however, just about every household has at least one car and they consider it to
4 be a prerequisite for everyday life. Early in the 1930's, the automotive industry
5 was beginning to feel its economic power and government regulators in the U.S.
6 began reacting by expressing concern that the General Motors Corporation might
7 be exploiting its dominant market position by implementing monopolistic pricing
8 practices. In order to defend its domestic pricing policies, General Motors commis-
9 sioned T. C. Court to estimate a quality-adjusted price trend for automobiles over
10 the period 1920 to 1939. In the course of Court's (1939) pioneering investigation,
11 he originated an innovative research methodology that subsequently became known
12 as the Adjacent Year Regression (AYR) method, (Griliches, 1971). To the present
13 day, when data is available only in limited quantities, the AYR method remains the
14 most frequently used quality-adjusted hedonic price measurement technique.
15
16
17
18
19
20
21
22

23 The AYR method, however, due to its reliance on pooled data from overlapping
24 time periods, has some serious inherent theoretical shortcomings. It is straightfor-
25 ward to show that the AYR method produces either biased estimates or at best
26 inefficiently estimated coefficients. To the best of our knowledge, these obvious
27 shortcomings have never been adequately elucidated. This is particularly surprising,
28 since the estimation deficiencies present in the AYR method can easily be prevented
29 by a simple procedural modification called the NOAYR method, where the letters
30 "NO" indicate the "non-overlapping" feature of this improved regression technique.
31
32
33
34
35

36 Up till now, a considerable amount of research has been carried out in order to
37 estimate the quality-adjusted price trends of computers. In comparison, the amount
38 of empirical evidence that is presently available concerning other IT products (e.g.,
39 computer peripherals) is rather limited. If the published price trends for computers
40 are taken as a proxy for all of the products in the IT sector, it is highly likely that
41 erroneous conclusions could be drawn. This paper, therefore, concerns itself with an
42 important IT product that, heretofore, has received inadequate attention. Utilizing
43 both the AYR and the NOAYR methods, a quality-adjusted long-term price trend
44 for laser printers was estimated. Certainly the product model enhancements that
45 have occurred in laser printers have had much less to do with advances in chip
46 technology than have computers. Therefore, the link between laser printer prices
47 and technical chip-progress should be less well defined and, as a result, a reduced
48 amount of upward measurement bias should be present in the published statistics.
49 This study attempts to discover whether or not this seemingly reasonable hypothesis
50 can be substantiated by actual price data.
51
52
53
54
55
56
57
58
59
60

1
2
3 This paper is organized into six sections. Section 2 presents the theoretical
4 framework for quality-adjusted hedonic price measurement. It begins by describing
5 the basic model. Next, the AYR method and its inherent theoretical shortcomings
6 are discussed. It is pointed out that a flexible variant of the AYR method removes
7 only part of these shortcomings. Subsequently, the NOAYR method is presented as
8 a solution to the estimation problems intrinsic to the AYR method and reference is
9 made to some related literature. Section 3 outlines the computation of estimated
10 price relatives and how these price relatives can be used for the estimation of long-
11 term price trends. A description of the data set employed in this study is contained in
12 Section 4. The empirical specification used and the estimation results are presented
13 in Section 5. Therein, a long-term price trend for laser printers is provided using
14 the AYR and the NOAYR methods. These estimated results are compared with
15 an officially published price trend for laser printers. Section 6 contains concluding
16 remarks.
17
18
19
20
21
22
23
24
25
26
27

28 **2 The Theoretical Framework for Quality-Adjusted** 29 **Hedonic Price Measurement**

30 **2.1 The Basic Model**

31
32
33
34
35
36
37 Assume that during some time period t ($t = 0, 1, 2, \dots, T$), the various manufac-
38 tures of an IT product (e.g., laser printers) are offering N^t different models to the
39 marketplace, each of these models identified by a subscript i ($i = 1, 2, \dots, N^t$), and
40 they are being sold for different prices, p_i^t . Waugh (1928), who was interested in
41 studying price differences in the Boston vegetable market, was the first to point out
42 that these diverse contemporaneous market prices could be formally related to vari-
43 ations in product quality. In order to accomplish this, however, certain distinctive
44 product quality characteristics need to be identified. For modeling purposes these
45 characteristics must be measurable, time invariant, and appropriate for all of the
46 various product models considered.
47
48
49
50
51
52

53
54 Consider a trivariate random vector (P, i, t) from a mixed population consisting
55 of the random variables product price, P , the product model identifier, i , and the
56 time period marker, t . Furthermore, assume that the univariate conditional random
57 variable $(P|i, t)$ has a Conditional Expectation Function (CEF) that is linearly re-
58 lated to K product quality characteristics, q_{ki} ($k = 1, 2, \dots, K$). According to this
59
60

assumption, the CEF is:

$$E(P|i, t) = \alpha_i^t + \sum_{k=1}^K \beta_{ki}^t q_{ki}, \quad t = 0, 1, 2, \dots, T \quad \text{and} \quad i = 1, 2, \dots, N^t.$$

The CEF is a single valued function that assigns to each product model a conditional mean price in each time period. The population intercept term, α_i^t , can be interpreted as capturing those factors in the marketplace that influence a product model's conditional mean price other than its own inherent product quality characteristics. The population slope coefficients, β_{ki}^t , on the other hand, can be interpreted as providing the implicit price contribution of each of the K product quality characteristics.

In order to enable sample estimation of the population coefficients in the CEF, the assumption is made that the intercept and slope coefficients remain constant for all of the models considered during time period t ($\alpha_i^t = \alpha^t$ and $\beta_{ki}^t = \beta_k^t$). According to this assumption the CEF is:

$$E(P|i, t) = \alpha^t + \sum_{k=1}^K \beta_k^t q_{ki}.$$

The estimated Sample Regression Function for model i in all time periods is:

$$\hat{p}_i^t = \hat{\alpha}^t + \sum_{k=1}^K \hat{\beta}_k^t q_{ki}. \quad (1)$$

Under the usual assumptions, unbiased estimates $\hat{\alpha}^t$ and $\hat{\beta}_k^t$, of the $K+1$ population coefficients are obtained by performing $T+1$ Ordinary Least Squares (OLS) regressions utilizing the observed prices, p_i^t , and product quality characteristics, q_{ki} . The estimate of the conditional mean price of model i in some specific time period $t = s$, \hat{p}_i^s , obtained from the sample is the best minimum Mean Squared Error (MSE) prediction of the market price of model i in time period $t = s$.

Those factors present in the marketplace that exert an influence upon the prices of all models are particularly prone to variations over time. For that reason, the value of the intercept term is likely to vary. In contrast, the population slope coefficients, β_k^t , relating the model's price to its product quality characteristics, could conceivably conform to three plausible alternative assumptions.

Assumption 1 *The population slope coefficients vary throughout the $T+1$ time periods considered.*

In this case, provided a sufficient quantity of data is available in each of the time periods, the unbiased estimates of the population coefficients in Sample Regression Function (1), $\hat{\alpha}^t$ and $\hat{\beta}_k^t$, can be obtained by performing $T+1$ separate OLS regressions. Alternatively, estimation could be accomplished by pooling the data from all of the time periods into a single data set where N different models are represented. This single equation estimation procedure requires time dummy variables, d_i^s , that are defined for certain specific time periods $t=s$. The estimated single Sample Regression Function is:

$$\hat{p}_i^t = \hat{\alpha} + \sum_{s=1}^T \hat{\gamma}^s d_i^s + \sum_{k=1}^K \hat{\beta}_k q_{ki} + \sum_{k=1}^K \sum_{s=1}^T \hat{\delta}_k^s d_i^s q_{ki}, \quad (2)$$

where

$$d_i^s = \begin{cases} 1 & \text{if model } i \text{ is observed in time period } s \\ 0 & \text{if model } i \text{ is observed in time period } s-1. \end{cases}$$

In this formulation, $\hat{\beta}_k$ is an unbiased estimate of the population slope coefficient β_k^0 , while $(\hat{\beta}_k + \hat{\delta}_k^s)$ is an unbiased estimate of β_k^s . Therefore, the estimated dummy variable coefficients, $\hat{\delta}_k^s$, provide unbiased estimates of the differences between the magnitudes of the population slope coefficients, β_k^t , in time periods $t=s$ and $t=0$. In a similar manner, the estimated dummy variable coefficients, $\hat{\gamma}^s$, provide unbiased estimates of the difference between the magnitude of the population intercept terms, α^t , in time periods $t=s$ and $t=0$. Sample Regression Function (1) and this single Sample Regression Function (2) will both produce the same unbiased estimates of the population slope coefficients.

Assumption 2 *The population slope coefficients remain constant throughout the $T+1$ time periods considered.*

Under this assumption the CEF is:

$$E(P|i, t) = \alpha^t + \sum_{k=1}^K \beta_k q_{ki}.$$

In this case, unbiased estimates of the population coefficients can be obtained once again by pooling all of the data from the $T+1$ time periods and estimating a simplified version of single Sample Regression Function (2):

$$\hat{p}_i^t = \hat{\alpha} + \sum_{s=1}^T \hat{\gamma}^s d_i^s + \sum_{k=1}^K \hat{\beta}_k q_{ki}, \quad (3)$$

where d_i^s is defined as in Sample Regression Function (2).

Assumption 3 *The population slope coefficients vary over some of the time periods considered while remaining constant over the others.*

This assumption is simply a combination of Assumptions 1 and 2. In this case, estimation of the population coefficients over the complete $T+1$ time periods is accomplished by selectively applying either single Sample Regression Function (2) or (3). The selection of these regression functions is made depending upon whether the population slope coefficients are assumed to vary or to remain constant over the range of time periods under consideration.

2.2 The AYR Method

In order to obtain the values of the estimated coefficients in Sample Regression Function (1), a sufficient amount of cross-sectional data is required in each of the time periods considered. Unfortunately, an ample quantity of data is not always available. A number of studies have found a solution for this problem by using scanner data (e.g., see Silver, 1999 or Silver and Heravi, 2001). Scanner data, however, is often prohibitively expensive. Furthermore, when attempting to estimate longer-term price trends retrospectively, scanner data is seldom available. These considerations necessitate the search for alternative sources of publicly available data. Product catalogues and the product test reports published in trade magazines often provide the answer. Unfortunately, however, the problems associated with a limited amount of data more often than not remain.

The AYR method was proposed as a pragmatic “solution” to this data shortage problem. The method enables estimation by pooling observations from a selected number of adjacent time periods. Estimation is conducted under the assumption that the population intercept term varies from period to period. In contrast, however, the slope coefficients are assumed to remain constant over the entire pooled time span.

For expository purposes, consider the case where the AYR method is implemented by pooling data from two adjacent time periods, $s-1$ and s for $1 \leq s \leq T$. In order to identify each of the T separate data sets that can be produced in this manner, an index τ is introduced that takes a numerical value equal to s . Therefore, the $(\tau = 1)^{\text{th}}$ data set contains pooled data from time periods 0 and 1; the $(\tau = 2)^{\text{th}}$ contains pooled data from time periods 1 and 2; and in general, the τ^{th} contains pooled data from time periods $s-1$ and s . Observed price data from N^τ different

models is contained in the τ^{th} pooled data set, each model identified by the subscript i ($i = 1, 2, \dots, N^\tau$). In these data sets, price data is available for some models in both time periods (p_i^{s-1} and p_i^s) while other models are only observed in one of the two time periods, e.g., N^s models are observed in time period s and N^{s-1} in time period $s-1$, where $N^\tau \leq N^s + N^{s-1}$.

A salient feature of the AYR method is the use of these overlapping pooled data sets in its estimation procedure. Estimation, however, requires the inclusion of a Time Dummy Variable (TDV), d_i^τ , in order to permit variation in the estimate of the population intercept term. The estimated two-period Sample Regression Function (hereafter referred to as a TDV regression) that utilizes the τ^{th} pooled data set in its estimation is:

$$\hat{p}_i^\tau = \hat{\alpha}^\tau + \hat{\gamma}^\tau d_i^\tau + \sum_{k=1}^K \hat{\beta}_k^\tau q_{ki}, \quad \tau = 1, 2, \dots, T, \quad (4)$$

where

$$d_i^\tau = \begin{cases} 1 & \text{if model } i \text{ is observed in time period } s \\ 0 & \text{if model } i \text{ is observed in time period } s-1. \end{cases}$$

TDV regression (4) yields an estimate of the conditional mean price of model i in both time periods because of the inclusion of the dummy variable, d_i^τ . When $d_i^\tau = 1$, the function produces an estimate of the conditional mean price of model i in time period s , $\hat{p}_i^\tau = \hat{p}_i^s$, and when $d_i^\tau = 0$, it yields an estimate of the conditional mean price of model i in time period $s-1$, $\hat{p}_i^\tau = \hat{p}_i^{s-1}$. Moreover, the possibility does exist to produce estimates of the conditional mean prices of models that were not observed in either of these time periods simply by inserting their quality characteristics into TDV regression (4). Estimation of the TDV regressions is conducted under the assumption that the population slope coefficients in the two pooled time periods are equal, $\beta_k^{s-1} = \beta_k^s$. If this assumption is correct, then the estimated coefficients, $\hat{\beta}_k^\tau$, are unbiased. Furthermore, the estimated intercept coefficient $\hat{\alpha}^\tau$ is an unbiased estimate of the population intercept term α^{s-1} , and $(\hat{\alpha}^\tau + \hat{\gamma}^\tau)$ is an unbiased estimate of the population intercept term α^s .

This regularly used overlapping estimation procedure, however, has some rather serious inherent theoretical shortcomings. Court (1939) and subsequently other authors have referred to some of the AYR method's various problems. Largely ignored heretofore, however, is the fact that the AYR method produces either biased or at best inefficient estimation results. This can be demonstrated by considering

the three alternative assumptions outlined hereinbefore in Section 2 (Assumptions 1, 2, and 3) pertaining to the population slope coefficients, β_k^t .

If Assumption 1 were true, the population slope coefficients would vary throughout the $T+1$ time periods considered: $\beta_k^s \neq \beta_k^{s-1}$ for $1 \leq s \leq T$. Unbiased estimates of the population coefficients of periods s and $s-1$ could only be obtained by utilizing a two-period version of the single Sample Regression Function (2),

$$\hat{p}_i^\tau = \hat{\alpha}^\tau + \hat{\gamma}^\tau d_i^\tau + \sum_{k=1}^K \hat{\beta}_k^\tau q_{ki} + \sum_{k=1}^K \hat{\delta}^\tau d_i^\tau q_{ki}, \quad \tau = 1, 2, \dots, T, \quad (5)$$

where d_i^τ is defined as in TDV regression (4). Therefore, the AYR method's TDV regression (4) is improperly specified. This is due to the well-known omission of relevant variables problem, specifically in this case, the K interaction terms, $d_i^\tau q_{ki}$. Consequently, every one of the estimated coefficients ($\hat{\alpha}^\tau$, $\hat{\gamma}^\tau$, and $\hat{\beta}_k^\tau$) from the T two-period pooled TDV regressions (4) is a biased estimate of its corresponding population coefficient, except for those coefficients relating to variables that are uncorrelated with the omitted variables $d_i^\tau q_{ki}$.

Result 1 *In the case of Assumption 1, the AYR method produces estimated coefficients, $\hat{\alpha}^\tau$, $\hat{\gamma}^\tau$, and $\hat{\beta}_k^\tau$, that are biased estimates of the corresponding population coefficients in all of the $T+1$ time periods.*

If Assumption 2 were true, the population slope coefficients would remain constant throughout the T time periods considered. The AYR method relies on overlapping TDV regressions that are conducted under the assumption that the population slope coefficients are only equal over the two time periods pooled. In this case, efficient estimates of the population slope coefficients can be obtained by pooling all of the data from the $T+1$ time periods and estimating the single Sample Regression Function (3). Consequently, the AYR method's use of TDV regression (4) is an inefficient estimation procedure.

Result 2 *In the case of Assumption 2, the AYR method produces estimated coefficients, $\hat{\alpha}^\tau$, $\hat{\gamma}^\tau$, and $\hat{\beta}_k^\tau$, that are unbiased but inefficient estimates of the corresponding population coefficients in all of the $T+1$ time periods.*

If Assumption 3 were true, then the population slope coefficients would vary over some of the time periods considered while remaining constant over others. As indicated heretofore, the AYR method's use of TDV regression (4) produces either

1
2
3 biased or inefficient estimates of the population slope coefficients. Throughout those
4 adjacent time periods where the population slope coefficients are varying, the AYR
5 method's overlapping TDV regressions would produce biased estimated coefficients.
6 On the other hand, over those adjacent time periods where the population slope
7 coefficients remain constant, the AYR method's estimated coefficients are unbiased,
8 but they are inefficiently estimated.
9

10 Due to the AYR method's use of overlapping data sets in the estimation of its
11 TDV regressions, two different sets of estimated slope coefficients are generated for
12 each of the overlapped time periods. Consider the case where the population slope
13 coefficients have remained constant for several time periods prior to and including
14 time period $s-1$, but then they change so that in time period s they are different.
15 In such a case, the TDV regression that was estimated using pooled data from time
16 periods $s-2$ and $s-1$ would produce inefficient estimates of the population slope
17 coefficients in time period $s-1$. The TDV regression that was estimated using pooled
18 data from time periods $s-1$ and s , on the other hand, would produce estimated
19 coefficients for time period $s-1$ (as well as time period s) that are biased.
20
21

22
23 **Result 3** *In the case of Assumption 3, the AYR method produces estimated coef-*
24 *ficients, $\hat{\alpha}^\tau$, $\hat{\gamma}^\tau$, and $\hat{\beta}_k^\tau$, that are biased estimates of the corresponding population*
25 *coefficients in those time periods where the population slope coefficients are varying,*
26 *and inefficient estimates in those time periods where they remain constant for more*
27 *than two periods.*
28
29

30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60

2.3 The Flexible AYR Method

To the best of our knowledge, the theoretical shortcomings inherent in the AYR method, even though they are rather obvious, have never been explicitly brought out into the open or adequately elucidated. Even in published studies utilizing a more flexible AYR approach, supported by hypothesis testing relating to parameter stability, these deficiencies have not been adequately addressed.

A well-known study by Berndt and Rappaport (2001) provides a good case in point. Investigating the price trend of personal computers (both desktop and laptop models) over the time period from 1976 to 1999, these authors conducted hypothesis tests on each pair of adjacent time periods in order to ascertain whether the constant slope coefficient assumption was a reasonable one before estimating the coefficients in the two-period TDV regression (4). In order to accomplish this, they first estimated

the coefficients in each of the T TDV regressions (5). For each of these regressions, F -Tests ($H_0: \delta_1^\tau = \delta_2^\tau = \dots = \delta_K^\tau = 0$) were conducted for the purpose of determining whether or not the K estimated coefficients, $\hat{\delta}_k^\tau$, were in fact close enough to zero to conclude that the population slope coefficients were actually constant over the two time periods considered, $\beta_k^s = \beta_k^{s-1}$. If this proved to be the case, then the coefficients in TDV regression (4) were estimated. If not, the two time periods were estimated separately yielding the coefficients in Sample Regression Function (1). These estimates are the same as those in TDV regression (5). When utilizing Regression Function (1), provided that sufficient data is available, a separate set of K slope coefficients, $\hat{\beta}_k^t$, are estimated for each of the time periods. This procedure could be called the Flexible AYR method.

Berndt and Rappaport estimated the coefficients in TDV regression (5) using pooled data from the time periods 1989-1990 and 1990-1991. In both of these regressions only laptops were considered and the null hypothesis could not be rejected. Therefore, the population slope coefficients were deemed to be equal over the time periods pooled: $\beta_k^{1989} = \beta_k^{1990}$ and $\beta_k^{1990} = \beta_k^{1991}$. As a consequence, they estimated the coefficients in TDV regression (4) using the pooled data sets $\tau = 1990$ and $\tau = 1991$. If in fact it is true that $\beta_k^{1989} = \beta_k^{1990}$ and $\beta_k^{1990} = \beta_k^{1991}$, then necessarily it is true that $\beta_k^{1989} = \beta_k^{1990} = \beta_k^{1991}$.

By conducting an F -test on a three period variant of TDV regression (5), one could statistically ascertain whether this latter equality is likely to be true. If equality could be (correctly) accepted, then using two separate TDV regressions instead of a single TDV regression covering the complete time span 1989-1991 would be inefficient. On the other hand, if equality could be (correctly) rejected, then at least one of the two TDV regressions performed by Berndt and Rappaport would suffer from omitted-variable bias. In other words, their Flexible AYR method does not resolve the inefficient estimation problem, but it avoids some of the situations where estimation bias could occur. Inefficiency, however, can be avoided by employing a minor procedural modification. This modification produces a procedure called the NOAYR method.

2.4 The NOAYR Method

Consider a population where Assumption 3 holds true. Under this set of circumstances, the population slope coefficients would vary over some of the time periods

considered while they would remain constant over the others. The estimation problems inherent in the AYR and Flexible AYR methods arise from the use of the overlapping time periods in the TDV regressions that are conducted. By simply avoiding these time period overlaps these problems can be eliminated. The improved method presented here is called the NOAYR method (“NO” for “non-overlapping”).

Estimation utilizing the NOAYR method is accomplished in the following manner. Consider the simple case where only one product quality characteristic, q_i , is considered and thus the estimated two-period TDV regression (4) simplifies to:

$$\hat{p}_i^\tau = \hat{\alpha}^\tau + \hat{\gamma}^\tau d_i^\tau + \hat{\beta}^\tau q_i, \quad \tau = 1, 2, \dots, T. \quad (6)$$

Furthermore, assume that the population slope coefficients vary only between time periods $s-1$ and s , such that $\beta^s > \beta^{s-1}$ for $1 \leq s \leq T$. The population slope coefficients remain constant in every other time period considered. The time-series of the population slope coefficient values, β^t , is indicated by the black dots in Figure 1.

Here Figure 1:

The AYR, Flexible AYR, and NOAYR Methods.

In Figure 1, the brackets indicate the time periods that were pooled when estimating the respective TDV regressions. For example, in the AYR method, observations from each pair of adjacent periods would be pooled into a separate data set. This would generate a sequence of overlapping data sets and associated with each data set an estimated TDV regression of the form (6). Estimation utilizing the τ^{th} data set would produce biased estimates of the population coefficients α^{s-1} and α^s as well as β^{s-1} and β^s .

If hypothesis testing for parameter stability correctly indicated that $\beta^{s-1} \neq \beta^s$, the Flexible AYR method would avoid the estimation bias inherent in the AYR method by not pooling these time periods and therefore not estimating this TDV regression (see Figure 1). The Flexible AYR method, however, does not incorporate the information that the population slope coefficients remain constant over the time span 0 to $s-1$ and also over the time span s to T . As a consequence, the Flexible AYR method does not produce efficient estimates.

In order to remove this inefficiency in this simple case, the NOAYR method estimates only two TDV regressions. The first TDV regression is estimated using

1
2
3 pooled data from all of the time periods between 0 and $s-1$. The second TDV
4 regression pools data from all the time periods between s and T (see Figure 1).
5 Unbiased, as well as efficient, estimates of the population coefficients are obtained
6 for all of the $T+1$ time periods by simply performing these two TDV regressions.
7 The modification necessary to turn the AYR method into the NOAYR method is
8 so trivial that one might rightly ask, Why is the AYR methodology, one that uses
9 overlapping pooled data sets to estimate the TDV regressions, still so popular? A
10 possible answer is provided in Section 3.

11
12 For all practical purposes, the appropriate guidelines for applying the NOAYR
13 method are as follows. First, hypothesis testing should be used to determine whether
14 the values of the slope coefficients change over time. Second, if these tests indicate
15 that the slope coefficients remain constant over two or more adjacent years, then all
16 of these years should be pooled into a single data set and estimation of the TDV
17 regression should be conducted using additional time dummy variables. Finally,
18 those adjacent years where tests indicate that the coefficients change, should not
19 be pooled. If in some period the values of the slope coefficients are such that no
20 adjacent time period exists with slope coefficients that possess the same values, then
21 for this period a separate single year regression utilizing Sample Regression Function
22 (1) must be estimated, provided sufficient data is available.

23
24 If no period exists in which the slope coefficients remain unchanged (as in As-
25 sumption 1), then all of the estimated coefficients of the AYR method's overlapping
26 TDV regressions are biased. In this case, however, the NOAYR method unfortu-
27 nately does not provide a solution either. The NOAYR method also relies on some
28 (although fewer) TDV regressions and therefore upon the assumption of constant
29 slope coefficients. If at least some periods exist in which the slope coefficients re-
30 main constant (as in Assumption 3), then the NOAYR method could represent a
31 useful and reliable method for long-term price measurement. In contrast, the AYR
32 method with its overlapping TDV regressions produces biased estimated coefficients
33 and therefore should be avoided.

34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 **2.5 Related Literature**

54
55 As indicated in Section 2.3, the Flexible AYR method employed by Berndt and
56 Rappaport (2001) avoids estimation bias by not estimating TDV regressions using
57 pooled data from time periods where hypothesis testing indicates that the popula-
58
59
60

tion slope coefficients, β_k^t , are changing. Their Flexible AYR method, however, fails to estimate using pooled data from consecutive time periods where the population slope coefficients were deemed to be invariant over time. As a consequence, the Flexible AYR method produces unbiased estimates, but unfortunately the estimation procedure is inefficient. The converse is true of the AYR methods employed in studies such as Berndt and Griliches (1993), Nelson *et al.* (1994), and Chwelos (2003). These studies pool data for use in several of their TDV regressions but the remaining ones rely on the traditional overlapping method. It was demonstrated in Section 2.3 that this methodology produces biased coefficient estimates.

A rather different approach assumes that the population slope coefficient values β_k^t change over time, but that they do so in a very systematic way that can be modeled according to some functional relationship. Studies utilizing this approach include Berndt and Griliches (1993), Nelson *et al.* (1994), Baker (1997), and more recently, Chwelos (2003) and Auer (2005). Stengos and Zacharias (2005) have proposed yet another estimation approach. They adopt a semi-parametric hedonic estimation framework.

3 Estimation of Long-Term Price Trends

In the following, an estimate of the individual price relative of model i over time periods s and $s-1$ is denoted by \hat{r}_i^τ . It is defined as the ratio of the estimated conditional mean price of model i in time period s to its estimated value in time period $s-1$:

$$\hat{r}_i^\tau = \hat{p}_i^s / \hat{p}_i^{s-1} \quad , \quad \tau, s = 1, 2, \dots, T \quad \text{and} \quad i = 1, 2, \dots, N^\tau \quad . \quad (7)$$

Normally, however, a more useful measure is an estimate of a comprehensive price index, \hat{r}^τ . This price index measures price changes for a whole product category (e.g., laser printers) and it can be derived in a variety of ways. The comprehensive price index is generally calculated as some function of the individual price relatives,

$$\hat{r}^\tau = g(\hat{r}_1^\tau, \hat{r}_2^\tau, \dots, \hat{r}_{N^\tau}^\tau) \quad .$$

A particularly useful functional form in this regard is the Jevons index:

$$\hat{r}^\tau = \exp \left[(1/N^\tau) \sum_{i=1}^{N^\tau} \ln \hat{r}_i^\tau \right] \quad . \quad (8)$$

Longer-term price trends can be calculated using the chained price index I^t . The value of this price index in time period $t = 0$ is $I^0 = 100$. In subsequent time periods its value is given by:

$$I^t = 100 \cdot \prod_{\tau=1}^t \hat{r}^\tau, \quad t = 1, 2, \dots, T. \quad (9)$$

3.1 Price Relatives in the AYR Method

More often than not, estimation of the coefficients in TDV regression (4) is accomplished using the natural logarithm of the observed prices, $\ln p_i^t$. This semi-logarithmic specification yields a particularly useful result. According to this specification, the estimated coefficient on the time dummy variable, $\hat{\gamma}^\tau$, is:

$$\hat{\gamma}^\tau = \ln \hat{p}_i^s - \ln \hat{p}_i^{s-1} = \ln \hat{r}^\tau.$$

According to this commonly used specification, all models now exhibit exactly the same price relative, $\hat{r}_i^\tau = \hat{r}^\tau$, and therefore the price index, \hat{r}^τ , is easily calculated as:

$$\exp(\hat{\gamma}^\tau) = \hat{r}^\tau. \quad (10)$$

It has been pointed out in various studies (e.g., Goldberger, 1968), however, that least squares estimation of this logarithmic specification leads to bias in the estimated conditional mean prices, \hat{p}_i^t . Fortunately, in hedonic regression analysis this bias is usually very small and therefore it is customarily ignored. This present study is no exception. Dalen and Bode (2004) provide a thorough discussion of this type of bias.

When the semi-logarithmic specification is used in the AYR method, Equation (10) states that the estimated price indices for all time periods, \hat{r}^τ , can be easily calculated using the estimated coefficients on the time dummy variable, $\hat{\gamma}^\tau$. The simplicity of this approach is one plausible explanation for the AYR method's widespread popularity. Unfortunately, when the Flexible AYR or the NOAYR method is employed there can be pairs of adjacent time periods that were not pooled for estimation purposes. Consequently, there will be no estimated time dummy variable coefficients, $\hat{\gamma}^\tau$, for these time periods. In this case, the estimation of the price index, \hat{r}^τ , will require some alternative approach. The following two methods are the approaches that were used in the empirical section of this paper.

3.2 Price Relatives from Estimated Prices

Suppose that the population slope coefficients, β_k^t , vary only between time periods s and $s-1$. It is assumed that they remain constant throughout all other time periods (as in Figure 1). Under this assumption, the NOAYR method would estimate the coefficients for only two TDV regressions of the type defined by Sample Regression Function (3). The first of these TDV regressions would employ a data set that pools all of the data from time periods 0 to $s-1$. Similarly, the second TDV regression would pool data from time periods s to T . In addition, assume that the semi-logarithmic specification was employed.

In this case, an estimate of the price index, \hat{r}^τ , cannot be calculated using Equation (10) because an estimate of the dummy variable coefficient, $\hat{\gamma}^\tau$, is not available. Estimates of the logarithms of the conditional mean prices of models $i = 1, 2, \dots, N^\tau$ in time period $s-1$, $\ln \hat{p}_i^{s-1}$, can be obtained directly from the first estimated TDV regression. Analogously, the logarithms of the conditional mean prices of model i in time period s , $\ln \hat{p}_i^s$, can be obtained directly from the second estimated TDV regression. The price index, \hat{r}^τ , can now be calculated using Equations (7) and (8).

3.3 Price Relatives from Observed and Estimated Prices

Recently, a “composite method” of calculating price relatives has become increasingly popular (e.g., Triplett, 2004). Assume that the same situation that was outlined in Section 3.2 prevails and again an estimate of the comprehensive price index, \hat{r}^τ , cannot be calculated using Equation (10). As before, pooling the data from periods s and $s-1$ generates the τ^{th} data set. The composite method divides this data set into two subsets according to whether the price of model i was observed in both time periods or in only one of them. In other words, a strict distinction is made between those M product models where the prices were observed in both time periods, hereafter called the “matched models”, and those $(N^\tau - M)$ other models where the price was observed in only one of the two time periods.

The price relatives of the M matched models are computed using their actual observed prices,

$$r_i^\tau = p_i^s / p_i^{s-1} \quad (i = 1, 2, \dots, M) . \quad (11)$$

A price index consisting of the matched models, called the Matched Model Index, r_M^τ , can be computed using the Jevons Index defined by Equation (8).

For the other $(N^\tau - M)$ product models, referred to as the non-matched models, only one of these two prices was observed. The price relatives for each of these models can either be calculated using Equation (7), that is, using only estimated prices, or it can be calculated as a mixture of an observed and an estimated price:

$$\hat{r}_i^\tau = \hat{p}_i^s / p_i^{s-1} \quad \text{or} \quad \hat{r}_i^\tau = p_i^s / \hat{p}_i^{s-1} \quad (i = M+1, M+2, \dots, N^\tau), \quad (12)$$

depending on which price was not observed. In both approaches, the estimated prices can be derived according to the procedure outlined in Section 3.2.

Which approach produces the “best” price relatives, Equation (7) or Equation (12)? If a particular model has individual qualitative characteristics that were not completely captured by the K product quality characteristics, q_{ki} , used in the estimation (e.g., a unique design), then estimation bias will occur as a result of misspecification. In this case, the use of Equation (12) would result in biased estimates. The use of Equation (7) would mitigate this type of bias.

From the relative prices of the $(N^\tau - M)$ non-matched models, a price index, \hat{r}_N^τ , can be generated by the Jevons Index defined in Equation (8). In order to arrive at the comprehensive price index, \hat{r}^τ , however, the price indices corresponding to the two subsets, r_M^τ and \hat{r}_N^τ , need to be combined. For this purpose, a weighted geometric average can be calculated where the weight, ω , attached to r_M^τ is the proportion of the matched model prices in the total number of observed prices in the two-period data set, $\omega = [2M / (N^\tau + M)]$. The weight assigned to \hat{r}_N^τ is $(1 - \omega)$. This weighting scheme assures that the informational content derived from the data is properly represented in the calculation of the comprehensive price index,

$$\hat{r}^\tau = (r_M^\tau)^\omega \cdot (\hat{r}_N^\tau)^{1-\omega}. \quad (13)$$

The use of the composite method is not confined to two adjacent time periods that were not pooled for estimation purposes. This method can also be applied if the two time periods were pooled for estimation purposes. Suppose, for example, that $\beta_k^{s-1} = \beta_k^s$ and therefore the coefficients in TDV regression (4) were estimated. Once again, the price relatives of the matched models can be computed using Equation (11) and the matched model index, r_M^τ , can be computed using Equation (8). The price index of the non-matched models, \hat{r}_N^τ , however, is set equal to $\exp(\hat{\gamma}^\tau)$, where $\hat{\gamma}^\tau$ is the estimated coefficient in TDV regression (4). Once again, the price index \hat{r}^τ is calculated using Equation (13).

In the empirical section of this paper, the approaches described in Sections 3.1 (AYR price trend), 3.2 (NOAYR price trend), and 3.3 (another NOAYR price trend)

1
2
3 were utilized. There are, however, various other conceivable methods. For example,
4 Berndt and Rappaport (2001) follow an alternative approach that was proposed by
5 Dhrymes (1971) and further developed by Cowling and Cubbin (1972).
6
7
8
9

10 4 Data Description

11
12
13 Currently, even the introductory macroeconomic textbooks (e.g., Blanchard, 2003,
14 p. 256) are expressing misgivings concerning the accuracy of the officially published
15 inflation statistics issued by national statistical agencies. These suspicions are in
16 accord with the published empirical results from numerous hedonic regression studies
17 that have been conducted using data on computer prices. Consequently, evidence
18 is mounting that the officially released price trends tend to underestimate the true
19 quality-adjusted ones. A fair and comprehensive assessment of this suspicion is to
20 be found in ILO *et al.* (2004).
21
22
23
24
25

26
27 Apart from computers, however, evidence of price trend measurement bias for
28 other products in the IT sector is for the most part lacking. For example, there are no
29 comparable studies using price data for computer monitors, printers, or for any other
30 such related product, Cole *et al.* (1986) and Moreau (1996) being notable exceptions.
31 This omission is somewhat surprising, since computer peripherals represent a sizable
32 part of the total expenditure on IT products. In an effort to alleviate this oversight,
33 this study provides a look at the quality-adjusted long-term price trend of laser
34 printers in the Federal Republic of Germany from 1993 to 2004. The estimated price
35 trend obtained has been compared with the officially published one released by the
36 *Statistisches Bundesamt* (Federal Statistical Office of Germany). The *Statistisches*
37 *Bundesamt* has computed and published a separate quality-adjusted price index for
38 laser printers only since 2000. Prior to 2000 the published price trend for ink jet
39 printers has been used as proxy for these missing values and prior to 1995 computer
40 prices have been used for this purpose.
41
42
43
44
45
46
47
48
49

50
51 The information utilized in this study was obtained from a sequence of test
52 reports that were published in a bi-weekly German language computer trade mag-
53 azine entitled: *c't - magazin für computertechnik*. Many experts regard these test
54 results to be among the most reliable and sophisticated in the field. For that reason,
55 test results appearing in other computer magazines were deliberately barred from
56 consideration in the study. This policy ensured an unusually coherent data set of
57 385 observations. Twelve time-invariant qualitative characteristics were utilized, q_{ki}
58
59
60

($k = 1$ to 12). Six of these were measured as continuous variables, while dichotomous variables represented the others. The continuous variables measure such characteristics as: printing speed, standard memory size, added memory size, maintenance cost per page, paper capacity, and printing resolution. The dichotomous variables represented the features: equipped with postscript, upgraded with postscript, equipped with network connectivity, upgraded with network connectivity, equipped with extra interface (provided that the printer has no network connectivity), and GDI-printer (Graphical Device Interface). A complete description of the data set used in this study, in conjunction with the data collection process, can be obtained on request from the authors.

Some selected summary statistics are presented in Table 1. The second column shows the number of observations per year. The third column gives the average price in EUR for the printers. The table demonstrates that in the years 1996 and 1997, primarily “low quality” printers were tested, whereas in the years 1998 and 1999, many of the “top-of-the-line” models received consideration. By and large, a clear downward trend in model prices was observed. The last column displays the average printing speed measured in pages per minute. These numbers vividly indicate that a significant amount of technical progress has been occurring during this time period.

Here Table 1:

Summary Statistics from the Laser Printer Data.

5 Empirical Specification and Results

Before estimation began, certain diagnostic tests were conducted. These tests suggested a semi-logarithmic functional form for the Sample Regression Function. Based on this functional form, three different long-term price trends were computed. One is based on the AYR method and the other two on the NOAYR method.

As described in Section 2.2, the AYR method that was applied is based upon OLS estimation using a sequence of overlapping pooled data sets, each of which pools two consecutive years, s and $s-1$. The estimated coefficient standard errors in the TDV regressions were corrected for heteroskedasticity.

The estimated two-period TDV regression that was employed is:

$$\ln \hat{p}_i^\tau = \hat{\alpha}^\tau + \hat{\gamma}^\tau d_i^\tau + \sum_{k=1}^K \hat{\beta}_k^\tau q_{ki}, \quad \tau = 1994, 1995, \dots, 2004. \quad (14)$$

where

$$d_i^\tau = \begin{cases} 1 & \text{if model } i \text{ is observed in time period } s \\ 0 & \text{if model } i \text{ is observed in time period } s-1. \end{cases}$$

The dependent variable in TDV regression (14), $\ln \hat{p}_i^\tau$, is in a natural logarithmic form. As a consequence, the estimated slope coefficients on the product quality variables, $\hat{\beta}_k^\tau$, are no longer indicating the implicit price contribution of these respective characteristics. Instead, they are estimating the semi-elasticities of each printer's conditional mean price with respect to these qualitative characteristics.

The AYR method, based on TDV regression (14), directly produces a time series of estimated coefficient values for each time period, $\hat{\gamma}^\tau$. Utilizing Equation (10), these coefficients can be used to produce the required price indices, \hat{r}^τ . Therefore, an AYR chained price index for use in measuring the quality-adjusted long-term price trend can be defined in terms of these price indices \hat{r}^τ . The value of this chained price index in time period t (define $I^{1993} = 100$) is:

$$I^t = 100 \cdot \prod_{\tau=1994}^t \hat{r}^\tau, \quad t = 1994, 1995, \dots, 2004.$$

The estimated AYR quality-adjusted long-term price trend for laser printers is depicted in Figure 2. According to this price trend, the prices of laser printers fell by 76.6 percent. The absolute value of the corresponding average annual growth rate is $|AAGR| = 12.4$ percent.

Here Figure 2:

Comparison of Laser Printer Price Trends.

The calculation of the AYR price trend was based upon estimates obtained from the TDV regressions that were conducted. Each of these TDV regressions utilized pooled data from two consecutive years. Hypothesis testing was conducted under the null hypothesis that the slope coefficients were equal over the two pooled years. These tests revealed that the null hypotheses could only be accepted at a 5 percent significance level in two of the data sets, namely $\tau = 2000$ and $\tau = 2004$. In fact, for all other pairs of adjacent years the null hypothesis that the slope coefficients were equal was rejected at a 1 percent significance level.

1
2
3 For that reason, a NOAYR price trend was computed utilizing the NOAYR
4 method outlined in Section 2.4 and Section 3.2. In particular, TDV regressions
5 were conducted using pooled data from the years 1999-2000 and 2003-04. Separate
6 regressions were conducted for the years 1993, 1994, 1995, 1996, 1997, 1998, 2001,
7 and 2002. According to this price trend, the prices of laser printers fell by 76.4
8 percent ($|AAGR| = 12.3$ percent). The resulting price trend is also depicted in
9 Figure 2. There is a divergence from the AYR price trend, but the deviation is not
10 large.

11
12 A second NOAYR price trend was computed utilizing estimated as well as ob-
13 served prices as outlined in Section 3.3. The resulting price trend is not depicted
14 in Figure 2, because it runs very close to the other two price trends. The prices of
15 laser printers using this method fell by 75.9 percent ($|AAGR| = 12.1$ percent). This
16 second NOAYR price trend, however, revealed an interesting aspect. On average
17 the matched model index r_M^T showed significantly smaller rates of decline than did
18 the estimated index of the non-matched models, \hat{r}_N^T . This divergence lends empiri-
19 cal support to the widely held belief that a pure matched model index is likely to
20 underestimate the rates of decline in the prices of IT related products.

21
22 The similarity of the AYR price trend and the two NOAYR price trends should
23 not be misconstrued as an excuse for utilizing the AYR method in situations where
24 the NOAYR method is feasible. In such situations the AYR method produces bi-
25 ased (or at best inefficient) estimates and these shortcomings should be avoided
26 by applying the NOAYR method. An interesting conclusion can be drawn, if one
27 compares the AYR and the NOAYR price trends of the years 1993 to 1998. For the
28 computation of the NOAYR price trend for each of these years separate regressions
29 were conducted and from those the NOAYR price trend was calculated. The results
30 were similar to those obtained from the overlapping TDV regressions conducted for
31 computing the AYR price trend. This result lends support to the conjecture that
32 in situations where a lack of data renders separate regressions for each time period
33 impossible and where the population slope coefficients vary throughout the time pe-
34 riods considered (Assumption 1 in Section 2.1), the measurement bias arising from
35 the utilization of the AYR method might not be very large.

36
37 The official price trend published by the *Statistisches Bundesamt* is also displayed
38 in Figure 2. Using the two quality-adjusted price trends (AYR and NOAYR) as a
39 point of reference, the official price trend does not appear to have reacted to the con-
40 siderable price decline that occurred between 1993 and 2004. The official price trend
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 falls by 52.7 percent over the entire time period considered ($|AAGR| = 6.4$ percent).
4 The quality-adjusted hedonic price trends measure a price decline of roughly 76 per-
5 cent ($|AAGR| > 12.0$ percent). This comparison provides some additional evidence
6 for the suspicion that not only computers are experiencing impressive downward
7 price trends but also other IT products as well. Furthermore, it underlines the fact
8 that frequently the officially published estimates are not accurately measuring these
9 price trend declines.
10
11
12
13
14

15 16 17 **6 Conclusions**

18
19 For those public agencies responsible for the accurate measurement of price trend
20 developments, adjusting for quality changes in many of today's technologically ad-
21 vanced products poses a serious challenge. In many cases, hedonic price measure-
22 ment techniques do offer a solution to this problem and can produce the quality-
23 adjusted price trend measurements required. This type of analysis, however, is labor
24 intensive and can frequently absorb a considerable amount of the agency's scarce
25 resources. For that reason, hedonic price measurement should be limited to those
26 selected product groups that are particularly prone to this type of measurement
27 bias. In order to identify these product categories, long-term price trends can be
28 estimated and compared to the officially reported price trends. For this purpose,
29 the AYR method has been the hedonic measurement technique most often utilized.
30 The AYR method uses a sequence of overlapping pooled data sets to estimate the
31 coefficients in the TDV regressions. This paper has demonstrated, however, that
32 these overlapping regressions should be avoided. A minor modification to the AYR
33 method has been introduced that avoids the theoretical shortcomings inherent in
34 the AYR method. This modification to the AYR method was called the NOAYR
35 method.
36
37
38
39
40
41
42
43
44
45
46
47

48 In the empirical section of this paper, both the AYR and the NOAYR method
49 were applied to data derived from publicly available test reports concerning laser
50 printers. These reports were published in a respected bi-weekly German computer
51 trade magazine. For each of these two hedonic regression methods, the price trend
52 of laser printers was computed for the time span 1993 to 2004. According to these
53 quality-adjusted price trends, the market prices of laser printers in Germany declined
54 by roughly 76 percent. Over this identical time period, however, the official price
55 trend reported by the *Statistisches Bundesamt* (Federal Statistical Office of Ger-
56
57
58
59
60

many) declined by less than 53 percent providing evidence of upward bias in their price trend measurement.

Deviations between hedonic price measurement and the published official price trends have been repeatedly reported in connection with computers. In response, various national statistical agencies have introduced hedonic techniques for measuring the price change of computers. The earliest example is the U.S. Bureau of Economic Analysis (see Cole *et al.*, 1986). Many other national statistical agencies have followed their lead. Recently, Linz and Eckert (2002) reported that the *Statistisches Bundesamt* has also taken serious steps towards a routinely computed hedonic price index for computers. The results presented in this study, however, indicate that it is prudent for national statistical agencies to also consider computing a hedonic price index for laser printers. Furthermore, additional studies need to be conducted in order to get a better picture concerning other major IT products.

Acknowledgements

We are grateful for the helpful suggestions that were received from an anonymous referee. Additionally, the benefit that this paper has enjoyed from the insightful comments offered by our colleagues here at the Faculty of Economics and Management, Otto-von-Guericke Universität Magdeburg, is gratefully acknowledged.

References

- Auer, L. von (2005) *Hedonic Price Measurement: The CCC Approach*, FEMM working paper series, No. 05023, Otto-von-Guericke-Universität Magdeburg.
- Baker, T. A. (1997) Quality-Adjusted Price Indexes for Portable Computers, *Applied Economics*, 29, 1115-23.
- Berndt, E. R. and Griliches, Z. (1993) Price Indexes for Microcomputers: An Exploratory Study", in *Price Measurements and Their Uses*, (Eds.) M. F. Foss, M. E. Manser and A. H. Young, NBER Studies in Income and Wealth, Vol. 57, University of Chicago Press, Chicago, 63-99.
- Berndt, E. R. and Rappaport, N. J. (2001) Price and Quality of Desktop and Mobile Personal Computers: A Quarter-Century Historical Overview, *American Economic Review*, 91(2), 268-73.

- 1
2
3 Blanchard, O. (2003) *Macroeconomics*, Prentice Hall, Upper Saddle River (New
4 Jersey).
- 5
6
7 Chwelos, P. (2003) Approaches to Performance Measurement in Hedonic Analysis:
8 Price Indexes for Laptop Computers in the 1990's, *Economics of Innovation
9 and New Technology*, 12(3), 199-224.
- 10
11
12
13 Cole, R., Chen, Y. C., Barquin-Stolleman, J. A., Dulberger, E., Helvacian, N. and
14 Hodge, J. H. (1986) Quality Adjusted Price Indexes for Computer Processors
15 and Selected Peripheral Equipment, *Survey of Current Business*, 66, January,
16 41-50.
- 17
18
19
20
21 Court, A. T. (1939) Hedonic Price Indexes with Automotive Examples, in *The
22 Dynamics of Automobile Demand*, General Motors Corporation, New York,
23 99-117.
- 24
25
26
27 Cowling, K. and Cubbin, J. (1972) Hedonic Price Indexes for United Kingdom
28 Cars, *Economic Journal*, 82, 963-78.
- 29
30
31
32 Dalen, J. van, Bode, B. (2004) *Estimation Biases in Quality-Adjusted Hedonic
33 Price Indices*, paper presented at the SSHRC-Conference, Vancouver 2004.
- 34
35
36
37 Dhrymes, P. J. (1971) Price and Quality Changes in Consumer Capital Goods:
38 Analysis of Quality Change, in *Price Indexes and Quality Change* (Ed.) Zvi
39 Griliches, Harvard University Press, Cambridge (Mass.), 88-149.
- 40
41
42
43 Goldberger, A. S. (1968) The Interpretation and Estimation of Cobb-Douglas Func-
44 tions, *Econometrica*, 36(3/4), 464-72.
- 45
46
47
48 Griliches, Z. (1971) Hedonic Price Indexes for Automobiles: An Econometric
49 Analysis of Quality Change, in *Price Indexes and Quality Change*, (Ed.) Z.
50 Griliches, Harvard University Press, Cambridge (Mass.), 55-87.
- 51
52
53
54 ILO, IMF, OECD, UNECE, Eurostat, The World Bank (2004) *Consumer Price
55 Index Manual: Theory and Practice*, International Labour Office, Geneva.
- 56
57
58
59
60 Linz, S. and Eckert, G. (2002) Zur Einführung hedonischer Methoden in die Preis-
statistik, *Wirtschaft und Statistik*, 10, 857-63.

- 1
2
3 Moreau, A. (1996) Methodology of the Price Index for Microcomputers and Print-
4 ers in France, in *Industry Productivity – International Comparison and Mea-*
5 *surement Issues*, OECD Proceedings, Paris Organization for Economic Co-
6 operation and Development, 99-118.
7
8
9
10 Nelson, R. A., Tanguay, T. L. and Patterson, C. D. (1994) A Quality-Adjusted Price
11 Index for Personal Computers, *Journal of Business and Economics Statistics*,
12 12(1), 23-31.
13
14
15
16 Silver, M. (1999) An Evaluation of the Use of Hedonic Regressions for Basic Compo-
17 nents of Consumer Price Indices, *Review of Income and Wealth*, 45(1), 41-56.
18
19
20 Silver, M. and Saeed, H. (2001) Scanner Data and the Measurement of Inflation,
21 *Economic Journal*, 111, 384-405.
22
23
24
25 Stengos, T. and Zacharias, E. (2005) Intertemporal Pricing and Price Discrimina-
26 tion: A Semiparametric Hedonic Analysis of the Personal Computer Market,
27 *Journal of Applied Econometrics*, forthcoming.
28
29
30
31 Triplett, J. E. (2004) *Handbook on Hedonic Indexes and Quality Adjustments in*
32 *Price Indexes: Special Application to Information Technology Products*, STI
33 Working Paper 2004/9, OECD, Paris.
34
35
36
37 Waugh, F. V. (1928) Quality Factors Influencing Vegetable Prices, *Journal of Farm*
38 *Economics*, 10(2), 185-96.
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Table 1: Summary Statistics from the Laser Printer Data.

Year	Observations	Price	Speed
1993	32	1072	5.4
1994	21	1009	6.0
1995	22	1003	6.4
1996	28	584	6.4
1997	18	504	6.1
1998	26	861	10.2
1999	66	768	11.1
2000	52	590	10.9
2001	53	416	10.6
2002	25	336	11.8
2003	22	286	14.8
2004	20	253	15.9
Total	385		

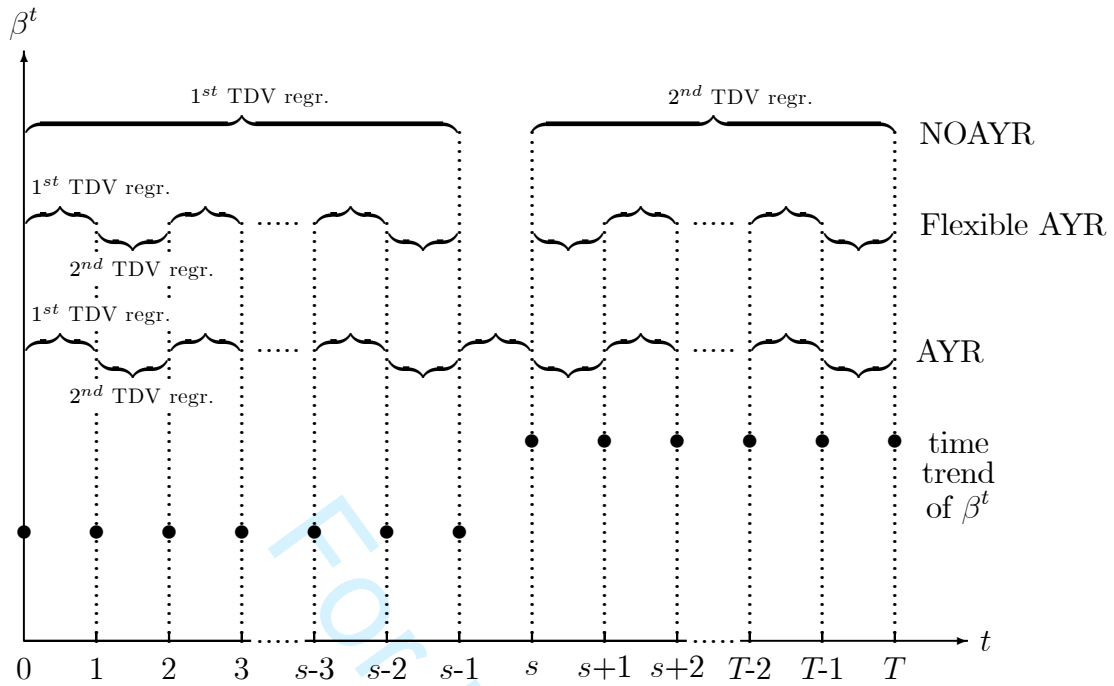


Figure 1: The AYR, Flexible AYR, and NOAYR Methods.

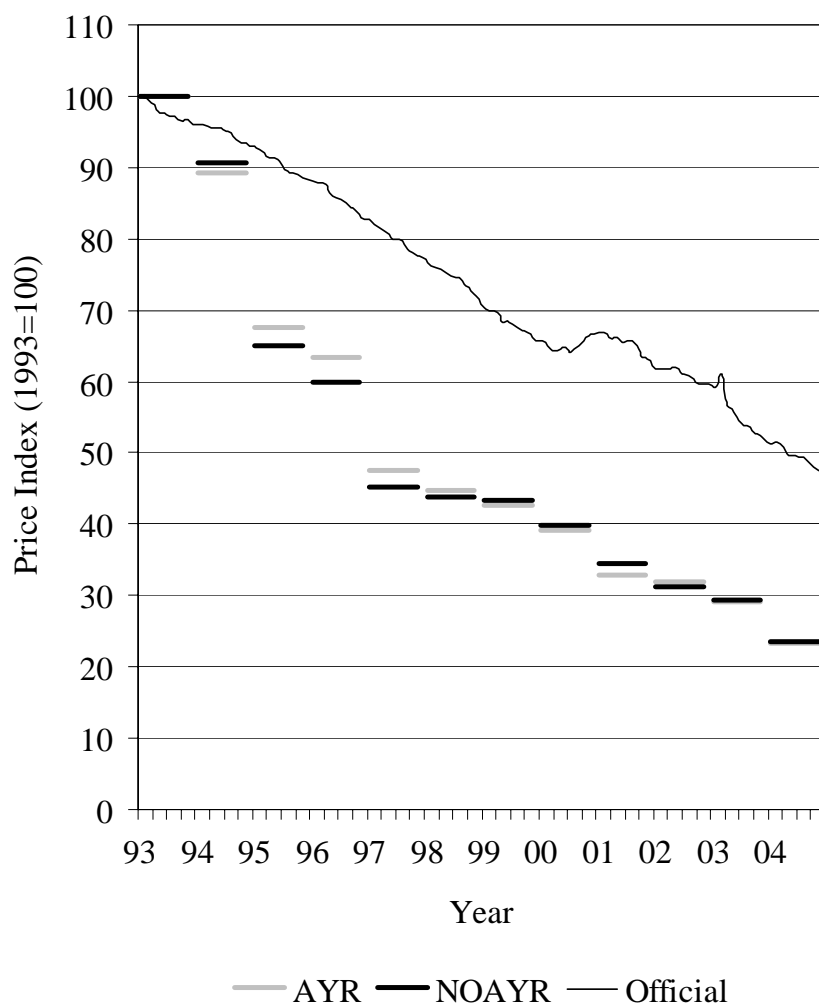


Figure 2: Comparison of Laser Printer Price Trends.