NEW PRODUCTS, QUALITY CHANGES, AND WELFARE MEASURES COMPUTED FROM ESTIMATED DEMAND SYSTEMS

Aviv Nevo*

Abstract—This paper examines the construction of a price index based on an estimated-demand system. In principle the method examined can produce a price index that takes account of the introduction of new products and quality changes in existing products. However, I isolate two key assumptions that have to be made in order to interpret the demand estimates into welfare measures. Using estimates of a brand-level demand system for ready-to-eat cereal, I demonstrate the empirical importance of the assumptions. For the data I use, depending on the interpretation of the demand estimates, a price index can range between a 35% increase over the five years examined to a 2.4% decrease.

I. Introduction

ACCURATE measurement of changes in the cost of living is one of the major tasks of applied economics. The results have implications for a variety of issues such as aggregate growth, industry productivity, real wages, and poverty rates. The Consumer Price Index (CPI) Commission, appointed by the Senate Finance Committee, found that the current CPI overstates the cost of living by about 1.1 percentage points (Boskin et al., 1996). The Commission also pointed to several causes of this bias. Two of the concerns raised, introduction of new products and quality changes (Gordon and Griliches, 1997), have been studied using estimated demand systems to evaluate welfare implications (for example, see Trajtenberg, 1989, 1990; Pakes, Berry, & Levinsohn, 1993; Hausman, 1996, 1997; Petrin, 2002).

This paper attempts to construct a price index that takes account of new-product introductions and quality changes. The basic idea is straightforward. I estimate a demand system, which allows for the inclusion of new products and for quality changes. Next, I use the estimates to construct a measure of changes in consumer welfare, which can be converted to a price index. However, in order to construct the welfare measure I show that one has to make assumptions about the interpretation of the estimation results. In particular, the researcher has to take a stand on the interpretation of (1) time dummy variables, if these are included in the regression (as they usually are), and (2) the error terms. Consider an estimated positive time trend (or an increase in the coefficients of time dummy variables). There are at least two reasons why the demand for a product could increase. The product could have improved, in which case, everything else being equal, consumer welfare has increased. On the other hand, it is possible that the alternatives got worse. Both of these interpretations are consistent with an estimated positive time effect, and if the purpose of the estimation is to compute own- and cross-price elasticities, then separating them is not important. But they have different implications for consumer welfare.

An additional assumption has to be made regarding the residual from the estimation. If the error term represents sampling error, then we want to use the "average" demand curve for the analysis. But in the analysis below I use aggregate data, in which case it is unlikely that the error term is driven by sampling error (Berry, Levinsohn, and Pakes, 1995). Instead the error term includes product characteristics, which are unobserved by the econometrician but valued by consumers. Therefore, the residual should be included in the welfare analysis. But when performing the welfare computation should the residual be allowed to vary between periods, or be held fixed? The answer, once again, depends on interpretation. If a change in the residual reflects a change in tastes (in the sense of Fisher & Shell, 1972), then the answer is no: the residual should not be allowed to vary in the welfare analysis. On the other hand, if the residual captures unobserved quality changes, it should be allowed to vary.

The rest of this paper discusses these two assumptions. In Section 2 I present the problem in the context of a discrete-choice model. It is important to note that the points made below are not specific to this model and are more general. The main claim is that neither of the above interpretations is identified from the price and quantity data used to estimate the demand system. Using data for ready-to-eat cereal and estimates of a demand system (taken from Nevo, 2001), I present, in section 3, the empirical importance of the various assumptions. I show that if one wants to produce a price index that allows for new brands and quality changes, then, depending on which assumptions are chosen, the results range from a 35% increase in the real price of cereal over 5 years to a 2.4% price decrease. For the data used here the assumption regarding the time trend is more important than the assumption on the error term, and explains almost all of this range. This might not always be the case.

Given this wide range, I argue that common practice should include reporting the range. In cases, like here, where the range is wide, I claim that the researcher should be clear about the assumptions made and justify them using prior knowledge. I point to situations where different sets of assumptions are more reasonable.

Received for publication June 12, 2000. Revision accepted for publication January 7, 2002.

* University of California, Berkeley, and NBER.

This paper was motivated by discussions with Zvi Griliches. I am grateful to Lanier Benkard, Ariel Pakes, Amil Petrin, two anonymous reviewers, the editor, and seminar participants at the Berkeley–Stanford IOfest, UC Berkeley, UC Davis, the University of Chicago, the NBER Productivity workshop, and the NBER/CRIW conference for comments and suggestions. I wish to thank Ronald Cotterill, the director of the Food Marketing Policy Center in the University of Connecticut, for providing the price and quantity data used in this study. Financial support from the UC-Berkeley Committee on Research Junior Faculty Fund is gratefully acknowledged.

The Review of Economics and Statistics, May 2003, 85(2): 266–275

© 2003 by the President and Fellows of Harvard College and the Massachusetts Institute of Technology

II. The Model

I demonstrate my point in the context of a random-coefficients discrete-choice model of demand. The problems outlined below are not specific to this model; as I show below, they have parallel in a classical demand model. Furthermore, similar issues arise when computing a price index based on a so-called index formula. I focus on the discrete-choice model for two reasons. First, this model has gained popularity lately and has been used in a variety of situations. Second, the somewhat more structural nature of the model (in particular the error term) allows one to think about the problem more conceptually.

A. *An Empirical Model of Demand*

Suppose we observe markets $t = 1, \ldots, T$, with consumers $i = 1, \ldots, M_t$, who are offered brands $j = 1, \ldots, J_t$. For simplicity we can think of t as indexing time. The conditional indirect utility of consumer i from product j at time t is

$$v_{ij}^t = x_j^t \beta_i^* - \alpha_i^* p_j^t + \xi_j^t + \varepsilon_{ij}^t \equiv V_{ij}^t + \varepsilon_{ij}^t, \tag{1}$$

where x_j^t is a vector of K observable product characteristics, p_j^t is the price of product j at time t, assumed common to all consumers, ξ_j^t is the valuation of unobserved (by the econometrician) product characteristics, and ε_{ij}^t is a mean-zero stochastic term. Finally, $(\alpha_i^* \ \beta_i^*)$ are $K + 1$ individual-specific coefficients.

I model the distribution of consumers taste parameters for the characteristics as

$$\begin{pmatrix} \alpha_i^* \\ \beta_i^* \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} + \Pi D_i + \Sigma \tau_i, \quad \tau_i \sim N(0, I_{K+1}), \tag{2}$$

where D_i is a $d \times 1$ vector of demographic variables, Π is a $(K + 1) \times d$ matrix of coefficients that measure how the taste characteristics vary with demographics, and Σ is a scaling matrix. This specification allows the individual characteristics to consist of demographics that are observed, denoted D_i, and additional characteristics that are unobserved, denoted τ_i, for which a normal distribution is assumed.

In principle, this specification allows for wealth effects, since equation (1) allows the marginal utility from income, α_i, to vary by individual, and equation (2) allows these coefficients to vary with demographics, one of which could be current income. In reality, if the price of the product is small relative to current income, one can assume that these coefficients do not change as a result of price changes. Both in order to simplify the presentation, and since my empirical example examines ready-to-eat cereal, which satisfies the requirement that price is much lower than income, for the rest of the paper I assume no wealth effects.[1]

The specification of the demand system is completed with the introduction of an *outside good;* the consumers may decide not to purchase any of the brands. The indirect utility from this outside option is

$$v_{i0}^t = \xi_0^t + \varepsilon_{i0}^t \equiv V_{i0}^t + \varepsilon_{i0}^t.$$

Consumers are assumed to purchase one unit of the good that gives the highest utility. This implicitly defines the set of unobserved variables that lead to the choice of good j. Given assumptions on the distribution of the unobserved variables the model can be estimated with aggregate data (Berry, 1994; and Berry, Levinsohn, & Pakes, 1995).

Various distributional assumptions on the unobserved variables would lead to different discrete-choice models. For example, assuming that $\alpha_i = \alpha$ and $\beta_i = \beta$ for all i and that ε_{ij}^t is distributed i.i.d. with a Type I extreme value distribution, we get the well-known (multinominal) logit model. The problems with the own- and cross-price elasticities implied by this model have been well documented (for example see McFadden, 1981, or Berry, Levinsohn, & Pakes, 1995). Furthermore, the logit model is also problematic for evaluating the welfare effects of economic changes (Petrin, forthcoming).

B. *Welfare Evaluation of Economic Changes Using Discrete-Choice Models of Demand*[2]

A consumer's well-being will increase as a result of a price change if and only if $u_i(p^t, w) - u_i(p^{t-1}, w) > 0$, where $u_i(\cdot)$ is consumer i's indirect utility function, p^t and p^{t-1} are the prices before and after the change, respectively, and $w > 0$ is the consumer's income. In order to express the welfare change in monetary terms, a particular class of indirect utility functions are used. These are constructed from the expenditure function and are called money metrics. Using an arbitrary price vector, $\bar{p} \gg 0$, define $e(\bar{p}, u_i(p, w))$. This function gives the expenditure the consumer needs to reach the utility level $u_i(p, w)$ when prices are $\bar{p}$. Thus, the consumer is better off if and only if

$$e(\bar{p}, u_i(p^t, w)) - e(\bar{p}, u_i(p^{t-1}, w)) > 0. \tag{3}$$

Furthermore, the left-hand side of equation (3) provides a monetary measure of the change in consumer welfare. This measure is important if we want to aggregate the change in

[1] This assumption does not drive any of the results, and although the details are slightly different, the exercise below can be repeated for products where wealth effects are significant. Note that this assumption still allows income to affect choice as a demographic variable, like age or education.

[2] This section reviews well-known definitions from welfare economics and shows how they apply to the model used here. For further discussion see, for example, Mas-Colell, Whinston, and Green (1995) or Small and Rosen (1981).

welfare over several consumers, or if we want to compare to changes in producers profit.

In principle any price vector $\bar{p}$ can be used to construct the measure in equation (3). A natural choice is p^{t-1}, which leads to the welfare measure suggested by Hicks (1939): the equivalent variation (EV).[3] The EV is the change in consumer wealth that would be equivalent to the change in consumer welfare due to the price change (expressed in monetary terms). Let $u_i^t = u_i(p^t, w)$; then

$$\begin{aligned} \mathrm{EV}_{it} &= e(p^{t-1}, u_i^t) - e(p^{t-1}, u_i^{t-1}) \\ &= e(p^{t-1}, u_i^t) - e(p^t, u_i^t). \end{aligned}$$

This measure will be negative when the consumer's well-being decreases as a result of the change, and positive when it increases. Note that the utility is not held constant between the two periods, but income is, i.e., $w = e(p^{t-1}, u_i^{t-1}) = e(p^t, u_i^t)$. Two ways to think about EV are (1) that it measures the difference in expenditure needed to achieve different utility levels in the two periods facing the initial prices, and (2) that it measures the difference in expenditure required for achieving the second-period utility when faced with the two different price vectors.[4]

As shown by McFadden (1981) and Small and Rosen (1981), for the model presented in the previous section the EV of individual i is given by

$$\mathrm{EV}_{it} = \frac{u_i^t - u_i^{t-1}}{\alpha_i}, \tag{4}$$

where u_i^{t-1} and u_i^t are the (unconditional) indirect utilities in periods $t-1$ and t, i.e., $u_i^t = \max_j v_{ij}^t$. The measure of total consumer welfare is given by

$$\begin{aligned} \mathrm{EV} &\equiv M \int \mathrm{EV}_{it}\, dP^*(D, \tau, \varepsilon) \\ &= M \int \mathrm{EV}_{it}\, dP_D^*(D)\, dP_\tau^*(\tau)\, dP_\varepsilon^*(\varepsilon), \end{aligned} \tag{5}$$

where M is the total mass of consumers and $P^*(\cdot)$ are distribution functions. An assumption of the independence of D, τ, and ε implies the equality.

[3] An alternative choice is p^t, which leads to the compensated variation (CV). Due to the assumption of no wealth effects, these two are equal.

[4] The equivalent variation can also be represented in terms of the compensated (Hicksian) demand curves:

$$\mathrm{EV}_i(p^{t-1}, p^t, w) = \int_{p^t}^{p^{t-1}} h(p, u_i^t)\, dp,$$

where $h(\cdot, \cdot)$ is the compensated demand curve. In words, the EV is equal to the area between p^{t-1} and p^t and is to the left of the compensated demand curve associated with the utility level u^t. Due to the assumed lack of wealth effects, this is equal to the area to the left of the compensated curve associated with the utility level u^{t-1}, as well as to the uncompensated Marshallian consumer surplus.

The integral in equation (5) can be computed in several ways. First, we could simulate draws from the distribution of the unobservable variables ε, v, and D, for each triplet compute the value of the integrand EV_{it}, and compute the (weighted) average of all the draws (Pakes, Berry, & Levinsohn, 1993). Although intuitive, this approach is an inefficient way to compute the value of the integral. If the marginal utility of income is fixed for each individual (i.e., it does not vary as a result of the price change we are evaluating), then an alternative is to use the derivation of McFadden (1981) to integrate analytically the extreme-value distribution of ε, so that the integral in equation (5) is given by

$$M \int \frac{\ln\left[\sum_{j=0}^{J_t} \exp(V_{ij}^t)\right] - \ln\left[\sum_{j=0}^{J_{t-1}} \exp(V_{ij}^{t-1})\right]}{\alpha_i} \times dP_\tau^*(\tau)\, dP_D^*(D). \tag{6}$$

If the marginal utility of income varies with the event that we are measuring, then the computation is more complex (McFadden, 1995).

The EV can be converted into a price index by finding the factor by which we have to multiply all prices in the base period in order to get the same welfare effect as computed by equation (6). Formally, we have to solve for φ_t such that

$$M \int \left[e(p^t, u_i^t) - e(\varphi^t \cdot p^{t-1}, u_i^t)\right] dP_D^*(D)\, dP_\tau^*(\tau) \times dP_\varepsilon^*(\varepsilon) = \mathrm{EV}_t$$

where EV_t is given by equation (6). In general this involves solving a nonlinear equation. Trajtenberg (1990) proposes approximating this somewhat difficult to compute quantity by the value computed in equation (6) divided by the average price level in period $t-1$. Finally, the above discussion was cast in terms of price changes, but it can easily be used to treat other economic changes, for example, introduction of new brands or a change in the quality of existing brands.

There are two other measures of welfare that are commonly used. The first is a price index based on so-called index formulas. The basic idea is to summarize many prices into a single index, which tries to measure the average price increase. Different weights, used to average across products, will yield different indices. For example, the quantities consumed in periods t or $t+1$ are a common choices for weights. Below I present one such price index as a preliminary analysis of the data.[5]

[5] See Diewert (2001) and references therein for various ways of constructing and deriving such an index, the correspondence between these measures and the EV, and various other issues in constructing an index number. It is important to note that the issues raised here are implicitly present in these price indices. The structure of the model used here allows us to discuss the issues.

The second approach is a quantity index. The intuition behind this index is that of revealed preference. For example, given the prices in $t + 1$, is the bundle consumed in t affordable? If so, and if it was not chosen, then the consumer is better off (in terms of revealed preference). This means that one can use the prices in $t + 1$ to average across quantities in both periods to construct a (revealed preference) quantity index. As in the case of a price index, there are many choices for the weights used to average the quantities. Below I present and discuss one such choice.

C. *Two Key Assumptions*

Whether we are interested in the aggregate EV or some other statistic, and whether this statistic is computed analytically or by simulation, there are two assumptions that have to be made: (1) What has happened to the utility from the outside good between the two periods? (2) What is captured by the unobserved characteristics ξ_j^t, and do these characteristics change between the two periods?

When estimating the model previously described, we cannot separately identify a change in the utility from the outside option from a change in the mean utility of the inside goods. For example, we cannot distinguish an increase in the desirability of the inside goods from a decrease in the utility of the outside good. But these two effects have opposite implications for consumer welfare. If the inside goods have improved, then consumer welfare has increased; if the quality of the outside good has deteriorated, then consumer welfare has decreased.

This identification problem is not unique to the discrete-choice model. For example, consider the case where a time dummy variable is included in a demand system. A positive coefficient on this variable implies an outward shift in the demand curve, which can be for either of two reasons: a decrease in the quality of alternative products, which implies a decrease in overall welfare, or an increase in the quality of the product being considered, which implies an increase in overall welfare. If the demand for the alternative commodities is modeled, then the changes in utility from these alternative products will be captured by the change in their demand curves. However, in practice we are never able to include all alternatives in the demand system.

A different problem is that the utility given in equation (1), required to compute the integral in equation (5), includes the unobserved product characteristic, ξ_j^t. By definition the unobserved characteristic captures the effect of all variables that are not included in the model, just like the error term in a classical demand system. Indeed, in the estimation these variables will constitute the error term, and therefore it can be recovered as the residual. If we use data from both periods t and $t - 1$ to estimate the model, then in general we will have two different values of ξ_j^t for products that were present in both periods. Which values of ξ_j^t should we use for the welfare analysis? Should we use the same value in both periods?

The answers depend on what we believe is included in the unobserved components. If these components are merely unobserved product characteristics, and a change in them is a quality change, then we would like to let them vary over time. On the other hand, if changes in these components capture taste changes (in the sense of Fisher & Shell, 1972), we would like to hold them fixed. In reality the unobserved components include both aspects, but which is more important depends on the case being analyzed, and cannot be identified without additional information.

The distinction made here between the time trend and the error terms is somewhat artificial. They both stem from the same problem: we do not know what is the unobserved term. The time trend just removes the trend from the unobserved term, which helps with the estimation, but does not solve the problem of interpretation. The reason I separate the two is to stress that the problem applies to all components of the model that are not fully specified, whether or not they are part of the econometric error term.

III. An Empirical Example

This section demonstrates the application of the methodology presented in the previous section to welfare analysis of price and quality changes and brand introduction in the ready-to-eat (RTE) cereal industry. I use the data described in the appendix to estimate the demand for cereal, using a random-coefficients discrete-choice model. Next, I use the estimated demand system to evaluate the welfare gains in this industry over the period considered.

A. *Trends in the Ready-to-Eat Cereal Industry*

RTE cereal was first introduced roughly 135 years ago. Initially it was mainly used to provide a healthy breakfast at various health resorts (the most well-known was the one run by Dr. John Harvey Kellogg). Before long the popularity of breakfast cereal spread beyond this limited use. Since World War II, the sales of RTE cereals have grown at a steady 3% annual rate. In 1997 the U.S. market consumed approximately 3 billion pounds of cereal, grossing roughly $9 billion in sales.

The driving force behind this growth was the successful introduction of new products. The number of new brands introduced by the top six manufacturers (Kellogg, General Mills, Post, Quaker, Ralston, and Nabisco) has increased dramatically over time. In the peak year, 1988, more brands were introduced than in the decade of 1950–1959. Starting from 1989 the level of introductions declined and matched the levels of the early 1980s.

During the period 1988–1992 of increased brand introductions, the price of cereal rose significantly. Table 1 presents various price indexes. The first two columns present the Consumer Price Index (CPI) and the change in the price of food and beverages. Both series are taken from the BLS. The next six columns present price indices com-

TABLE 1.—PRICE INDICES (88Q1 = 100)

Year and Quarter	CPI	Food & Beverage	24 Cereal Brands	K Corn Flakes	K Frosted Flakes	GM Cheerios	Post Raisin Bran	N Shredded Wheat
88Q2	101.3	101.1	101.3	102.8	101.5	101.3	100.6	101.5
88Q4	103.7	103.8	107.5	109.5	107.2	109.6	107.2	107.8
89Q2	106.7	107.7	112.3	115.1	113.2	112.0	107.2	113.9
89Q4	108.5	109.4	117.5	114.7	118.3	119.7	116.4	113.3
90Q2	111.4	113.2	120.5	119.9	119.7	125.5	119.6	118.1
90Q4	115.3	115.5	125.1	118.4	122.5	130.3	116.2	124.3
91Q2	116.9	118.6	127.6	119.5	123.8	130.1	121.1	132.0
91Q4	118.8	118.2	129.2	118.0	119.0	133.3	121.3	147.1
92Q2	120.4	119.4	128.8	110.5	117.4	138.3	123.3	154.1
92Q4	122.4	120.1	132.4	121.5	122.9	137.9	124.8	160.3

Source: IRI Infoscan Data Base, University of Connecticut, Food Marketing Center, and BLS series.

TABLE 2.—QUANTITY INDICES (88Q1 = 100)

Year and Quarter	24 Cereal Brands	K Corn Flakes	K Frosted Flakes	GM Cheerios	Post Raisin Bran	N Shredded Wheat
88Q2	99.9	98.8	95.3	92.8	99.7	106.3
88Q4	88.4	84.1	81.9	90.9	71.2	92.5
89Q2	95.8	96.2	97.7	108.0	78.4	85.5
89Q4	88.5	94.1	83.2	98.5	57.7	73.7
90Q2	99.7	97.4	94.0	106.5	59.7	78.2
90Q4	95.9	106.3	101.0	95.8	70.1	73.0
91Q2	103.0	111.2	93.2	116.0	74.5	72.9
91Q4	98.9	110.3	100.3	103.1	73.9	61.4
92Q2	106.0	120.9	106.2	120.1	77.8	68.1
92Q4	97.1	83.0	89.2	121.8	83.2	71.1

Source: IRI Infoscan Data Base, University of Connecticut, Food Marketing Center.

puted from the data used in the estimation below. The column labeled "24 Cereal Brands" computes a fixed-weights index for the 24 products that are present in all time periods.[6] As mentioned in Section 2.2, this type of price index tries to summarize the "average" price increase. Indeed, the average price went up, and roughly speaking consumers should be worse off.

In order to complete the picture, table 2 presents quantity indices, computed from the data described in the appendix. The indices are computed as the numbers of pounds per capita consumed relative to the number consumed in the first quarter of 1988. As was noted above, these indices could be considered as revealed-preference indices: if consumers consume more cereal, they are better off in a revealed-preference sense. The most striking feature seen when comparing tables 1 and 2 is that despite a nominal price increase, the per capita consumption of cereal increases. For example, in the second quarter of 1989 the weighted-average price of the 24 brands went up, relative to the last quarter of 1988 (112.3 versus 107.5), but at the same time so did the quantity index (95.8 versus 88.4). Consumers might have been consuming more because the quality of cereal had gone up, in which case, despite the higher prices, they were better off. Alternatively, they might have been consuming more because the alternative, the outside good, was worse, in which case they were hurt by the price increase. Depending on the interpretation, we should put more weight on the price index or the quantity index as an indicator of consumer welfare. In a sense, it is the relative importance placed on the revealed-preference quantity index, which states that consumers are roughly no worse off, versus the price index, which states that consumers are strictly worse off, that will explain some of the results we observe below.

[6] I used as weights the base-period quantities. There are numerous other alternatives one could use, but since the indices are given only for descriptive purposes, I do not present these.

B. *The Estimated-Demand System*

The first step of the analysis described in the previous section is estimation of demand. The empirical model described in section IIA was used. The instrumental variables include average regional prices in all quarters and the cost proxies discussed in the appendix. The results presented here are taken from Nevo (2001). See there for a detailed discussion. Here I focus on the main results.[7]

The results of the estimation are presented in table 3. The first column displays the means of the taste parameters, α and β. Generally speaking the coefficients seem to have the

[7] My goal in this section is to demonstrate the empirical importance of the assumptions previously discussed. Therefore, I will not argue here for the validity of the identifying assumptions, the data construction, and the modeling assumptions (see Nevo, 2001, for such arguments). All that I claim here is that the estimated demand system is reasonable, in the sense that using the estimates to demonstrate the effects of the assumptions is meaningful.

TABLE 3.—ESTIMATED-DEMAND PARAMETERS

Variable	Mean β	Standard Deviation σ	Interactions with Demographic Variables			
			Income	Income^2	Age	Child
Price	−27.198	2.453	315.894	−18.200	—	7.634
	(5.248)	(2.978)	(110.385)	(5.914)		(2.238)
Advertising	0.020	—	—	—	—	—
	(0.005)					
Constant	−3.592	0.330	5.482	—	0.204	—
	(0.138)	(0.609)	(1.504)		(0.341)	
Cal. from Fat	1.146	1.624	—	—	—	—
	(0.128)	(2.809)				
Sugar	5.742	1.661	−24.931	—	5.105	—
	(0.581)	(5.866)	(9.167)		(3.418)	
Mushy	−0.565	0.244	1.265	—	0.809	—
	(0.052)	(0.623)	(0.737)		(0.385)	
Fiber	1.627	0.195	—	—	—	−0.110
	(0.263)	(3.541)				(0.0513)
All-family	0.781	0.1330	—	—	—	—
	(0.075)	(1.365)				
Kids	1.021	2.031	—	—	—	—
	(0.168)	(0.448)				
Adults	1.972	0.247	—	—	—	—
	(0.186)	(1.636)				
GMM objective (degrees of freedom)			5.05 (8)			

Based on 27,862 observations. Except where noted, parameters are GMM estimates. All regressions include brand and time dummy variables. Asymptotic robust standard errors are given in parenthesis.

expected signs. The next five columns present the parameters that measure heterogeneity in the population: standard deviations, and interaction with log of income, log of income squared, log of age, and a dummy variable that is equal to 1 if age is under 16. The estimates of standard deviations of the taste parameters are insignificant at conventional significance levels for all characteristics except for the Kids segment dummy variable. Most interactions with demographics are significant. Marginal utility from sugar decreases with income. Marginal valuation of sogginess increases with income. In other words, wealthier (and possibly more health-conscious) consumers are less sensitive to the crispness of a cereal.

The estimated time dummy variables (not displayed in the table) show a positive trend with usually lower demand in the fourth quarter.[8] The interpretation of this trend will be a focus of the next section. An additional focus will be on interpreting the change in the residuals. The average of $\xi_j^t - \xi_j^{t-1}$ is zero with a standard deviation of 0.67 and 90% of the distribution between −1.1 and 1.1. This is roughly of the same order of magnitude as the estimated time dummy variables, so at least in principle it is not clear which of the two assumptions previously discussed will have a greater effect.

Further discussion of the results and their implications, including price elasticities and implied markups, can be found in Nevo (2001).

C. *Welfare Gains*

Equivalent variations are presented in table 4. For each set of assumptions two series of equivalent variations are computed: those using the full model and a BLS-like measure. The columns labeled "Full" are computed by multiplying the population-weighted average of the median EV in each city by 260 million (the U.S. population) and 365 days. The median EV in a city was used so to reduce the impact of outliers.[9] The results were multiplied by the number of days because the "product" is defined as one serving of cereal per capita per day. Therefore, in order to convert the welfare gains into annual welfare gains we need to multiply by the number of choices made, i.e., the number of days. Since this is a per capita measure, it is multiplied by

[8] The time dummy variables could be picking up changes in income over time. These changes should, in principle, be captured by the model, since income can influence choice (see footnote 1). In reality, however, since household data are not observed, this effect can be masked by simulation error. Nevo (2001) went to some length to ensure this was not the case. Here I also examined a specification that included the average income by market. The results are essentially identical.

[9] I also explored eliminating the effects of outliers by taking the average EV of all individuals excluding the top and bottom 5%, or by averaging only over individuals with EV within two standard deviations of the city mean. Admittedly, all these solutions are very ad hoc. As pointed out by Petrin (2002), one has to be sure that the welfare results are not driven by the behavior in the tails. Since this point is somewhat orthogonal to the point I am making here, I focus on measures that are not driven by the behavior in the tails of the distribution, rather than offering a general solution.

TABLE 4.—EQUIVALENT VARIATION

Year and Quarter	Equivalent Variation (million 1988Q1 $/year)							
	Utility from Outside-Good Changes				Utility with Outside Good Fixed			
	ξ_{jt} Change		ξ_{jt} Fixed		ξ_{jt} Change		ξ_{jt} Fixed	
	Full	BLS	Full	BLS	Full	BLS	Full	BLS
88Q2	32	−97	−22	−29	88	−45	26	21
88Q4	−342	−454	−370	−383	−52	−142	−47	−58
89Q2	−884	−1258	−1004	−1408	242	−66	120	−81
89Q4	−862	−1240	−890	−1233	−1	−353	−79	−307
90Q2	−1835	−2121	−1580	−2110	135	−74	320	13
90Q4	−1469	−1741	−1359	−1789	72	−209	110	−123
91Q2	−2234	−2612	−2013	−2508	230	−57	407	64
91Q4	−1785	−2059	−1546	−2056	60	−180	256	−78
92Q2	−1669	−2112	−1319	−1938	97	−143	430	4
92Q4	−1852	−2270	−1518	−2104	112	−160	406	36

All measures are based on the results of table 3. The columns labeled "Full" present results from the model; the columns labeled "BLS" present a BLS-like measure. Different columns make different assumptions on the interpretation of the regression results (see text for details).

TABLE 5.—PRICE INDICES COMPUTED FROM THE MODEL (88Q1 = 100)

Year and Quarter	Utility from Outside Good Changes				Utility from Outside Good Changes with the CPI				Utility from Outside Good Fixed			
	ξ_{jt} Change		ξ_{jt} Fixed		ξ_{jt} Change		ξ_{jt} Fixed		ξ_{jt} Change		ξ_{jt} Fixed	
	Full	BLS	Full	BLS	Full	BLS	Full	BLS	Full	BLS	Full	BLS
88Q2	101.1	101.9	101.4	101.5	100.8	101.6	101.1	101.2	99.5	100.3	99.8	99.9
88Q4	105.8	106.4	105.9	106.0	104.0	104.6	104.0	104.1	100.3	100.8	100.3	100.3
89Q2	112.2	114.5	112.9	115.5	105.2	107.1	106.0	107.2	98.6	100.4	99.3	100.5
89Q4	114.0	116.4	114.1	116.3	108.5	110.8	109.0	110.5	100.0	102.1	100.5	101.8
90Q2	123.3	125.1	121.6	125.0	110.5	111.9	109.3	111.3	99.2	100.4	98.1	99.9
90Q4	125.2	127.0	124.5	127.3	114.9	116.7	114.6	116.2	99.6	101.2	99.4	100.7
91Q2	132.1	134.6	130.6	133.9	115.3	117.3	114.1	116.5	98.7	100.3	97.6	99.6
91Q4	131.1	133.0	129.5	133.0	118.4	120.0	117.0	119.3	99.7	101.0	98.5	100.5
92Q2	132.1	135.2	129.7	134.0	119.8	121.4	117.4	120.4	99.4	100.8	97.5	100.0
92Q4	135.6	138.6	133.2	137.4	121.6	123.6	119.5	122.2	99.4	100.9	97.6	99.8

All measures are based on the results of table 3. The columns labeled "Full" present results from the model; the columns labeled "BLS" present a BLS-like measure. Different columns make different assumptions on the interpretation of the regression results (see text for details).

the U.S. population in order to obtain the aggregate measure.

The columns labeled "BLS" display a BLS-like measure. The BLS price index uses fixed base-year weights. The equivalent in our framework is to assume that the individual probabilities of purchase of each product do not change, but that the conditional indirect utility obtained from a product does (Pakes, Berry, & Levinsohn, 1993). For each individual this quantity is computed by substituting into the integral in equation (6) the difference in utility relative to the base year, rather than the welfare derived from a given choice set. Once again the median for each city was multiplied by the U.S. population and the number of days.

The two measures were computed under different assumptions on what is a change in utility versus what is a change of tastes. As was discussed in Section 2, the "error" in the regression, ξ_j^t, is the perceived quality of the product (that is, the level of utility needed to rationalize the observed purchased quantities). A change in this variable can be considered either a change in tastes or a change in the utility consumers obtain from the product.[10] There is nothing in the data or the basic model to separate these two interpretations. Rather than making additional assumptions, in table 4 we examine the results the two extreme cases. The columns labeled "ξ_j^t Change" assume utility from the product changes, whereas the columns labeled "ξ_j^t Fixed" attribute the change in the perceived quality to a change in tastes and therefore fix the value of ξ_j^t at its base-year value.

The first four columns attribute all the change over time to changes in the valuation of the outside good. As was discussed in Section 2, this is equivalent to not including time dummies in the computation. In contrast, the last four columns attribute the change in the time dummies to increases in the utility of the "inside" products.[11] The actual changes over time are probably a combination of both these effects, but the data cannot separate which part is due to which effect.

Comparing across the different sets of assumptions, the equivalent variations are seen to vary significantly. This is

[10] An alternative is to assume that this captures the effects of variables that influence choice but not utility. This requires a separation between choice and utility. Thus, fundamentally undermining the foundation of the revealed preference approach to consumer welfare.

[11] All the columns are based on the same estimation results, which includes time dummies; the difference is only in the interpretation of the results.

seen even more dramatically when we examine the price index, computed as described in section II and presented in table 5. The estimates range from a 35.6% increase in prices to a 2.4% decrease in prices. This range should not be surprising given the difference between the price and quantity indices presented above. The different sets of assumptions sort out the rise in prices and slight rise (or no change) in quantities by attributing them to different causes. For example, the columns that assume the utility from the outside good is fixed attribute more weight to the revealed-preference quantity index, and (since consumption is roughly constant even though prices increase) consumer welfare has not decreased.

For the data used here the more important of the two assumptions is the one regarding the change in the utility of the outside good (which is just the allocation of the time dummy variables). There is no reason why this should always be the case. In particular there are several features that make this example unique. First, there are few product or quality improvements in ready-to-eat cereal. The picture might be different if we examined other products like computers or automobiles. Second, the analysis here was limited to a small set of well-established products. The relative importance of the two assumptions might differ if we include more new products.

There are a variety of intermediate assumptions one could make for attributing the time effects to changes in the inside goods versus a change in the outside good. One such assumption is used by Pakes, Berry, and Levinsohn (1993), who assume that on average the unobserved quality of continuing products is the same. With a nearly fixed set of products, such as I have here, this is almost equivalent to the columns that allow the utility from the outside good to change (if the set of products were exactly fixed, it would be exactly the same). An alternative assumption is exploited in the middle columns of table 5. Here I assume the utility from the outside good varies with the CPI.[12] For the data used here this yields results similar to splitting the time effect equally between the inside and outside goods.

The columns labeled "BLS" in tables 4 and 5 display, for each set of assumptions, the results from the suggested measure, holding the set of products and the purchasing probabilities fixed at the first of the 1988 values. In general the results in these columns vary like those from the full model, when compared across different assumptions. For a given set of assumptions the difference between the "Full" and "BLS" columns includes the bias due to using a fixed set of weights and the exclusion of new products. The latter seems to have a greater effect. The only new product in the sample is Post Honey Bunches of Oats, which was introduced in the second quarter of 1989. This is also the first point where the two columns start to differ systematically. Note also that this difference might overestimate the bias due to new products because it does not allow for the disappearance of products. In principle one could try to separate the various effects that contribute to this difference by using a model like that suggested by Hausman (1996).

[12] In some sense this is cheating. If the purpose of the analysis is to compute a CPI, then I cannot use it to generate the results.

IV. Discussion

This paper discusses the use of estimated brand-level demand systems to construct a price index. I use a discrete-choice model to point out two key assumptions regarding (1) the interpretation of the changes over time in the average demand for all products considered, and (2) the interpretation of a change in the unobserved components effecting demand for each product. Neither of these points is specific to the discrete-choice model. Data from the RTE cereal industry are used to demonstrate the empirical importance of the assumptions. For this data set the breakdown of the time trend (that is, the average change of all products) was more important than the interpretation of the brand-specific unobserved component.

Where do these results leave us? I hope that it will become common practice at the very least to do two things. First, clearly state the assumptions made and provide reasons why these are the correct assumptions for the case being considered. Unfortunately, I do not believe that the various issues can be resolved on theoretical grounds; therefore, reporting the range of possible results of the assumptions, as presented in tables 4 and 5, would help. Hopefully in some cases the sensitivity to the assumptions will be smaller than the example provided here.

In those cases where the range is not tight, more care should be placed on some of the details. For example, the effect of alternatives (the changes in the outside good) might be reduced by including more products in the analysis. Even if this does not affect the estimated demand system, it could ease the interpretation of the results. Placing more structure on the unobserved components of demand, and their evolution over time, is another way in which additional information can be incorporated into the analysis.

While the results of this paper are mainly negative, they have implications for the direction of future work. The results suggest that certain parts of the model, which can be left unspecified if the main focus is the matrix of cross-price elasticities (or its implications), have to be better characterized for welfare analysis. One such characterization is given by Berry and Pakes (1999), who, rather than leaving the unobserved quality as a residual, model how it evolves. The performance of this new method and its sensitivity to the above assumptions is yet unknown.

As mentioned above, the assumptions discussed here are not unique to the discrete-choice model; they are present in other demand models as well. Furthermore, some of the same issues also arise when computing a price index using an index formula. For example, Triplett (1998) reports that, using data from the Canadian soft-drink market, Shultz

(1994) finds that price indices computed from six different index formulas yield a range of 15.1 to 925.7 over a period of five years. While not directly comparable to the range we find here, it does make the latter seem tight.

A somewhat different use of an economic model to predict the welfare changes due to an economic event has recently been used by Hausman (1996, 1997) and Petrin (2002). They both examine the welfare implications of new brand introductions. The essential idea is to use an economic model of demand and supply to produce the counterfactual of what the market equilibrium would have looked like in the absence of the new product. With this prediction in hand, one could compare consumer welfare in these two outcomes. This approach is valid if the subject of research is an isolated event (e.g., an introduction of a new product, a merger, or a change in regulatory environment) that can be modeled. In such a case this approach partly sidesteps the problems discussed here. Since the purpose of the analysis is to produce a counterfactual for a particular observed outcome, no assumptions have to be made regarding the interpretation of time effects. Although if one would like to link these effects into a price index, the problem is reintroduced. Also, a natural assumption to make with regards to the unobserved characteristics is that they do not vary. This assumption implies that the only reaction to the brand introduction is in prices and not in other decision variables that enter the error term. For example, if a reaction to a brand introduction can be a change in promotion activity or a change in the unobserved characteristics, then this assumption will be violated.

REFERENCES

Berry, S., "Estimating Discrete-Choice Models of Product Differentiation," *Rand Journal of Economics* 25 (1994), 242–262.

Berry, S., J. Levinsohn, and A. Pakes, "Automobile Prices in Market Equilibrium," *Econometrica* 63 (1995), 841–890.

Berry, S., and A. Pakes, "Estimating the Pure Hedonic Discrete Choice Model," Yale University mimeograph (1999).

Boskin, M. J., E. R. Dulberger, R. J. Gordon, Z. Griliches, and D. Jorgenson, "Toward a More Accurate Measure of the Cost of Living," Final Report to the Senate Finance Committee from the Advisory Commission to Study the Consumer Price Index (December 4, 1996).

Diewert, E. W., "The Consumer Price Index and Index Number Theory: A Survey," Department of Economics discussion paper 01-02, University of British Columbia (2001).

Gordon, R. J., and Z. Griliches, "Quality Changes and New Products," *American Economic Review* 87:2 (1997), 84–88.

Fisher, F., and K. Shell, *The Economic Theory of Price Changes: Two Essays on the Effects of Taste Quality and Technological Change* (New York: Academic Press, 1972).

Hausman, J., "Valuation of New Goods under Perfect and Imperfect Competition," in T. Bresnahan and R. Gordon (Eds.), *The Economics of New Goods,* Studies in Income and Wealth Vol. 58 (Chicago: National Bureau of Economic Research, 1996).

——— "Valuing the Effects of Regulation on New Services in Telecommunications," Brookings Papers on Economic Activity, Microeconomics (1997).

Hicks, J. R., *Value and Capital* (Oxford: Clarendon Press, 1939).

Mas-Colell, A., M. Whinston, and J. Green, *Microeconomic Theory* (Oxford University Press, 1995).

McFadden, D., "Econometric Models of Probabilistic Choice," (pp. 198–272), in C. Manski and D. McFadden (Eds.), *Structural Analysis of Discrete Data* (Cambridge, MA: The MIT Press, 1981).

——— "Computing Willingness-to-Pay in Random Utility Models," University of California, Berkeley mimeo (1995), available from http://emlab.berkeley.edu/~mcfadden.

Nevo, A., "A Practitioner's Guide to Estimation of Random Coefficients Logit Models of Demand," *Journal of Economics & Management Strategy* 9:4 (2000), 513–548.

——— "Measuring Market Power in the Ready-to-Eat Cereal Industry," *Econometrica* 69:2 (2001), 307–342.

Pakes, A., S. Berry, and J. Levinsohn, "Applications and Limitations of Some Recent Advances in Empirical Industrial Organization: Price Indexes and the Analysis of Environmental Change," *American Economic Review, Papers and Proceedings* 83 (1993), 240–246.

Petrin, A., "Benefits of New Products," *Journal of Political Economy* 110 (2002), 705–729.

Schultz, B., "Choice of Price Index Formula at the Micro-Aggregation Level: The Canadian Empirical Evidence," in *International Conference on Price Indices: Papers and Final Report,* First Meeting of the International Working Group on Price Indices (1994).

Small, K. A., and H. S. Rosen, "Applied Welfare Analysis with Discrete Choice Models," *Econometrica* 49 (1981), 105–130.

Trajtenberg, M., "The Welfare Analysis of Product Innovations, with an Application to Computed Tomography Scanners," *Journal of Political Economy* 97 (1989), 444–479.

——— *Economic Analysis of Product Innovation: The Case of CT Scanners* (Cambridge, MA: Harvard University Press, 1990).

Triplett, J. E., "Elementary Indexes for a Consumer Price Index," Brookings Institute (1998).

APPENDIX

Data and Estimation

1. Data

Market shares and prices were obtained from the IRI Infoscan Data Base at the University of Connecticut. These data were collected by Information Resources, Inc. (IRI), a marketing firm in Chicago, using scanning devices in a national random sample of supermarkets located in various size metropolitan areas and rural towns. These data are aggregated by brand (for example, different-size boxes are considered one brand), city, and quarter. The data cover up to 65 different cities, and range from the first quarter of 1988 to the last quarter of 1992. The results presented are computed using the 25 brands with the largest national market shares in the last quarter of 1992 (detailed in table A1). For all except one, there are 1124 observations (that is, they are present in all quarters and all cities). The exception is Post Honey Bunches of Oats, which appears in the data only in the second quarter of 1989. The combined city-level market share of the brands in the sample varies between 43% and 62% of the total volume of cereal sold in each city and quarter. Combined national market shares vary between 55% and 60%.

Market shares are defined by converting volume sales into number of servings sold (using the manufacturer's suggested serving size) and dividing by the total potential number of servings in a city in a quarter. This potential was assumed to be one serving per capita per day. The outside-good market share was defined as the difference between 1 and the sum of the observed market shares. A price variable was created by dividing the dollar sales by the number of servings sold, and was deflated using a regional urban-consumer CPI. The dollar sales reflect the price paid by consumers at the cashier, generating an average real per serving transaction price. However, the sales data do not take into account any coupons used after purchase.

The Infoscan data were matched with a few other sources. First, advertising data were taken from the Leading National Advertising data base, which contains quarterly national advertising expenditures by brand

TABLE A1.—BRANDS USED FOR ESTIMATING DEMAND

All Family + Basic Segment	Taste Enhanced Wholesome Segment	Simple Health Nutrition Segment	Kids Segment
K Corn Flakes	K Raisin Bran	K Special K	K Frosted Flakes
K Rice Krispies	K Frosted Mini Wheats	GM Total	K Froot Loops
K Crispix	P Raisin Bran	P Grape Nuts	K Corn Pops
GM Cheerios	P Honey Bunches of Oats	N Shredded Wheat	GM H-N Cheerios
GM Wheaties	GM Raisin Nut		GM KIX
	Q 100% Natural		GM Lucky Charms
			GM Trix
			GM Cinn Toast Crunch
			Q CapN Crunch
			Q Life

K = Kellogg, GM = General Mills, P = Post, Q = Quaker Oats, N = Nabisco.

TABLE A2.—SAMPLE STATISTICS

Description	Mean	Median	Std. Dev.	Min.	Max.
Prices (¢/serving)	19.4	18.9	4.8	7.6	40.9
Advertising (M$/quarter)	3.56	3.04	2.03	0	9.95
Share within Cereal Market (%)	2.2	1.6	1.6	0.1	11.6
Calories	137.6	120	36.32	110	220
Fat calories (/100)	0.124	0.100	0.139	0	0.60
Sodium (% RDA/100)	0.087	0.090	0.042	0	0.150
Fiber (% RDA/100)	0.095	0.050	0.094	0	0.310
Sugar (g/100)	0.084	0.070	0.060	0	0.200
Mushy (=1 if cereal gets soggy in milk)	0.35	—	—	0	1
Serving weight (g)	35.1	30	9.81	25	58
Income ($)	13,083	10,475	11,182	14	275,372
Age (years)	29.99	28	23.14	1	90
Child (=1 if age $<$ 16)	0.23	—	—	0	1

Source: IRI Infoscan Data Base, University of Connecticut, Food Marketing Center, Cereal boxes and samples from the CPS.

collected from 10 media sources.[13] I used only the total of the 10 types of media. Product characteristics were collected in local supermarkets by examining cereal boxes. This implicitly assumes that the characteristics have not changed since 1988. Although this is not exactly true, it seems a reasonable first approximation. Each cereal was classified into "mushy" or not, depending on its sogginess in milk.[14] There might be some measurement error in this classification. Information on the distribution of demographics was obtained by sampling individuals from the March Current Population Survey for each year. Individual income was obtained by dividing household income by household size. Finally, instrumental variables were constructed using two additional data sources. An average of wages paid in the supermarket sector in each city was constructed from the NBER CPS Monthly Earning Extracts. Estimates of city density were taken from the BLS, as were regional price indices.

Summary statistics for all variables are displayed in Table A2.

[13] The sources include: magazines, Sunday magazines, newspapers, outdoor, network television, spot television, syndicated television, cable networks, network radio, and national spot radio.

[14] I wish to thank Sandy Black for suggesting this variable and helping me classify the various brands.

2. Estimation

I estimate the parameters of the model described in section II using the data described in the previous section by following the method of Berry (1994) and Berry, Levinsohn, and Pakes (1995). The key point of the estimation is to exploit a population moment condition that is a product of instrumental variables and a (structural) error term, to form a (nonlinear) GMM estimator (for details see Berry, 1994; Berry, Levinsohn, & Pakes, 1995). The error term is defined as the city-quarter-specific unobserved product characteristics, ξ_j^t (the estimation includes a brand fixed effect, which captures the overall mean of each brand's unobserved component). As instrumental variables I use the average price of the brand in other cities in the same region, as well as city-level cost proxies. These instrumental variables are justified and examined carefully in Nevo (2001). The market-share function is computed by assuming an i.i.d. extreme-value distribution for ε_{ij}^t, a normal distribution with a diagonal variance matrix for τ_i, and the empirical distribution of demographics D_i as approximated by draws from the March CPS for each year. The extreme-value distribution was integrated analytically; the normal and empirical distributions were approximated by 40 individuals per market. For more details on the estimation, including computer code, see Nevo (2000).