Valuing New Goods in the Presence of Complementarities: Online Newspapers

Matthew Gentzkow*
Harvard University
(Preliminary; please do not cite without permission)

November 19, 2002

Abstract

I develop a multiple-discrete choice demand model in which goods may be either substitutes or complements. I estimate the model using a new micro data set on print and online newspaper consumption in Washington DC, and use the results to quantify the effect of introducing an online edition on both consumer and producer surplus. I find that print and online newspapers are very weak substitutes, while two print newspapers with different ideologies are weak complements. Estimated canibalization of print demand by the online product is minimal: for the $30 million of revenue generated by the online edition, the parent company sacrificed only $4.6 million of circulation and advertising revenue from lost print sales. Looking at consumer welfare, I find that the average surplus per-online-reader-per-day is $.27, compared to $.41 per-print-reader, and the total benefit to consumers from the online edition is $32.2 million. Rough estimates of operating costs suggest that barring large future cost reductions or changes in technology, the social benefits of this online newspaper do not substantially exceed its costs.

*I would like to offer a special thanks to Bob Cohen and Jim Collins of Scarborough Research for giving me access to the data for this study. I am also grateful to John Asker, Richard Caves, Raj Chetty, Liran Einav, Yael Hadass, Jorge Huitron, Ulrich Kaiser, David Laibson, Markus Mobius, Julie Mortimer, Al Roth, Minjae Song, and especially to Ariel Pakes, for advice and encouragement. I also thank the Social Science Research Council for financial support.

"Twenty, thirty, at the outside forty years from now, we will look back on the print media the way we look back on travel by horse and carriage."

Dan Okrent, Editor at Large of Time, Inc., speaking in 1999.

1 Introduction

The effect of online news on the fortunes of print media companies has been the subject of much speculation and debate. Some have predicted that dramatic growth of internet news will curtail or even eliminate the market for traditional newspapers and books.[1] Others have disagreed, arguing that consumption of online news does not necessarily crowd out print consumption, and may even complement it.[2] What little evidence there is, however, is fragmentary and difficult to interpret, and the question has usually been argued at the level of competing hypothetical scenarios. Setting aside broad questions about future changes in tastes or technology, much of this debate boils down to a basic economic question: are print and online news products substitutes or complements?

Methods for estimating the welfare effects of a new good are in many respects very well developed. The standard approach has been to estimate the parameters of a discrete-choice demand system, allowing explicitly for heterogeneity among consumers, and then consider the counter-factual

[1] In a 2000 speech, Warren Buffet said, "I love newspapers... But that is not the way the world is going... Newspapers are very threatened by the internet" (USA Today 2000). Dan Okrent said in the same 1999 speech quoted above, "I believe... all forms of print are dead. Finished. Over" (Okrent 1999). Dick Brass, Microsoft's Vice President of Technology Development, predicted that the last print issue of the New York Times will appear in 2018 (Industry Standard 2000). A study of the U.K. market by the firm Screen Digest concludes that the "dramatic decline in consumption of newspapers and books will continue as more consumers switch to electronic media" (Screen Digest 2002).

[2] A widely cited 2000 survey of U.S. online readers by Belden Associates found that 21% reported buying more print copies since they began using the internet, 7% reported buying less, and 72% reported no change. The same survey found that 7% had started a new print subscription since beginning to read online, while only 3% had stopped a print subscription (Newspaper Association of America 2001).

effects of removing the product of interest from the choice set. This technique has been applied successfully to analyze the welfare impacts of, among other things, medical technology (Trajtenberg 1989), computers (Greenstein 1997), minivans (Petrin 2002), direct broadcast satelites (Goolsbee and Petrin 2002), and semiconductors (Song 2002).

All of these models, however, begin with the assumption that the goods in question are substitutes. Standard discrete-choice models specifcy that consumers choose exactly one good from the set available, so the choice of one good must crowd out demand for another. In the aggregate, the degree of substitutability between two goods will be moderated depending on the extent to which they appeal to similar or different consumers, but actual complementarity will be impossible. While the modern empirical demand literature has made great methodological strides, the concept of complementarity–a central issue in earlier generations of demand theory–has been essentially absent.

The goal of this study is to develop an estimable discrete-choice demand model that allows goods to be either substitutes or complements, and then use the model to estimate the effects of an online newspaper on both firm and consumer welfare. The basic approach is quite simple: I extend a standard discrete choice model by defining the choice set to be all possible *bundles* of goods rather than the individual goods themselves. This provides a natural framework in which to estimate the interaction of different goods when consumed simultaneously, as well as the dependence of utility on consumer characteristics.

While the model is theoretically straightforward, its empirical implementation poses some challenges. First, the number of possible bundles increases exponentially in the number of goods, making the number of parameters and the resulting estimation algorithm potentially intractable even for moderately small choice sets. I address this in part by using a flexible parameterization that allows the degree of heterogeneity in the utility of the underlying goods to be controlled separately from the degree of heterogeneity in their interactions. Thus, the rate of growth in the parameter space can be made close to the linear growth of the number of goods, rather than the exponential growth of the number of bundles. I also point out that using simulated maximum likelihood rather than simulated method of moments can limit the estimation cost of the expanding choice set. While in this application it will make sense to consider a choice set of only 3 goods, I believe the model in its present form would be tractable with up to 8 or 9 goods, and I suggest some possible changes that would make it estimable with much larger choice sets.

A second empirical issue is properly accounting for unobserved consumer characteristics. This has been a major focus of the recent demand literature, and I will apply the random coefficients techniques developed by Berry (1993), Berry, Levinsohn, and Pakes (1995), and others. In the model with complementarities, however, I show that such a flexible error structure is doubly important. In the standard framework, we worry that not controlling for unobserved heterogeneity may lead us to incorrect conclusions about the extent to which two goods appeal to similar or different consumers. Here, omitting such heterogeneity may also cause bias in estimating whether goods are substitutes or complements. Both of these things are crucial in obtaining accurate estimates of the cross-price elasticities. Along with the flexible error structure, I improve identification of the unobservables by incorporating both excluded consumer characteristics (that plausibly affect the utility of one good but not another), and a limited amount of panel structure from the data.

I estimate the model using new micro-data on the print and online newspaper readership of 8,032 consumers in Washington DC. I analyze the effect of the major online newspaper, the washingtonpost.com, on demand for the Washington Post print edition, as well as the city's competing daily, the Washington Times. I also use a pricing equation for the two Washington Post products to estiamate their own-price elasticities of demand. This in turn allows me to calculate the benefit of the online edition to consumers.

The first important result is that accounting for both observed and unobserved heterogeneity is crucial to obtaining plausible estimates of the degree of complementarity or substitutability among products. Both reduced form regressions and a structural model without heterogeneity suggest that the washingtonpost.com is a strong complement to the Washington Post. Adding observed consumer characteristics to the structural model reduces the estimated interaction somewhat, combining it with a flexible correlation structure for the unobservables reduces it still further, and adding panel structure from the data changes the sign so that the two products are estimated to be weak substitutes.

In the full model, I find little evidence that the growth of the online edition has caused a substantial reduction in print readership. All three products are nearly independent, with the Post and post.com perhaps weak substitutes, and the Post and Times (print newspapers with sharply differing ideologies) slightly complementary. Thus, of the 132,500 consumers who read the washingtonpost.com but not the print edition, only about 15,000 would switch to the Washington Post if the online edition were not available. Interestingly, because of consumer heterogeneity, this result is not symmetric: removing the Washington Post from the market would *decrease* post.com readership by 22,000, or more than 20%. Overall, the $30.4 million in revenue the post.com generates comes at the cost of $4.6 million in lost print revenue, for a net revenue gain of $25.8 million. The effect of the post.com on the revenue of the Times is negligible.

I find that the surplus of the average consumer who reads the post.com for a day is \$.27, compared to a surplus of \$.41 for the average Post reader and \$.08 for the average Times reader. This translates into a total consumer welfare gain from the post.com of \$32.2 million.

Thus, from a social standpoint, we would view the post.com to be a valuable innovation if its yearly operating costs were less than \$58 million. While costs clearly exceeded this figure at the height of the dot-com boom, there is some evidence that in the long run they may fall below it. However, these results stand in sharp contrast to the estimates for products such as minivans and semiconductors whose consumer benefits appeared to be an order of magnitude greater than their costs. In the Washington DC market, at least, the online newspaper does not provide a clear net gain to society.

2 A First Look at the Problem

2.1 The Newspaper Product

Before turning to the specific market to be analyzed, it will be worth considering in principle why different newspaper products might be either substitutes or complements. In the case of a print and an online edition of the same paper, a common intuition that the products are substitutes comes from the fact that online editions generally contain close to 100% of the content available in the affiliated print edition. Viewing newspapers as a bundle of information, and recognizing that consuming the same information twice should have close to zero value, we might predict that the marginal value of reading the print edition to someone who has already read the online edition should be close to zero.

Offsetting this, however, are a number of factors that make the products differentiated. Most obviously, one or the other medium may be convenient at different times and places–print newspapers at home over breakfast, say, and online newspapers at one's desk at work. The print edition offers particular advantages such as portability, lower eye strain, and the tactile experience of a printed page. Similarly, the online edition offers access to breaking news, continually updated financial information, archived stories, searchable classifieds, and special interactive features. All these factors would tend to weaken the degree of substitutability among the products. If cross-platform advertising is important, if reading a story about a news item in the morning increases the value of hearing breaking updates on the item later in the day, if consumers like to start a story at work and finish it at home, or if they buy the paper to read in depth analysis of the market movements or sports scores they had been tracking online, the two could even be complementary.

Another set of factors come into play when we consider the interaction of two print editions. One one hand, their content is different, weakening the argument for substitutability. Consumers might well gain from reading two newspapers that are particularly strong in different subject areas, or offer different perspectives on the same story. On the other hand, they are not differentiated in terms of convenience, as print and online editions are. We might well expect that a pair of "up-market" and "down-market" newspapers such as the New York Times and Daily News or the Boston Globe and Herald might be either substitutes or complements, but that their interaction would be very weak since they appeal to very different segments of the population. A pair of newspapers that appeal to more similar consumers but have different strengths and ideologies might be more likely to be complements, as reading both could provide a variety of viewpoints on the same issue, but this effect would be small if people have a strong preference for information sources that confirm their pre-existing opinions.

2.2 The Washington DC Market

Washington DC has two major daily newspapers, the Washington Post and the Washington Times. The former is dominant: average daily circulation of the Post was 762,009 in 2001, compared to 102,957 for the Times. Both have online editions. The washingtonpost.com received approximately 370,000 Washington area visitors per day in 2001, as compared to 38,000 per day for the washingtontimes.com.[3]

Figure 1 displays the daily circulation of Washington DC's print and online newspapers since 1961 (with the exception of the times.com for which I have no historical data). The first thing to note is that the rapid increase in post.com readership since its introduction in 1996 has not been accompanied by a significant drop in Post circulation. A simple OLS regression of Post circulation on post.com circulation gives a significantly negative coefficient of -.099 (with a standard error of .029), but suggests that it takes 10 post.com readers to reduce Post circulation by 1. This certainly is not the picture one would expect to see if the post.com were perfectly crowding out Post readership; however, there is no way to tell what changes were taking place in demand or the characteristics of the products over this period. It might well be, for example, that the post.com does crowd out Post readership, but that this was offeset by growth in overall demand or improvement in the Post's quality.

Looking at the relationship among the print newspapers, one notable feature of the graph is the

[3]Print edition circulation is from Audit Bureau of Circulations figures. Online circulation is based on the Scarborough survey used in this study.

increase in Post circulation after the exit of the Washington Star in 1981. Similarly, the purchase of the Washington News by the Washington Star in 1973 led to a clear increase in Star circulation. In both cases, however, the exits led to declines in total circulation, and fewer than half of the readers of the exiting paper apparently switched to one of the remaining papers. In contrast, there is no evidence of a negative relationship between circulation of the Post and the Times. A linear regression of Post circulation on Times circulation actually gives a positive coefficient of 1.28 (with a standard error of .57), and even when a time trend is included in the regression the coefficient remains positive, though insignificant. However, this relationship may be similarly confounded by changes in demand or characteristics of the products over time.

A natural way to disentangle these effects is to turn to micro-data on demand. The empirical analysis in this paper will be based on a single cross-section of 8,039 Washington DC consumers sampled in 2000-2001. One thing that is immediately apparent from this data is that readership of multiple newspaper editions on the same day is quite common. Of the consumers who report reading some print newspaper in the last 24 hours, 34% said they read two or more. Even more striking, of consumers who reported reading some online newspaper in the last 24 hours, 80% reported also reading a print newspaper in the same period. Note, though, that these figures include not only the Washington papers mentioned above, but several smaller suburban papers as well as national papers like the New York Times and Wall Street Journal.

Table 1 presents cross-tabultions of readership for the Post, Times, and post.com specifically. Here, too, we see that multiple readership is important; 38 percent of respondents read one of these 3 products in the last 24 hours, 8 percent read two, and just under half a percent read all three. Furthermore, 12.1% of Post readers read the post.com, as compared to 5.8% of consumers who did not read the Post, and 8.8% of Post readers read the Times, as opposed to 3.3% of those who did not read the Post. The fact that on average reading the Post on a given day makes one more likely to read both the post.com and the Times suggests intuitively that both pairs of products may be complements.

Reflecting on the validity of this conclusion, however, highlights the importance of what will be one of the central empirical issues we grapple with in the analysis below: consumer heterogeneity. Imagine, for example, results similar to Table 1 for consumption of Bordeaux wine and Mercedes cars in a random sample of US consumers. We might well find that the fraction of Mercedes drivers who drink Bordeaux wine is higher than the corresponding fraction of non-Mercedes drivers. We would not, however, conclude that consumption of expensive wine is complementary to driving expensive cars. Obviously, the apparent relationship is driven by consumer characteristics (income, for example) that increase the propensity to consume each of the two goods independently. One possibility, then, is that all we are observing in Table 1 is the effect of some consumer characteristics that are associated with a propensity to read both print and online newspapers. For this reason, a believable answer to the question of how these goods are related will require us to estimate a richer demand model that explicitly controls for both observed and unobserved consumer heterogeneity.

2.3 Reduced Form Results

Before developing the structural demand model, I briefly present some reduced form regression results as a first look at how controlling for heterogeneity may affect the conclusion that the products are complements. The main reason for doing so is that in a simple linear model, I can use instrumental variables techniques to control for unobserved heterogeneity in a way that will not be possible in the full model. I will therefore run two regressions: (1) a linear IV regression of Post consumption on consuption of the post.com, consumption of the Times, and observable characteristics; and (2) an analagous regression where post.com consumption is the dependent variable and Post consumption is on the right-hand side. In the first case, I instrument for post.com consumption with whether or not the consumer has internet access at work, whether she has a high-speed connection, the amont spent on internet purchases in the last month, and the number of non-news tasks for which she uses the internet. In the second, I instrument for Post consumption using whether or not the consumer takes the subway to work (a plausible instrument if it decreases the time cost of newspaper reading but has no effect on utility from the online edition). I do not have any plausible instruments that would determine Times consumption separately from Post consumption, so the Times dummy is left uninstrumented.

The coefficients on the media consumption terms from these regressions are presented in Table 2. Interpreting the magnitude of the coefficients in a linear probability model is problematic, but these regressions do show that the apparent complementarity among the different products does not disappear in this simple specification. Basically, these regressions show that subway riders are more likely to read online newspapers, and that those who have easy access to the internet are more likely to read print newspapers. Both of these facts are fairly striking, and support the conclusion that the products are complementarity.

Of course there are many reasons why these regressions are mis-specified. The linear probability model ignores the limited nature of the dependent variable and therefore is likely to be biased. The marginal effect of the right-hand side media variables may be different for various consumers (though we could in principle allow this by including interaction terms). Finally, since these regressions look at one product conditional on other endogenous consumption decisions, and because they are not specified in terms of utility, they provide no means by which to calculate producer or consumer surplus from market changes. Having obtained a first look at the problem, therefore, I will now turn to the full structural model.

3 Model

3.1 Demand

As mentioned in the introduction, the basic approach will be to extend a standard multinomial choice model by defining choices to be bundles of the underlying goods, rather than the goods themselves. I begin with a population of N consumers whom I will index by $i = \{1, ..., N\}$. There are G underlying goods; I define the choice set to be the power set of these goods, whose 2^G elements will be indexed $j = \{0, ..., J\}$. The good indexed by 0 will be defined to be the empty set, indicating that the consumer chose to consume none of the G goods in the model, and will be referred to as the outside alternative. For simplicity, I will suppose that the next G elements (indexed $j = \{1, ..., G\}$) are the bundles containing only one good, and the remaining $J - G - 1$ elements are bundles with multiple goods.

Once we define choices as bundles, there is some subtlety to the interpretation of the outside good. In a standard discrete-choice model, it is common to interpret the utility of the outside good as the max over all goods not explicitly included in the model. Once we allow for multiple choices, the issue arises how to treat the interaction of the outside goods with inside goods. The natural extension of the standard interpretation is to view the utility of *all* bundles as the max over all outside goods conditional on consuming the goods in the bundle. This is the approach I will adopt here.

Standard demand specifications allow a consumer's utility for a particular choice to depend on the interaction between a vector of consumer characteristics and a vector of product characteristics. Because meaningful observable characteristics of media products are hard to come by, and because the number of goods in this application will be small, I define the product characteristics to be dummy variables for each good, meaning each consumer characteristic is allowed to interact freely with each of the goods (rather than interacting with a smaller-dimensinal space of characteristics). Thus, I will work in "product space" rather than "characteristic space." This is strictly more general than a characteristics specification.

The product dummies interact in the utility of consumer i with a k-dimensional vector of observable consumer characteristics x_i and a G-dimensional vector of unobservable consumer characteristics ν_i. As I will be using repeated observations on the choices of individual consumers over a week, I will subscript time periods by t, with t being a single day. I will assume that the observable and unobservable characteristics are constant over this period, and that all day-to-day variation is captured by a type-I extreme value error, ε_{ijt}. Thus, consumer i's utility for choice j can be written as:

$$\tilde{u}_{ijt} = \alpha_i(y_i - p_j) + \tilde{\delta}_j + \sum_k x_{ik}\tilde{\beta}_{jk} + \nu_{ij} + \varepsilon_{ijt}, \qquad j = 0, ..., G \tag{1}$$

where $\tilde{\delta}_j$ is the mean utility of good j across all consumers, y_i is income, p_j is the price of good j, and α_i is a consumer-specific price-elasticity. The $\alpha_i(y_i - p_j)$ term represents utility from consumption of all other goods. The quasilinear specification means that I abstract from wealth effects–i.e., I assume that price differences among goods are small enough that they do not change the marginal utility of income. This seems a reasonable approximation in a context where the maximum price is 25 cents. The ν_{ij} will be assumed to have a G-dimensional multivariate normal distribution with free correlation across choices; the $(G+1)G/2$ parameters of this distribution will be denoted σ. Aside from the incorporation of bundles into the choice set, this is a standard random-coefficients specification common in the literature following Berry, Levinsohn, and Pakes (1995); as estimated from micro data, it can also be viewed as a mixed multinomial logit in the sense of McFadden and Train (2000).

I will define the utility of a bundle $j > G$ to be the sum of the mean utilities of the underlying goods plus a consumer-specific interaction term. Letting B_j be the set of indices of the goods in bundle j (i.e. $\{B_j = g : \text{choice } j \text{ includes good } g\}$), this is:

$$\tilde{u}_{ijt} = \sum_{g \in B_j} \left[\alpha_i (y_i - p_g) + \tilde{\delta}_g + \sum_k x_{ik} \tilde{\beta}_{gk} + \nu_{ig} \right] + \Gamma_{ij} + \varepsilon_{ijt}, \qquad j > G \tag{2}$$

Since utility is only defined up to a linear transformation, I impose two standard normalizations. First, I fix the scale parameter of the ε_{ijt} distribution at one. Second, I fix the mean utility of the outside alternative at zero. Noting that the price of the outside good is zero, the normalized utilities can be written as:

$$u_{i0t} = \varepsilon_{i0t}, \tag{3}$$

$$u_{ijt} = -\alpha_i p_j + \delta_j + \sum_k x_{ik}\beta_{jk} + \nu_{ij} + \varepsilon_{ijt} \qquad \text{for } 0 < j \leq G \tag{4}$$

$$u_{ijt} = \sum_{g \in B_j} \left[-\alpha_i p_g + \delta_g + \sum_k x_{ik}\beta_{gk} + \nu_{ig} \right] + \Gamma_{ij} + \varepsilon_{ijt} \;\; \text{for } j > G. \tag{5}$$

Here, $\delta_j = \tilde{\delta}_j - \tilde{\delta}_0$, and $\beta_{jk} = \tilde{\beta}_{jk} - \tilde{\beta}_{0k}$.

I will allow the interaction terms to vary with a separate set of consumer characteristics z_i[4]:

$$\Gamma_{ij} = \sum z_{ik}\gamma_{jk} \tag{6}$$

Finally, because the model lacks sufficient price variation to freely estimate heterogeneous price elasticities, I will define α_i to be a constant scaled by the inverse of income:[5]

$$\alpha_i = \frac{\alpha}{y_i}. \tag{7}$$

Letting θ be the stacked vector of all the utility parameters, $[\delta \ \beta \ \gamma \ \sigma]$, and letting $w_i = [x_i \ z_i]$, we can write the mean utility of consumer i for good j (excluding the logit error ε_{ijt}) as $\bar{u}_j(w_i, \nu_i, \theta)$. Integrating out over the distribution of ε_{ijt} then gives the probability of consumer i selecting choice

[4]Note that in the present draft, this heterogeneity in the interaction terms is NOT estimated. I plan to add these results in the next version.

[5]Heterogeneity in α_i is also not included in the present results.

j in any given period:

$$P_j(w_i, \nu_i, \theta) = \frac{\exp[\bar{u}_j(w_i, \nu_i, \theta)]}{1 + \sum_k \exp[\bar{u}_k(w_i, \nu_i, \theta)]}. \tag{8}$$

The expected total demand per day for good g is then simply:

$$D_g(x, \nu, \theta) = \sum_i \sum_{j:g \in B_j} P_j(w_i, \nu_i, \theta). \tag{9}$$

I will also write the econometrician's expectation of these values over the unobservables ν_i as:

$$\bar{P}_{ij}(\theta) = \int P_j(w_i, \nu, \theta)\phi(\nu|\theta)d\nu, \tag{10}$$

$$\bar{D}_g(w, \theta) = \sum_i \sum_{j:g \in B_j} \bar{P}_{ij}(\theta). \tag{11}$$

Here, $\Phi(\nu|\theta)$ is the multivariate normal distribution.

3.2 Substitutes and Complements

In Gentzkow (2002), I provide a detailed discussion of how the parameters of the above model are related to concepts of substitutability and complementarity in classical demand theory.[6] In this section, I will briefly review some of the relevant results, referring the reader to the other paper for proofs.

The most common classical definition of complementarity is based on derivatives of the quantities consumed by an individual consumer with respect to the goods' prices. Strictly speaking, in a discrete-choice random utility model, the cross-price elasticities are zero for all goods and all

[6]The first draft of this paper is still in preparation.

consumers, except at the measure-zero set of ε_{ijt} where the consumer is exactly indifferent between two goods (at such points, the elasticities of the products in question would be infinitite). Thus, in either the uncompensated or the compensated sense, no goods in this model are either substitutes or complements.

For the questions that motivate discrete-choice analysis, however, it is the cross-elasticities of different goods in *aggregate* demand that are most important. Indeed, accurate estimation of these cross-elasticities has been one of the central goals of the modern literature on demand estimation. In the standard discrete-choice framework (where consumers can choose at most one good), this means measuring the average quality of the two goods in question and the extent to which they appeal to different or similar consumers. To see this, recall that if the choice set above were restricted to single goods, a standard result gives the uncompensated cross-price derivative of aggregate demand as:

$$\frac{\partial D_j}{\partial p_k} = \sum_i \alpha_i P_{ij} P_{ik} \quad \text{for } j \neq k, \tag{12}$$

defining $P_{ij} = P_j(w_i, \nu_i, \theta)$. Letting $\hat{P}_j = \frac{1}{N}\sum_i P_{ij}$, and assuming for a moment that the price elasticity is constant with $\alpha_i = \alpha$ for all consumers, we can expand this as:

$$\frac{\partial D_j}{\partial p_k} = N\alpha \left[\hat{P}_j \hat{P}_k + Cov(P_{ij}, P_{ik})\right] \tag{13}$$

where $Cov(P_{ij}, P_{ik})$ is the sample covariance over i. Thus, holding constant the average quality of the goods (as measured by the $\hat{P}$), the cross-price elasticity will become increasingly positive as the covariance of the probabilities increases, and approach zero when as the covariance becomes large and negative. In other words, when the same consumers who have high utility for good j also have high utility for good k, j and k will be strong aggregate substitutes; when the goods appeal to very different groups, they will approach independence.

When we expand the choice set to include bundles of goods, we add a new dimension that is independent of the covariance of the goods' utilities across the population. The Γ_{ij} terms in the model capture the extent to which different goods interact in the utility of a single consumer. A positive Γ_{ij} for the bundle conaining goods g_1 and g_2 means that the added gain in utility from consuming g_2 is higher if g_1 is consumed as well; a negative Γ_{ij} means the added gain is lower.

Although this does not change the fact that cross-price elasticities for a single consumer will be generically zero, it captures a different (and, indeed, older) notion of complementarity. Undergraduates are usually introduced to this idea by Fisher's indifference curve representation of perfect substitutes and complements: at the polar extremes, goods are perfect complements if they are only valuable in combination (L-shaped indifference curves), and perfect substitutes if combining them gives no added value (linear indifference curves). The perfect complements case corresponds in the present model to individual goods having arbitrarily negative mean utilities and the Γ for their combinations approaching infinitity. The perfect substitutes case corresponds to fixed mean utilities for the individual goods and a Γ approaching minus infinitity. The Fisher definition only specifies the polar cases. More generally, what Samuelson (1974) calls the Pareto-Edgeworth definition of complements is precisely the cross-derivative of utility (as opposed to the cross-derivative of demand), caputured in our discrete setting by the Γ's.

The Pareto-Edgeworth definition was abandoned in favor of the usual compensated demand elasticity definition because it relies on cardinal features of the utility function (i.e. the definition is not invariant to non-linear monotone transformations). In a random utility specification, however, we introduce cardinality to the *mean* utilities through the functional form chosen for the random unobservables (though the actual utility of a consumer whose random utility components have been realized is still purely ordinal). Saying that $\Gamma_{ij} > 0$ when j is the bundle with goods g_1 and g_2 is equivalent to the statement: "conditional on all observable characteristics, the probability that i's (normalized) utility for bundle j exceeds 0 is greater than the product of the respective probabilities that the utilities of g_1 alone and g_2 alone exceed 0."

We thus have two meaningful definitions of g_1 being complementary to g_2 in the model described above: (1) a negative cross-price elasticity of aggreagate demand ($\partial D_2/\partial p_1 < 0$); (2) a positive interaction term Γ_j for the bundle containing goods g_1 and g_2. In Gentzkow (2002), I show that these definitions exactly coincide when the two goods are independent in a particular sense from all other goods g_3.

There, I show that the aggregate cross-price elasticity is equal to:

$$\frac{\partial D_1}{\partial p_2} = \sum_i \left[P_i^1 P_i^2 - P_i^{12} P_i^{-12}\right], \tag{14}$$

where P_i^{12}, P_i^1, P_i^2, and P_i^{-12} are the probabilities that consumer i chooses some bundle with both 1 and 2, with 1 but not 2, with 2 but not 1, and with neither 1 nor 2 respectively. Intuitively, an increase in the price of 2 will increase demand for 1 to the extent that it causes consumers who were choosing bundles with only 2 to switch to bundles with only 1 (captured by the term $P_i^1 P_i^2$) and decrease demand for 1 to the extent that it causes consumers who were choosing both 1 and 2 to switch to bundles with neither (captured by the term $P_i^{12}P_i^{-12}$).[7]

As in classical demand theory, interactions can become quite complex and counter-intuitive in sets of more than two goods. However, in the special case where at least one of goods 1 and 2 is independent of all other goods, and we assume that the price elasticity and interaction of 1 and 2 are constant across consumers (i.e. $\alpha_i = \alpha$ and $\Gamma_{ji} = \Gamma_j$ for all i where j is the bundle of 1 and 2), the cross-derivative takes a particularly intuitive form:

$$\frac{\partial D_j}{\partial p_k} = \alpha N(1 - e^{\Gamma_j}) \left[\hat{P}^1 \hat{P}^2 + +cov(P_i^1, P_i^2)\right] \tag{15}$$

In this expression, $\hat{P}^1$ and $\hat{P}^2$ are the population averages of P_i^1 and P_i^2. Thus, the sign of the cross-derivative will depend exactly on the sign of Γ_j, and the two definitions of complementarity exactly coincide. Furthermore, covariance of the utilities of the two goods will increase the intensity of their interaction, making them either stronger substitutes *or* stronger complements. Again, as the covariance becomes large and negative, the goods will be independent. When the price elasticities are heterogeneous and both goods 1 and 2 interact with other goods in the choice set (as will be true in the implementation of the model here), the cross-derivatives are more complex; the basic intuition from Equation **??** will, however, continue to hold.[8]

[7] It can be shown that switches among bundles with both j and k and bundles with only k cancel out in expectation. A change in the price of k causes no movement between bundles with neither j nor k and bundles with only j.

[8] Interactions with third goods create additional "indirect" substitutability or complementarity as analyzed by Ogaki (1990) for classical demand functions. When price elasticities are heterogeneous, either substitutability or complementarity will increase when α_i covaries positively with $\left(P_i^1 P_i^2 - P_i^{12} P_i^{-12}\right)$.

3.3 Supply

One of the drawbacks of using a cross-section of data from a single market is that there is no variation in prices. Even a study looking across multiple markets would have difficulty getting good estimates of price elasticities, because there is a very limited degree of price competition in newspaper markets. Newspaper cover prices tend to move in increments of at least 5 cents, driven by the practical limitations of vending machines and the benefits of easy over-the-counter payment; in a 2001 sample of 1569 US newspapers, 74% charged a price of exactly 50 cents (Editor and Publisher 2001). Similarly, all major online editions with the exception of the Wall Street Journal are provided free. This discreteness and lack of variation would likely make standard methods of instrumenting to separate price elasticities from unobserved quality unusable.

In the present context, I will instead use industry information to approximate advertising revenue and cost parameters, and then use the firm-side pricing equation for the Washington Post to back out the price coefficient. Specifically, I find the unique coefficient that renders the observed cover price of 25 cents optimal (assuming the minimum increment that prices can change is 5 cents).

One distinction that has not arisen yet, but will be important here, is the difference between the readership of the paper and the number of copies sold. Comparing the micro data used in this study (which measures readership) to circulation figures from the Audit Bureau of Circulations suggests that the average issue of the Post is read by 2.2 adults. This seems roughly intuitive, considering the large fraction of papers that are deliverd either to multi-occupant households or to businesses. I will assume that the propensity to share issues, and hence the number of additional copies sold per additional reader, does not vary across different types of consumer.

I suppose that firm costs are made up of a fixed first-copy cost for both the print and web editions. I assume that marginal costs for an additional web reader are zero in a reasonable neighborhood of the observed number of readers, and marginal costs for the print edition are constant. Finally, I will use a highly simplified model for advertising demand, namely a constant revenue-per-reader for both print and online. Putting these together, I can specify profits for the Washington Post Company:

$$\Pi(p) = a_p N_p(p) + (p - c_p) N_p(p)/\lambda + a_w N_w(p) - \Psi \tag{16}$$

where p is the print edition's price; a_p and a_w are advertising revenue per consumer for print and online respectively; $N_p(p)$ and $N_w(p)$ are the number of print and online readers; c_p is the marginal printing and distribution cost per copy; λ is the number of readers per copy; and Ψ is the fixed first copy cost.

For a price of .25 to be optimal, we must have $\Pi^{WP}(.25)$ greater than or equal to both $\Pi^{WP}(.30)$ and $\Pi^{WP}(.20)$. Plugging in the demand estimates plus approximations of the marginal advertising revenue and marginal cost terms, this condition will define a range of permissible values of the price-elasticity scale parameter, α.[9]

[9] In the current draft, I ignore the discreteness of price changes and simply equate the Post Company's marginal cost and marginal revenue.

4 Empirical Specification

4.1 Data

The data for this study come from a survey of 8,039 consumers in the Washington DC Designated Market Area conducted by Scarborough Research between March, 2000, and August, 2001. The Washington DMA includes the District of Columbia, as well as neighboring counties in Virginia, West Virginia, Pennsylvania, and Maryland. The data includes a wide range of individual and household characteristics of the respondents, as well as detailed data on various consumption decisions. Most important for our purposes, these include a complete enumeration of all print newspapers read over the last 24 hours and 7 days, as well as readership of the washingtonpost.com and washingtontimes.com over the same periods.

Because readership of the washingtontimes.com is less than 100,000 per day, however, there are only 68 times.com readers in the sample–too few to estimate the parameters of its utility. I will therefore define the goods in the model to be daily editions of the Washington Post, the Washington Times, and the washingtonpost.com.[10] The outside alternative includes other print and online newspapers, other news sources such as television and radio, and the choice not to consume news at all. As mentioned before, I interpret *all* choices in the model to represent an implicity maximization over these outside goods–the observed choice to read the Post only, for example, includes consumers who read the Post and the New York Times, read the Post and watched TV news, etcetera.

[10]The model is computationally tractable with a choice set up to 8 or 9 underlying goods (or more if one was willing to impose additional restrictions). In later revisions, I may present results for a larger choice set that includes the major national newspapers, though I am not sure that this gives additional results that are of particular interest.

The vector of consumer characteristics that affect utilities for goods, x_i, contains the following variables: sex, age, household income, how long the consumer has lived in the DC area, and dummies for whether the consumer attended college, attended graduate school, is non-white, is a registered Democrat, is a registered Republican, and lives outside of the DC metropolitan area.[11] It also includes the variables used earlier as instruments: whether the consumer takes the subway to work (which should affect utility for the print but not online editions), the number of non-news tasks the consumer uses the internet for, and whether the consumer has internet access at work or at home (the latter affecting online but not print utility).

The vector of characteristics affecting the interactions, z_i, contains the variables for internet access at home and at work, dummies for various tasks the consumer reports performing on the internet, and dummies for readership of different newspaper sections. I did not have a strong intuition for which variables should affect the interactions, but these ones seem to capture (1) the notion that reading the print at home and the online at work would be a particularly useful combination and (2) the sense that certain kinds of news may have especially high complementarities.[12]

While Scarborough attempts to make their sample representative, there is nonetheless significant over-representation of certain groups. Scarborough provide sampling weights, however, that allow me to map behavior in the pool of respondents to expected effects in the entire market. I will use the unweighted data for the estimation, and then add weighting for all results relating to aggregate market effects.

[11] Note that by including income here, I allow for the possibility that income has a differential effect on the utility of the various goods in addition to the $\alpha_i(p_j - y_i)$ term explicitly included in the utility specification.

[12] In the results presented in the current draft, these variables are NOT included. They assume that the Γ_j are equal across consumers. I plan to include heterogeneous interactions in the next revision.

4.2 Panel Structure and Identification

Recall that in the discussion above, I mentioned the potential confounding of unobserved correlation and complementarity in the example of French wine and expensive cars. Intuitively, the model will predict a large number of consumers choosing two goods together when either the Γ_{ij} for bundles containing both goods are high, or when utilities for the two goods are strongly correlated. It is therefore important to consider how the model will separately identify the interaction terms from the covariance matrix of ν_i.

Keane (1992) discusses this question in the context of a multinomial Probit model. He shows that the covariance matrix is formally identified (up to a single normalization) with only data on characteristics and observed choices, so long as there is sufficient variation in the characteristics. While he does not explicitly consider a choice set that includes bundles, there is no reason that the present model cannot be considered as a regular discrete choice model for the purposes of identification. Thus, though I do not prove this formally, we would expect the model based only on one-day choices and a general vector of characteristics to be identified.[13] My own experiments with the data suggest that this is indeed the case.

In the same paper, however, Keane also shows that identification of the unobservables can be quite fragile without some additional structure, and the problems he encounters should only be exacerbated when one tries to estimate complementarity. Keane suggests including variables that can be excluded *a priori* from the utility of one or more goods and shows using monte carlo experiments that the effect of such variables can be quite substantial. Intuitively, this is exactly the logic

[13] I intend to verify this more formally and add some intuition for how exactly the simple model is identified.

that drove the instrumental variables regressions above. Variables such as internet familiarity or subway ridership duplicate the natural experiment of giving consumers an exogenous improvement in the utility of one good. If the goods are complements, this should also increase the likelihood of consuming the other good as well; if they simply have correlated utilities, it should not. While I do not impose the exclusion restrictions a priori, these variables should improve identification.[14]

The second feature of the data I exploit is the fact that choices are observed in both 24-hour and 7-day windows. This allows me to add a small element of panel structure. Since the ν_i are assumed to be constant across the 7-day period, they will imply restrictions on the correlation of choices over time. If ν_{i1} and ν_{i2} covary strongly, we would expect to see choices of good 1 followed by choices of good 2 and vice-versa. This latter would not follow if the goods were complementary, since a high interaction would lead to them being consumed together in the same period, and *decrease* the likelihood of their being consumed sequentially. In the results that follow, I show that the panel data is indeed extremely powerful in separating correlated utilities from complementarity.

A high variance of the components of ν_i should also imply a greater degree of day-to-day consistency in choices than would be implied by observable characteristics alone. While the model should be able to choose this variance to fit an arbitrarily large degree of consistency, I found that the model's fit is substantially improved when I add an additional parameter to directly control the degree of persistence in choices. I assume, therefore, that on a given day the probability that consumer i has a new ε_{ijt} shock (and thus makes a new choice) is π, while with probability $(1-\pi)$

[14]Note that the most natural candidate for excluded variables would be independent price variation, since a change in the price of one good clearly does not change the utility of the other. While the current setting does not provide it, variation in prices or other characteristics may prove important in other settings, and allow complementarity to be estimated in the absence of the kind of panel structure that is exploited here.

i does not draw a new ε_{ijt} and thus repeats the previous day's choice. Thus, π could be be thought of as controlling an additional degree of serial correlation in unobserved utilities. Alternatively, it could be interpreted as determining how many underlying consumption choices the span of a week actually represents.

To incorporate the panel data into the model specification, I will need some additional notation for the probability for a given combination of 1- and 7-day consumption patterns. Note, first, that the space of possible 7-day consumption observations is the same as the 1-day space, namely the J-element power set of the G goods. Observing that consumer i consumed bundle j in the last 7 days means that each good in j was consumed at least once during that period. Ignoring for a moment the question of whether the 1- and 7-day observations are consistent, we will have $(J+1)^2$ possible 1- and 7-day pairs $\{j_1, j_7\}$, which I will index by $r = \{1, ..., (J+1)^2\}$. Let ω denote a vector of sequential choices, where each element represents a bundle of goods: $\omega_t \in \{0, ..., J\}$. Let $I(\omega)$ be the dimension of ω and let $L(\omega) \in \{1, ..., (J+1)^2\}$ be the index of the pair $\{j_1, j_7\}$ that would be recorded as the aggregate choice if the consumer chose ω_1 on the first day, ω_2 on the second day, and so forth. In other words, $L(\omega)$ is the index of the $\{j_1, j_7\}$ such that $\omega_1 = j_1$ and $g \in B_{j_7}$ if and only if $g \in B_{\omega_t}$ for some $1 \leq t \leq I(\omega)$.

Because the data distinguishes between weekday and weekend editions of the newspapers, and I am only using data on the weekday editions, I will assume that the 7-day observation includes at most 5 different independent choices. Thus, I will only consider consumpton vectors ω with $1 \leq I(\omega) \leq 5$. Let $\Omega(r)$ be the set of all such vectors that are consistent with observation r:

$$\Omega(r) = \{\omega : 1 \leq I(\omega) \leq 5 \text{ and } L(\omega) = r\}. \tag{17}$$

I can thus write the probability of observing consumer i choose consumption r as:

$$Q_r(w_i, \nu_i, \theta) = \sum_{\omega \in \Omega(r)} \pi^{I(\omega)-1}(1-\pi)^{5-I(\omega)} \prod_{t=1}^{I(\omega)} P_{\omega_t}(w_i, \nu_i, \theta). \tag{18}$$

The exponents on π and $(1-\pi)$ range from 0 to 4 (rather than 0 to 5) because I assume the data always represents at least one independent choice. The expectation over the unobservables ν_i is defined as before:

$$\bar{Q}_{ir}(\theta) = \int Q_r(w_i, \nu, \theta)\phi(\nu|\theta)d\nu \tag{19}$$

4.3 Estimation

Given $w_i = [x_i \; z_i]$ and the observed choices for each consumer, a natural way to estimate the model parameters, θ, would be to find the value that maximizes the log-likelihood:

$$L(d_i, w_i, \theta) = \sum_i \sum_r d_{ir} \ln \bar{Q}_{ir}(\theta), \tag{20}$$

where d_{ir} is a dummy variable equal to 1 if consumer i chose (combined 1- and 7-day) choice r and 0 otherwise. As usual with a random-coefficients model, however, the integral that defines $\bar{Q}_{ir}$ in Equation **??** does not have a closed-form solution. I will therefore use simulation draws on the distribution of the ν_i to form consistent estimates of these probabilities.The simplest way to do this would be to form simulated probabilities by averaging the conditional probabilities $Q_r(w_i, \nu_i, \theta)$ over S draws from the multivariate normal distribution of ν. In the actual estimation, I use an importance-sampling variant of this simulator to generate approximations $\bar{Q}^S_{ir}$ of the true probabilities. The simulator reweights the normal distribution to increase the likelihood of drawing ν_i for which the observed choices are relatively likely.[15]

An extensive literature considers techniques for constructing simulation estimators in discrete-choice models. The most intuitive approach would be to simply maximize the simulated analogue of Equation ??. This is the Simulated Maximum Likelihood (SML) estimator first applied by Lerman and Manski (1981):

$$\hat{\theta}_{SML} = \arg\max_{\theta} \sum_i \sum_q d_{ir} \ln \bar{Q}^s_{ir}(\theta). \tag{23}$$

An alternative approach which is now widely applied is the Method of Simulated Moments (MSM) proposed by Pakes and Pollard (1989) and McFadden (1989). It uses a transformation of the first-order condition from Equation ?? to generate a moment condition:

$$E\left[Z_i\left(d_i - \bar{Q}^s_i(\theta)\right)\right] = 0 \tag{24}$$

where d_i and $\bar{Q}^s_i(\theta)$ are stacked vectors of all d_{ir} and $\bar{Q}^s_{ir}(\theta)$, and Z_i is a matrix of instrumental

[15] To take importance sampling draws for a particular consumer i, I use a preliminary estimate of the parameters, θ_0, to generate a simulated approximation of the probability $\bar{Q}^0_{ir}(\theta_0)$ for i's observed choice r. Since this is calculated only once for each i, it can be approximated with a large number of simulation draws. I then draw ν_i from the distribution:

$$H(\nu|w_i, r, \theta, \theta_0) = \frac{Q_r(w_i, \nu, \theta_0)\phi(\nu|\theta)}{\bar{Q}^0_{ir}(\theta_0)}. \tag{21}$$

The simulated probabilities are then formed as a weighted average over S draws from $H(\nu|w_i, r, \theta, \theta_0)$:

$$\bar{Q}^s_{ir}(\theta) = \frac{1}{S}\sum_s Q_r(w_i, \nu_s, \theta) W(w_i, \nu_s, \theta, \theta_0) \tag{22}$$

where $W(w_i,\nu_s, \theta, \theta_0) = \bar{Q}^0_{ir}(\theta_0)\phi(\nu|\theta)/Q_r(w_i, \nu, \theta_0)$ Drawing from $H()$ is simple since it is just the distribution of ν conditional on choice r, and draws can thus be taken using an acceptance-rejection method. In a later draft, I will provide estimates of the variance reduction of importance sampling over a crude simulator.

variables with column-dimension equal to the total number of choices r and row-dimension equal to the number of parameters in θ; the identifying assumption is that Z_i is orthogonal to both the simulation error and the unobservables. McFadden (1989) shows that the estimator will be efficient when the instruments are equal to the derivative of the choice probabilities:

$$Z_i^* = \frac{\partial \bar{Q}_i^s(\theta)'}{\partial \theta}. \tag{25}$$

Standard practice is thus to form an approximation of the optimal instruments, Z_i^s, by averaging the derivative at a preliminary parameter estimate $\hat{\theta}_0$ over a large number of simulation draws of ν. Since the number of moment equations will be equal to the number of parameters, we can then define the MSM estimator:

$$\hat{\theta}_{MSM} = \arg\max_{\theta} \; \left[d - \bar{Q}^s(\theta)\right] Z^s Z^{s\prime} \left[d - \bar{Q}^s(\theta)\right], \tag{26}$$

letting variables without i subscripts represent vectors stacked across all observations.

The main attraction of the MSM estimator is the fact that the simulated probabilities $\bar{Q}^s(\theta)$ enter linearly, and the simulation error will therefore cancel out as the summation is taken over larger and larger N. This means that the MSM estimator is consistent as $N \to \infty$ for a fixed number of simulation draws, while SML also requires $S \to \infty$ with $\sqrt{S}/N = O(1)$ for consistency. On the other hand, efficiency requires the optimal instruments to be calculated exactly, and performance of MSM can be quite poor when the preliminary parameter estimate used to generate them is inaccurate (see Gourieroux and Monfort 1996 for a discussion). Furthermore, calculating the sample analogue of the moments in Equation **??** requires simulations of $\bar{Q}_r^s(\theta)$ and its derivatives for all choices r, rather than just the observed choices which are required for SML. The latter is particularly costly in models with panel data, and in my own experiments I find the computation time of MSM to be more than 10 times that of SML.

For the results presented here, I will use SML to obtain a preliminary estimator to calculate the MSM instruments and then use MSM for the final estimator.[16] This is tractable in the current setup because the choice set has only three goods. In applications with larger choice sets, however, I expect estimation will need to procede using SML. In my experience thus far, the bias can be rendered negligible for a reasonable number of simulation draws, and there are tests available to verify that the number of SML draws is sufficient.[17] I use 100 draws per observation for the SML estimates and 5 draws per observation for the MSM.[18]

5 Results

5.1 Parameter Estimates

Table 3 displays MSM estimates for the parameters of the full model. Almost all the coefficients in the utility of the Post and post.com are significant, as are half the coefficients in the utility of the Times. On the whole, the coefficients correspond closely to what we would have expectated, both for readership of print and online newspapers in general, and for the specific case of the Post and

[16]In the SML estimates, I use the first-order bias correction suggested by Gourieroux and Monfort (1996).

[17]See, for example, Hajivassiliou (2000).

[18]In a later draft, I will either increase the number of MSM draws or abandon it in favor of SML. Even starting with a good preliminary estimate, MSM with 5 draws took approximately 20 times as long as SML with 100 draws started from a random value.

the Times.

The coefficients in the utility of the Post are consistent with its reputation as a relatively high-brow, liberal newspaper. Education, sex, and income coefficients are all positive and significant. Considering a consumer with characteristics at the mean of the data, college attendance, graduate school attendance, and an extra $10,000 of household income increase the probability of choosing the Post by 17%, 4%, and 1% respectively, while being male increases the probability by 8% (only the graduate school coefficient misses significance at the 5% level). Being a registered Democrat increases the probability by 6% on the margin, while being a registered Republican has a negative (though insignificant) effect of 3%. Age has a positive impact, with an additional 10 years of age adding 7% to the probability, and living 5 years longer in the DC area adds 2%. The single strongest effect is living outside of the DC metro area, which makes one 44% less likely to read the Post.[19]

The coefficients of post.com utility are of the same sign as the Post coefficients, though of slightly smaller magnitudes. The notable exceptions are age and time lived in DC: adding ten years of age *decreases* the probability of a mean consumer choosing the post.com by .6% and living in DC five years longer decreases the probability by .3% (note that all of the post.com marginal effects are much smaller than for the Post since at the mean the probability of choosing the post.com, 2.5%, is itself much smaller). Both of these effects are consistent with a widely cited belief in the industry that the online edition has the potential to reach out to younger, more mobile consumers who do not fit the profile of the typical Post reader, though the magnitudes suggest that this strategy may

[19] One additional result that is perhaps surprising is that non-whites are significantly more likely to read all three papers. This may be accurate, or else the non-white dummy may be picking up a finer effect of geographic variation than is captured by the non-metro dummy (the fraction of non-whites is substantially higher in the central city). In a later draft, I will attempt to disentangle this and I will discuss it at more length.

have limited effectiveness.

The coefficients on utilty of the Times are also consistent with its reputation as slightly less high-brow and much more conservative than the Post. Most notably, registered Republicans are significantly more likely, and registered Democrats significantly less likely to choose the Times; the marginal effects are +.09% and -.08% respectively, which are quite large viewed relative to the base probability of the mean consumer choosing the Times, which is .08%. While the effects sex and living outside of the DC metro area are similar to those in the Post's utility, education and income are now insignificant, and the sign on the college dummy is actually negative.

Results are mixed on the variables that I hypothesized could be viewed as excluded from the utility of either print or online newspapers. On one hand, the predicted effects on the goods whose utility they were expected to affect are strong and significant. Subway ridership is strongly related to readership of the Post, with an effect of 13% at the margin, and all of the web-related variables strongly predict readership of the post.com. On the other hand, subway ridership has a significantly positive effect on post.com readership, and having internet access at home has a significantly positive effect on reading the Post, both of which effects were expected to be zero. There are a few possible explanations for this. One is simple omitted variables bias. This is certainly possible, though many of the most obvious potentially correlated variables (occupation, income) are already included. However, we might think that unobserved variation in how much time people spend at home could affect both the choice to install internet access and newspaper readership. A second explanation is that the control variables included in γ_j do not pick up all the heterogeneity in Γ_{ij}, and that this heterogeneity is being picked up by the coefficients on "web at home" and "subway." I discuss this below in the section on robustness checks; the results there suggest that this is not the explanation.[20]

A final, and more interesting explanation is that the "web at home" and "subway" coefficients reflect an unmodelled dynamic demand relationship. Specifically, we might hypothesize that reading the Post one day increases the utility of reading the post.com *on subsequent days*, and that a similar effect would flow from the post.com to the Post. This could be because readers become familiar with the format and style of the paper, get to know particular reporters and columnists, or develop brand loyalty, all of which could be transferable between the print and online editions. If this were the case, any characteristic that affects Post readership would have an upwardly biased coefficient in the post.com regression and vice versa. The length of the panel in this data is too short to allow such an effect to be estimated directly. However, the reduced form IV results above suggest that there is some durable complementarity between the Post and post.com. Because, as shown below, the estimates in the full model show that no complementarity exists within the timeframe of a single day, there is at least a suggestion that such dynamic effects may be important.

Table 4 shows the estimated covariance matrix of the unobservables ν_i, along with the sample covariance of the estimated utility from observables. I estimate that the utility of the Post and post.com are more correlated than the observables alone would predict, and that the utilities of both the Post and post.com are less positively correlated than the observables would predict. The table also shows that the variance of the unobservable component of the Times utility is quite large, reflecting the fact that Times consumption is much more consistent over days than the small probabilities from the model without unobservables would allow. Comparing the relative

[20]Note that in the present draft there are no γ_j coefficients estimated, nor is there a robustness section. Both will be included in subsequent versions. Preliminary estimation allowing Γ_{ij} to depend on the subway and web variables, however, suggest that this does not eliminate the counter-intuitive effects.

magnitudes of the observables and unobservable utility components shows that the unobservables play a key role in allowing the model to fit the data. I do not currently have standard errors for the covariance terms, but inspecting the estimates suggests that the variance terms are significantly different from zero, while the co-variance terms generally are not.

The overall fit of the model can be measured in a number of ways. One is to look at the fit of the aggregate predictions. A simulation of choices from the model matches the aggregate shares quite closely: for the 8 different bundles, the simulated share never differs from the actual share by more than .6%, and the overall MSE is $1.8 * 10^{-5}$. We can also look on a more micro level at how well the model is able to fit consumer heterogeneity. One way to see this is to compare the predicted probabilities of a particular choice for consumers who actually consumed that choice to the predicted probabilities for those who did not. Table 5 displays this comparison. Overall, the fit appears quite good, with the predicted probability of choosing a particular choice nearly 10 times greater on average for those who chose it than those who did not.[21]

5.2 Supply Parameters and Price Coefficient

As mentioned earlier, the lack of price variation in the data will prevent me from endogenously estimating the supply-side paramters in Equation **??**. The firm's optimality conditions are very important, however, as they allow me to obtain a rough estimate of the price-elasticity for use in welfare computations.

To calculate the marginal advertising revenue a_p and a_w, I assume that the quantity of adver-

[21]The average here is over choices. Note that the fitted values include both observables and expected values of the unobservables.

tising space both print and online is fixed (at least over small variations in readership), and the print and online advertising markets are competitive with a fixed prices per-reader-per-day. Note that this abstracts from many important features of the advertising market, including differential values for different types of consumers, the extent to which the same reader on two consecutive days is valued differently from two different readers, and possible market power of the Washington Post.

To calculate the marginal advertising revenue from an additional reader per-day, we can thus divide total advertising revenue by the number of daily readers. For 2001, the Washington Post had total print advertising revenue of \$574.3 million dollars, or \$1.57 million per day (Washington Post Company 2001). With a combined daily and Sunday circulation of 1,075,918, this gives $\alpha_p = 1.46$.[22] Online advertising revenue is not made public, but I employ two sources of information to get a ballpark figure. First, total revenue for the Washington Post's online division was \$30.4 million for 2001 (Washington Post Company 2001), of which the majority was apparently advertising revenue for the washingtonpost.com. If this revenue were entirely from post.com advertising, the daily online readership of 370,000 would imply $\alpha_w = .225$, which we can interpret as an upper bound. Second, Competitive Media Research tracks online advertising spending for major websites. While they do not track the post.com, they were willing to give me June 2001 advertising revenue for the New York Times online edition. If we assume that this month was representative, nytimes.com revenue for 2001 was \$32.4 million, which, combined with nytimes.com readership statistics from Media Metrix, yields $\alpha_w = .160$.[23] I will use this second figure as a point estimate, but also check

[22] Note that this needs to be adjusted to take account of Sunday and daily circulation differently.

[23] CMR reports that June, 2001 advertising revenue for nytimes.com was \$2,701,085. Media Metrix reports the number of unique readers of nytimes.com in July, 2001 at 5,034,000. Assuming that the ratio of monthly to daily

the robustness of the results to varying α_w from .1 to .225.[24]

The final parameter we need is the marginal cost of printing and distribution for an additional print copy. According to industry sources, the largest componenet of marginal cost is newsprint. The Washingon Post's newsprint consumption in 2001 was 226,796 metric tons (Editor and Publisher 2001), and average prices were \$625 per metric ton. Using these figures along with Post circulation and average number of pages per issue, I estimate that the newsprint cost was 42 cents per daily issue. Considering that there are additional marginal costs of ink and distribution, I will estimate $c_p = .5$ and check robustness to values ranging from .4 to .6. Combining these figures with the \$.25 cover price of the Washington Post print edition, I estimate that the Washington Post Company's marginal profit per day from an additional print reader is \$1.21.

I estimate the price elasticity directly from the optimality condition of the Washington Post Company: that the profit in Equation # not be higher for increases or decreases in the cover price.[25] From the demand system, I find that the own-price elasticity of the Washington Post is 626,493α, where α is the price coefficient to be estimated. Plugging this into the firm's first-order condition, along with the assumed advertising and cost parameters, I find that we must have $\alpha = -2.98$. This means that a 10 cent change in price has a similar effect on Post utility to being a democrat rather than unregistered, being male rather than female, 10 years of age, and about 40% of the effect of

readership at the nytimes.com was the same as at the washingtonpost.com, this implies a daily nytimes.com readership of 562,883 (this is probably a lower bound, since the daily figure used for the washingtonpost.com only counts readers in the DC metro area, and so overstates the monthly-daily ratio).

[24]Note that the robustness checks are not yet written up in the present draft.

[25]A few notes are in order. First, as mentioned before, in this draft I actually calculate this from the first order condition and not a discrete variation. Second, I will need to be a little more careful about how to treat subscriptions. The current estimates assume that the subscription price per day is equal to the cover price and that it has the same elasticity. In reality, both the price and the elasticity are probably lower (the latter because of the bundling effect), and it's not clear what the ratio is. I do know which consumers are subscribers, so in a future version I may use the optimality condition for single-copy sales only, ignoring possible substitution to or away from subscriptions. In this version, I also do not allow for heterogeneity in α.

having a college education.

5.3 Substitutes or Complements?

One of the major questions this paper seeks to answer is whether the introduction of the post.com has a significant crowding out effect on demand for the Post, and more generally, to what extent all of the media products are either substitutes or complements. The reduced form results in Section II provided some evidence: based on simple specifications, it appears that far from crowding each other out, the Post and post.com are actually strong complements; both the Post and the post.com also appeared to be strong complements to the Times. We can now see how these conclusions change in the full model.

Table 6 shows the estimated interaction terms, Γ_j. As discussed above, these can be viewed as an intuitive measure of the extent to which goods are substitutes or complements, though they do not necessaily have a simple relationship to the cross-price elasticities. Most strikingly, the apparent complementarity between the Post and post.com has disappeared, and the two products are now estimated to be substitutes, with a Γ_j for the Post-post.com bundle of -.4397. On the other hand, the complementarity between both Post editions and the Times remains, their Γ_j being 1.87. However, the relatively large standard errors suggest that neither interaction is signifiantly different from zero (in contrast to the reduced form regressions where the interaction was strongly significant).[26] A conservative conclusion would be that the products are roughly independent, noting that there is some evidence for Post-Times complementarity.

[26]These standard errors are inflated, however, because they are based on MSM estimates with only 5 simulation draws; in future versions, I will either increase the number of draws or report results from SML (where I can use a much larger number of draws).

In Table 7, I report the estimated price derivatives of demand. The own-price derivatives imply that a 1 cent increase in own price would lead to a loss of 19,546 Post readers, 8,350 post.com readers, and 4071 Times readers. This is equivalent to 1.1% of Post circulation, 2.5% of post.com circulation, and 1.7% of Times circulation, and implies own-price elasticities of .29 and .43 for the Post and Times respectively (we cannot calculate an elasticity for the post.com since its price is zero). Turning to the cross-derivatives, we find that their signs follow the pattern of the Γ_j exactly: the Post and post.com are substitutes (with a positive cross-price elasticity) and both the pairs Post-Times and post.com-Times are complements (with a positive cross-price elasticity). Relative to both the own-price derivatives and total circulation, however, the effects are quite small. A 1-cent increase in the price of the post.com leads to an increase in Post readership of 643 (or .04%) and a decrease in Times readership of 501 (or .2%). Thus, once again, the overall conclusion would seem to be that there is no evidence for substantial crowding out, and all three products appear to be roughly independent.

To see what aspects of the full model led the conclusion of the reduced-form analysis to be reversed, Table 8 presents Γ_j terms and cross-price elasticities for a sequence of simpler models building up to the full one. Adding observed characteristics and panel data each reduce the estimated complementarity of all pairs of products. Allowing for unobservable characteristics further reduces the Post-post.com complementarity while increasing the estimated complementarity of the Post editions and the Times. Thus, demand for the newspapers in this data follows a similar pattern to the expensive wine-expensive cars example given earlier: they are frequently consumed together and thus appear to be complements in a simple analysis, but in fact the apparent complementarity is mainly driven by consumer heterogeneity.

5.4 Welfare Analysis: Consumer Surplus

From the estimated demand parameters, it is a fairly straightforward exercise to calculate the welfare effects of an individual product. For consumer surplus we need can calculate the gain in utility terms as:

$$\Delta U_i(j) = E(\hat{u}_i(\bar{u}_{i1}, ..., \bar{u}_{iJ})|d_i) - E(\hat{u}_i(\bar{u}_{i1}, ..., \bar{u}_{ij-1,}\bar{u}_{ij+1,...}\bar{u}_{iJ})|d_i), \tag{27}$$

where $\hat{u}_i$ is the expected per-day utility of consumer i given a set of mean utilities (each of which is a function of the ν_i). This is therefore conditioned on ν_i. A standard result shows that the expected utility in a multinomial logit model is equal to[27]

$$\hat{u}_i(\bar{u}_{i1}, ..., \bar{u}_{iJ}) = \ln\left[\sum_j e^{\bar{u}_{ij}}\right]. \tag{28}$$

The expectations in Equation **??** are integrated over ν_i conditional on observed choices (as opposed to the integrals over ε_{ijt} which, since the errors are independent, are unconditional).We can therefore estimate the above expectations by simulation:

$$E(\hat{u}_i(\bar{u}_{i1}, ..., \bar{u}_{iJ})|d_i) = \sum_{s=1}^{S} \ln\left[\sum_j e^{\bar{u}_{ij}(\nu_{is})}\right] P(\nu_{is}|d_i) \tag{29}$$

[27]See Anderson, de Palma, adn Thisse for a derivation.

where each ν_{is} is an independent draw from the estimated (unconditional) distribution of unobservables, and the conditional probability is:

$$P(\nu_{is}|d_i) = \frac{P_{ij}(\nu_{is})f(\nu_{is})}{\bar{P}_{ij}}. \tag{30}$$

The figures in utils can be converted to dollars by dividing by the price elasticity and then aggregated:

$$\Delta CS(j) = \sum_i \alpha \Delta U_i(j). \tag{31}$$

A number of caveats should be emphasized about the validity of these figures, however. First, we don't actually see prices varying over a large range so we are depending heavily on the assumption of quasi-linearity that allows utils to be converted into dollars at a constant rate. It is variation in consumer characteristics that allows us to map out the demand curve against the scale given by the normalized ε_{ijt} distribution. Second, using a single firm's optimality condition to estimate the price elasticity can be no more than a rough guide to consumers' sensitivity to price, and it depends heavily on the accuracy of the approximations to advertising revenue and marginal costs.

The results of this calculation are presented in the first part of Table 9. We can get some idea of the welfare effects of the post.com independent of the estimated price coefficient by looking at the surplus to post.com readers relative to the surplus of print readers. I find that the average post.com reader's surplus is 66% of the average print reader's surplus. Aggregating over readers, this means that the total consumer welfare gain from the post.com is about 13% of the total print surplus. In dollars, I find that the average Post reader would be willing to pay \$.41 over the \$.25 cover price to read the paper. The average post.com reader would be willing to pay \$.27, and the average Times reader would be willing to pay \$.08. The total consumer surplus gain to the post.com is \$88,174 per day, or \$32.2 million per year.

5.5 Welfare Analysis: Producer Surplus

The calculation of producer surplus is even easier. Given estimates for the advertising and marginal cost parameters, the gain to a given company from the addition of a particular product can be calculated by simulating the change in demand when that product is removed, and calculating the resulting change in profit from Equation **??**.

Thus, I first simulate the changes in demand. The results are presented in Table 10. They reflect the conclusion above that the three products are nearly independent in demand, as the change in demand for remaining products when some product is removed is never more than 15%. The answer to the question of how the introduction of the post.com has affected readership of the print newspapers is: "not much." Of the 329,000 post.com readers, only 15,500 (on net) would switch to the Post if the post.com were removed, and the increase in Post circulation would be only .9%. The effect on Times readership would be even smaller.

One feature that may seem strage at first is that the effects are asymmetric: removing the post.com increases Post readership, but removing the Post *decreases* post.com readership. This reflects an intuitive fact about consumer heterogeneity in the market: those who choose to read only the post.com on average have a high utility for the Post (making it their second choice) whereas those who read only the Post have a relatively low utility for the post.com (meaning they are more likely to switch to the outside alternative).

These demand effects are translated into dollars in Table 11. I estimate that the introduction of the post.com does canibalize Post demand. However, the cost of this canibalization, at \$4.6 million, is substantially less than the \$30.4 million of revenue generated by the post.com, making the net effect on the Washington Post Company a gain of \$25.8 million. The effect on the Washington Times is negligible, with the few customers gained from the post.com-Times complementarity increasing revenues by \$328,000.

Based on these results, can we understand the decision of the Post Company to introduce an online edition? The answer, of course, hinges on the operating costs of the online division. While these numbers are not made public, one report estimated the post.com's operating costs for 1999 at \$90 million (American Spectator 2000). By any measure, costs at this level would appear to be unsustainable for the firm: even if they could price discriminate perfectly and capture all of the surplus from the online edition, they would still be losing roughly \$30 million. Another way to say this is that the post.com would clearly appear to be costing society more than it is worth.

Some anecdotal evidence, however, suggests that the long-run operating costs may be substantially lower than this. Much of this \$90 million expense went to marketing and promotional activities that were viewed as a long-term investment in building a readership base. For many online papers, they also included purchases of dot-com startups, experiments with new technologies, and other costs associated with the technology bubble of the late 1990's. Some evidence for this contention comes from an analysis of the financial reports of New York Times Digital, which does make its costs public. They reported costs of \$57.8 million, \$136.6 million, and \$67.6 million for 1999, 2000, and 2001 respectively (on revenues of \$44 million, \$67 million, and \$60 million). They have also reported a profit for the first three quarters of 2002 (New York Times Company

2001, 2002). While differences in accounting structure make comparing these numbers across firms difficult[28], they do reflect a similar balooning of costs around the end of the dot-com bubble, and suggest that costs for the post.com may similarly come down, at least to the \$60 million level recently achieved by the Times. In order for the online edition to be profitable, however, there must either be a rebound in the online advertising market or a reduction of costs well below \$60 million.

6 Conclusions

The main contribution of this paper has been to estimate a multiple-discrete choice demand model in which goods were allowed to be either substitutes or complements. This demand system was then used to estimate the value of an online newspaper to both consumers and firms.

To the question of whether print media will survive the growth of online news, this data delivers a strong yes. While the print and online editions of the Washington Post are estimated to be substitutes, the interaction is not significantly different from zero, and the estimated number of print readers lost is a small fraction of either the print or the online edition's total readership. At no point in either the reduced form or structural results is their any evidence of strong substitutability.

The more important question, it seems, is whether the online edition itself will survive. This hinges on the future trajectory of the post.com's operating costs, a subject about which little data is available. In light of the figures on both consumer and producer surplus, however, these costs would have to be reduced substantially from their historical levels for the post.com to achieve profitability

[28]See for example online report

(barring a dramatic increase in online advertising demand). This conclusion would remain true even if the post.com were able to begin charging positive prices and capture a significant share of the consumer surplus.

From the perspective of society, the online newspaper as it stands is not a product whose benefits clearly outweigh its costs. It may well be that cost reductions, technology improvements, and growth in advertising will all combine to change this conclusion, and we may view the high costs of the late 1990's as a valuable long-term investment. On the other hand, we may find that many aspects of online newspapers–interactive features, independent reporting staffs and editors, top-quality graphic design–provide benefits insufficient to justify their high costs, and that the online product will have to be reduced to a more stripped down presentation of the print edition's content supplemented by breaking stories from the news wires.

I am hopeful that some of the methodology here may find application in other settings. The question of estimating complementarity is clearly important, for example, in markets for pharmaceuticals, electronics, food products, automobiles, and any setting where goods have a hardware-software component. All of these could potentially be addressed with a direct translation of this framework, though some aspects of the data such as its panel structure may be difficult to replicate elsewhere.

There are also a number of natural extensions. One would be to consider how in other settings the Γ_{ij} interaction terms could be respecified in terms of measurable product characteristics. Another would be to develop a strategy closer to Berry Levinsohn and Pakes (1995) that would allow complementarity to be estimated from macro data.

References

[1] American Spectator, The. 2000. Print discovers web and in doing so, it's finding new ways to lose money. May.

[2] Berry, Steven, James Levinsohn, and Ariel Pakes. 1995. Automobile prices in market equilibrium. *Econometrica.* 63 (4). 841-890.

[3] Berry, Steven T.. 1994. Estimating discrete-choice models of product differentiation. *RAND Journal of Economics.* 25 (2). 242-262.

[4] Editor & Publisher. 2001. *Editor & Publisher International Yearbook.* New York: Editor & Publisher Co..

[5] Gentzkow, Matthew. 2002. A multiple-discrete choice model with complementarity. Mimeo.

[6] Goolsbee, Austan and Amil Petrin. 2002. The consumer gains from direct broadcast satellites and the competition with cable TV. Mimeo.

[7] Gourieroux, Christian and Alain Monfort. 1996. *Simulation-Based Econometric Methods.* New York: Oxford U. Press.

[8] Greenstein, Shane M.. 1997. From superminis to supercomputers: estimating surplus in the computing market. In Timothy F. Bresnahan and Robert J. Gordon, eds.. *The Economics of New Goods.* Chicago and London: U Chicago Press. 329-374.

[9] Hajivassiliou, Vassilis A.. 2000. Some practical issues in maximum simulated likelihood. In Roberto Mariano, Til Schuermann, and Melvyn J. Weeks, eds.. *Simulation-Based Inference in Econometrics.* Cambridge, England: Cambridge U. Press.

[10] Industry Standard, The. 2000. E-book evangelist: Microsoft's brash Dick Brass leads the revolution that could bury dead trees. Article posted online. September 25. *www.thestandard.net.*

[11] Keane, Michael P.. 1992. A note on identification in the multinomial probit model. *Journal of Business & Economic Statistics.* 10 (2). 193-200.

[12] Lerman, Steven R. and Charles F. Manski. 1981. On the use of simulated frequencies to approximate choice probabilities. In Charles F. Manski and Daniel McFadden, eds.. *Structural Analysis of Discrete Data with Econometric Applications.* Cambridge, MA: MIT Press.

[13] McFadden, Daniel. 1989. A method of simulated moments for estimation of discrete response models without numberical integration. *Econometrica.* 57 (5) 995-1026.

[14] McFadden, Daniel and Kenneth Train. 2000. Mixed MNL models for discrete response. *Journal of Applied Economics.* 15. 447-470.

[15] Newspaper Association of America. January, 2002. Yes circulation directors, they're still buying the paper: single-copy purchases up during Belden Q3 survey. Article posted online. *www.naa.org.*

[16] New York Times Company. 2001. *Annual Report 2001.*

[17] New York Times Company. 2002. The New York Times company reports improved results for third quarter 2002. Press release. September 20.

[18] Ogaki, Masao. 1990. The indirect and direct substitution effects. *American Economic Review. 80 (5). 1271-1275.*

[19] Okrent, Dan. 1999. The death of print? Hearst New Media Lecture. Columbia University. Transcript posted online. *www.jrn.columbia.edu.*

[20] Pakes, Ariel and David Pollard. 1989. Simulation and the asymptotics of optimization estimators. *Econometrica.* 57 (5). 1027-1057.

[21] Petrin, Amil. 2002. Quantifying the benefits of new products: the case of the minivan. *Journal of Political Economy.* 110 (4). 705-729.

[22] Samuelson, Paul A.. 1974. Complementarity: an essay on the 40th anniversary of the Hicks-Allen revolution in demand theory. *Journal of Economic Literature.* 12 (4). 1255-1289.

[23] Screen Digest. 2002. Mediaphile 2010: UK media market could be worth 84.26 pounds by 2010. Report.

[24] Song, Minjae. 2002. Price trend and welfare changes in the CPU market. Mimeo.

[25] Trajtenberg, Manuel. 1989. The welfare analysis of product innovations, with an application to computed tomography scanners. *Journal of Political Economy.* 97 (2). 444-479.

[26] USA Today. 2000. Is Buffett too quick to write off newspapers? May 4. Pg. 3B.

[27] USC Annenberg Online Journalism Review. 2002. Newspapers in the digital age: forget the death of newspapers: the men running America's top chains say papers will survive the digital revolution. Article posted online. May 2. *www.ojr.org.*

[28] Washington Post Company. 2001. *Annual Report 2001.*

Figure 1: *Circulation of Newspapers in Washington DC (1961-present)*

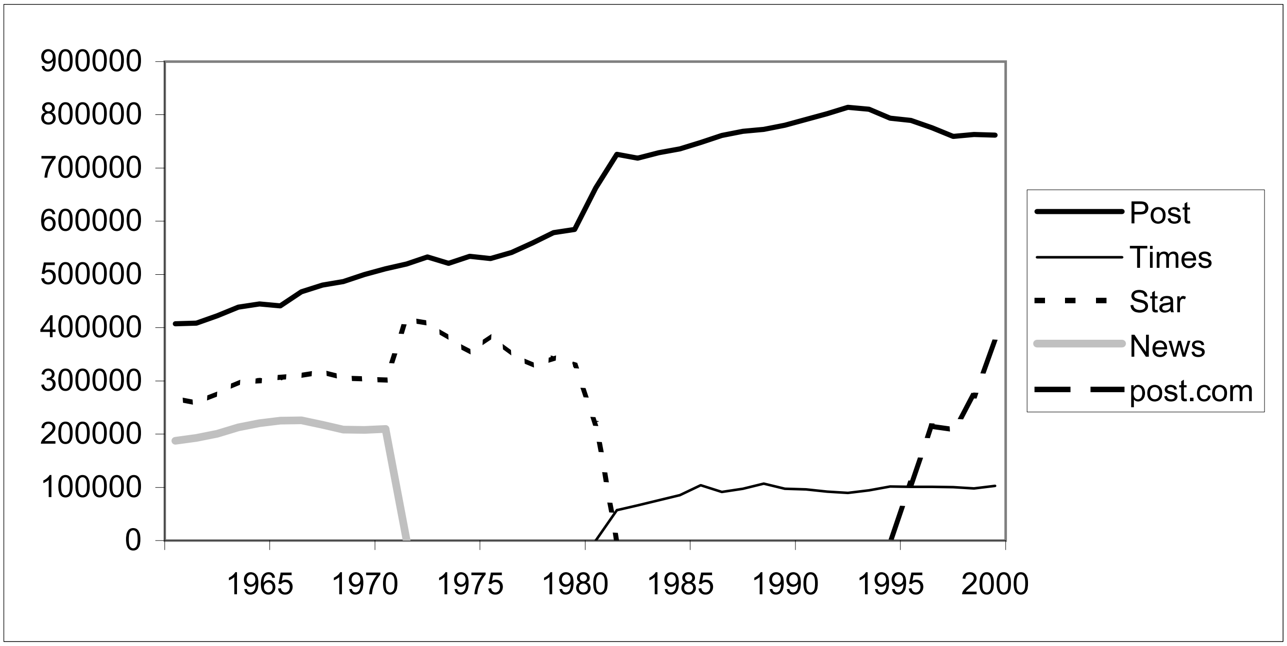

Table 1: *Post, post.com, and Times Readership*

(number of consumers)

		Read post.com:	
		No	*Yes*
Read Post:	*No*	4481	255
	Yes	2932	371

		Read Times	
		No	*Yes*
Read Post:	*No*	4580	156
	Yes	3012	291

Table 2: *Reduced-Form Regression Coefficients*

Dependent Variable: Read Post	
Read post.com	.3284
	(.0557)
Read Times	.3877
	(.0648)
Dependent Variable: Read post.com	
Read Post	.7592
	(.1964)
Read Times	-.0968
	(.0341)
N	8039

Notes: Standard errors in parentheses. Results are from 2SLS regressions with listed RHS variables as well as consumer charactersics listed in Table 3. Instruments for post.com are fast internet connection, web at work, sum of other uses, and amoun spent on internet purchases.

Table 3: *Parameter Estimates (Full Model)*

	W Post	post.com	W Times
Male	**.377**	**.0595**	**.8965**
	(.1202)	(.1238)	(.3814)
Age	**.029**	**-.026**	**.0126**
	(.0104)	(.009)	(.0115)
College	**.7797**	**.6073**	**-.227**
	(.2191)	(.1728)	(.3421)
Grad School	**.171**	**.0314**	**.3438**
	(.0971)	(.1159)	(.3471)
Non-White	**.3646**	**.2716**	**1.2697**
	(.1088)	(.144)	(.521)
Democrat	**.2578**	**.2551**	**-.9328**
	(.0716)	(.1096)	(.3683)
Republican	**-.1272**	**.2132**	**1.1367**
	(.1516)	(.155)	(.4524)
HH Income	**.0604**	**.0058**	**.0391**
	(.0194)	(.0086)	(.0224)
Non-Metro	**-1.9463**	**-.5074**	**-2.1796**
	(.59)	(.1751)	(.9016)
Time Lived in DC	**.0989**	**-.122**	**-.1285**
	(.0458)	(.0669)	(.1533)
Subway	**.5812**	**.3917**	**.1231**
	(.1522)	(.1525)	(.3574)
Web at Work	**.0869**	**.5596**	**.2472**
	(.0883)	(.1636)	(.3016)
Web at Home	**.4741**	**.4219**	**.0394**
	(.151)	(.1444)	(.3422)
Web Uses	**.0162**	**.197**	**.0127**
	(.0219)	(.0612)	(.0859)
Professional	**.1782**	**.4975**	**.0188**
	(.0979)	(.1525)	(.2998)
Constant	**-3.333**	**-3.091**	**-7.3103**
	(.9878)	(.9012)	(2.2909)
Unobs 1	**.73**	**.**	**.**
	(2.864)	(.)	(.)
Unobs 2	**-1.0619**	**-.9659**	**.**
	(.9203)	(1.0817)	(.)
Unobs 3	**-.2117**	**-.4768**	**2.8076**
	(.4957)	(.6226)	(1.2363)
N	8039		

Notes: standard errors in parentheses

Table 4: *Covariance of Consumer Characteristics*

Observable Characteristics

	W Post	*post.com*	*W Times*
W Post	2.093		
post.com	.965	1.767	
W Times	1.397	.777	2.224

Unobservable Characteristics

	W Post	*post.com*	*W Times*
W Post	1.705		
post.com	1.128	1.163	
W Times	-.593	-1.344	7.905

Table 5: *Predicted Probabilities of Each Choice*

	did choose *j*	did not choose *j*
outside good	0.731	0.3456
Post	0.5186	0.2313
post.com	0.0946	0.0322
Times	0.1661	0.0109
Post & post.com	0.2212	0.0304
Post & Times	0.2566	0.0179
post.com & Times	0.0416	0.0033
all 3 goods	0.0998	0.0042

Table 6: *Interaction Terms*

	coeff	std. error
Post & post.com	-.4397	.4323
post.com & Times	1.1518	.782
Post & Times	1.8655	1.4389
Post & post.com & Times	2.0318	1.1397

Table 7: *Cross-Derivatives of Demand (Full Model)*

(Change in demand per $.01 change in price)

	W Post	*post.com*	*W Times*
W Post	-19546.4		
post.com	643.	-8350.2	
W Times	-608.5	-500.8	-4070.6

Table 8: *Cross-Derivatives of Demand (Decomposition)*

No Consumer Heterogeneity, no Dynamics:	
Post-post.com:	-14625.6
Post-Times:	-13453.6
Observables Only, no Dynamics:	
Post-post.com:	-9030.5
Post-Times:	-10541.1
Observables Only, with Dynamics	
Post-post.com:	-4213.2
Post-Times:	-7222.3
Full Model	
Post-post.com:	643.0
Post-Times:	-608.5

Table 9: *Consumer Surplus Gain from post.com*

Per Consumer as % of Print	66%
Per Consumer in $ per day	$0.27
Total as % of Print	13%
Total in $ per day	$88,174
Total in $ per year	$32,183,510

Table 10: *Changes in Circulation*

Take out Washington Post		
Washington Post Readers	1,712,270	
Change in Circ of post.com	-21,823	(-6.6%)
Change in Circ of Times	-34,715	(-14.5%)
Take out washingtonpost.com		
washingtonpost.com Readers	329,045	
Change in Circ of Post	15,558	(.9%)
Change in Circ of Times	-1,112	(-.4%)
Take out Washington Times		
Washington Times Readers	239,380	
Change in Circ of Post	-66,419	(-3.9%)
Change in Circ of post.com	3,641	(1.1%)

Table 11: *Producer Surplus Effects of post.com*

Producer Surplus:	
Δ W. Post Revenue	-$4,586,556
Δ post.com Revenue	$30,400,000
Δ W. Times Revenue	$327,838
Δ Washington Post Co.	$25,813,444