SIEPR Discussion Paper No. 03-18

The Effect of Napster on Recorded Music Sales: Evidence from the Consumer Expenditure Survey

By

Seung-Hyun Hong
Stanford University

January 30, 2004

Stanford Institute for Economic Policy Research
Stanford University
Stanford, CA 94305
(650) 725-1874

The Effect of Napster on Recorded Music Sales: Evidence from the Consumer Expenditure Survey

Seung-Hyun Hong *

30 January 2004

Abstract

This paper quantifies the magnitude of changes in household-level expenditures on recorded music in the United States, particularly attributed to the emergence of Napster. Exploiting the rich information contained in the Consumer Expenditure Survey, I use three approaches to measure the effect of Napster. The difference-in-difference kernel matching (DDM) method directly quantifies the effect. I find that the quarterly music expenditure of the average U.S. household has declined by two dollars and forty-six cents as a result of using the Internet and plausibly Napster. This accounts for 33% of the decrease in total recording sales in 2000. The second approach estimates a demand system for entertainment goods. The estimated cross-price elasticities imply that changes in prices of other entertainment goods also explain the slump in recorded music sales. In 2000, roughly 37% of the decline in recording sales is due to such changes in prices. The final method constructs synthetic cohorts. The results indicate that transition from LPs to CDs might describe the increase in music sales during the 1990's as well as the recent slowdown. These two other methods indirectly measure the effect of Napster in that they explicate that more than 80% of music sales decrease in 2000 might have resulted from factors aside from Napster. This implies that the estimated magnitude using DDM may quantify changes in the household-level music expenditure due to not only Napster but also factors other than file-sharing of copyrighted music.

*Department of Economics, Stanford University. I am grateful to Frank Wolak and Liran Einav for invaluable insights and encouragement. I also thank Susan Athey, Ben Malin, Mark Jacobson, Roger Noll, Azeem Shaikh, Yongs Shin, Ed Vytlacil for helpful discussion and comments. I am responsible for remaining errors.

1. Introduction

This paper quantifies the magnitude of changes in U.S. household-level music expenditure attributed to the emergence of Napster. Napster was introduced in June 1999 and quickly became popular before it was practically closed in early 2001 by the ruling of U.S. Court of Appeals for the Ninth Circuit. Notwithstanding its short life, Napster seems to have facilitated demand for a new medium of recorded music, namely, MP3's and online music services, which are likely to have reshaped demand for existing media of recorded music and have possibly affected music supply. This effect of Napster, particularly on household-level demand for recorded music, is analyzed in this paper. Although I mainly measure the changes in music expenditure due to Napster as well as other possible factors, I argue that the estimated effect on the household-level expenditure essentially quantifies the shift in demand for recorded music ascribed to the introduction of Napster.[1]

In at least two respects, it is important to quantify the extent to which the shift in the household-level demand for recorded music is attributable to Napster. Firstly, it will contribute to the current debate on file-sharing in the recording industry. The emergence of Napster coincided with the recent slowdown in recording sales. The Recording Industry Association of America (RIAA) thus blamed file-sharing for its slump. Though it appears to be plausible, this claim is not substantiated by mere coincidence. It is frequently argued that free downloading of some songs could promote the public awareness of new or unknown music and increase the desire to purchase those albums. Moreover, it is possible that other factors affected the recent downturn in music sales. Consequently, it is desirable to have reliable measure for or against the alleged negative effect of Napster on music sales.

Secondly, evidence for the effect of Napster will enhance our understanding of copyrights. Under copyright protection, distributing or copying songs without permission from copyright holders is prohibited. Whether this protection will encourage consumers to purchase more legitimate copies, however, is not evident.[2] Nonetheless, as Breyer (1970) pointed out in his seminal paper on copyrights, one fundamental premise of copyright protection is that it secures revenues of copyright holders and the related industry by prohibiting unauthorized copies of music. Because this premise underlies the essential economic incentive of copyright protection, it needs to be verified before any further discussion regarding copyrights. Napster weakened the copyright protection by enabling millions of its users[3] to share and download numerous copyrighted songs for free. Therefore, the emergence of Napster provides an unprecedented natural experiment of temporarily loose enforcement of copyrights,

[1] Consumer Price Index for recorded music has not changed much compared to that for other entertainment goods. See section 5.1 for more detail.

[2] Without illegal copies of music, some consumers would purchase legitimate copies, while others could spend less or no dollars on buying music.

[3] According to BBC News (March 26, 2002), there were 70 millions of Napster users at one point.

from which the preceding premise can be verified or refuted.

Accordingly, one might wish to measure the effect of Napster on music revenues from both demand- and supply-side aspects because revenues are determined by changes in demand as well as those in supply. However, I focus on the demand-side effect not only because of its importance, but also because of the lack of supply-side information.[4] Specifically, I attempt to investigate whether household-level demand for recorded music was shifted by the emergence of Napster, or whether the slope of the demand changed after Napster was introduced. Accurate measurement of such changes in demand will inform us of the effect of Napster on revenues for recorded music.

Given available data and methodology, there are several empirical strategies to quantify the magnitude of the change in the household-level music expenditure and ultimately demand that can be attributed to Napster. Section 2 begins by examining aggregate-level changes in recorded music sales from three different sources. Though there are some discrepancies between these sources, the overall trend in recorded music indicates that total recording sales declined around 2000. An aggregate economic downturn, however, does not provide any reliable measure for the effect of Napster. Different consumers have different demand for recorded music, and music is only minor entertainment for many consumers. The impact of Napster on individual consumption, if any, varies substantially for different consumers. Therefore, one should account for the heterogeneity in individual music demand, which requires micro-level data with information on demographics and music consumption. For this reason, I use the *Consumer Expenditure Survey* (CEX) by the U.S. Bureau of Labor Statistics (BLS). The CEX is the unique data set with detailed information on demographics and expenditures. It is publicly available and consists of random samples of households designed to be representative of the total U.S. population.[5]

Using the CEX, section 4 quantifies the effect of Napster on household-level music expenditure. The CEX contains rich data including expenditures on CDs and Internet access, and demographics, such as whether the household is living in a college dormitory. However, it does not include any information about downloading music. Despite the absence of such ideal data, I attempt to identify the effect of Napster by exploiting the CEX and treating the introduction of Napster as a natural experiment. The essential idea of the identification is as follows. I first consider households with Internet access as a treatment group and those without Internet

[4]Nevertheless, there are three reasons for considering only the demand-side effect. First, significant changes in music supply, especially increases in online music services by recording companies had not occurred until the Apple's i-Tunes music stores' success in 2003. Second, detailed information on recording supply is not publicly available. Third, and more importantly, the fundamental premise of copyrights, particularly in the case of Napster, is more related to demand-side changes in the sense that the copying technology was available to individual consumers. In the short-run, the demand of these consumers is more likely to have changed than the supply of recording companies.

[5]There are several consumer-level surveys by a few private research firms. However, most of them are proprietary, and representativeness of their samples is often questionable.

access as a control group. I then suppose that for the treatment group, any change in music expenditure after the introduction of Napster is due to both a time effect and the effect of Napster, whereas such changes for the control group only result from the time effect. Assuming a linear relationship between independent variables, I compute difference-in-difference (DID) estimates for music expenditure between the treatment group and the control group. This will difference out the time effect, so that the resulting DID estimate implies the effect of Napster.

The preceding DID approach, nevertheless, may contain several potential problems. A major problem is that the treatment group may not be comparable to the control group, not only because households in the treatment group tend to be richer than those in the control group but also because the ratio of Internet users to non-users has increased over time. To account for this selection bias, I use a DID kernel matching (DDM) method in a way to make both groups comparable by controlling for the probability of having Internet access. According to the estimation results, the quarterly music expenditure of the average U.S. household has declined by two dollars and forty-six cents as a result of using the Internet and plausibly Napster, which accounts for 33% of the decrease in total recording sales in 2000.[6]

When interpreting the estimated results as the magnitude of negative impact of Napster on music sales, however, one should keep in mind the following caveats. First, the DDM method might not completely remove the bias. Second, not only Napster but also other factors related to using the Internet might have affected music expenditure for the treatment group after the emergence of Napster. Third, other factors such as prices of other goods, which determine music expenditure, may enter demand function nonlinearly, so that the differencing strategy in the DID framework might fail to work.

In this regard, I consider two other methods in the following sections. Both approaches use the CEX and focus on other factors that might have affected recording sales. The idea is that the DID *directly* measures the effect of Napster on music revenues, whereas the other approaches *indirectly* measure the effect in that they explicate how much of the recent slump in the recording industry is *not* due to Napster.

Section 5 estimates a household-level demand system for entertainment goods. As recurrently argued in news articles, consumers might have started spending more dollars and time on other entertainment goods such as movie DVDs and computer games, hence spending less on purchasing music. By estimating a demand system which is consistent with utility maximizing behavior, I recover cross-price elasticities between different entertainment goods, from which I can verify the aforesaid effect of other entertainment goods on music demand. The results from the estimation indicate that CDs and Video are substitutes, whereas CDs and Admission are

[6]See section 4.3 for more detail.

complements.[7] Since 1998, prices for CDs have not changed much while prices for Video have decreased, and those for Admission have increased, which explains why demand for CDs has declined. I find that approximately 37% of the decline in total recording shipments from 1999 to 2000 is due to such changes in prices.

Estimation of the demand system for entertainment goods additionally identifies whether changes in demand are due to a shift in the demand curve or change in its slope. The CEX does not contain detailed data on quantities and prices of CDs purchased. By using detailed *Consumer Price Index* (CPI) data and estimating the demand system on which the restrictions implied by utility maximization are imposed, however, I can recover parameters of an indirect utility function, which enables me to compute own- and cross-price elasticities. I find that own-price elasticities for CDs are similar for both before and after the emergence of Napster. Together with the fact that prices for CDs have not changed much, this implies that Napster might have shifted household-level music demand downwards, rather than changing its slope.

Section 6 examines the possibility of a transition effect on recorded music sales. After CDs became popular in the late 1980's, adult music buyers who used to own many LPs might have purchased a considerable number of CDs during the 1990's, just to update their old LPs. This transition from LPs to CDs could have slowed down around 2000, however, as most adult music buyers would have finished their updating by then. This transition effect could explain the increase in total recording sales during the 1990's and the decrease around 2000. Measuring such an effect, though, requires panel data of the same sample, but the CEX is repeated cross-section data. To overcome this limitation, I construct a synthetic cohort (or pseudo-panel data) using the CEX. The analysis indicates that transition effect also explains the slump in recording sales.

Summarizing the results from the demand system estimation and synthetic cohort approach, I find that more than 80% of recording sales decline in 2000 might be attributable to factors aside from Napster. This implies that the DDM estimate may have quantified the extent to which declines in music expenditure are attributed to both Napster and factors other than Napster.

The following sections detail the aforementioned empirical strategies and present results. Section 2 overviews aggregate-level changes in recording sales. Section 3 then investigates possible reasons for those changes. Section 4 directly quantifies the effect of Napster using the DID method. Section 5 and 6 indirectly measures the effect of Napster by demand system estimation of entertainment goods and a synthetic cohort approach, respectively. Section 7 concludes and suggests directions for future research.

[7] Video means video tapes and disc purchased and rented. Admission includes tickets to movies, plays, opera, and concerts. These goods are defined more precisely in section 5.1.

2. Aggregate-level recorded music sales

According to the RIAA[8], total recording sales have declined since 1999. As shown in figure 1, total real value[9] of shipments in the United States had reached its peak of \$14,268 million in 1999, and then decreased by 4.9 percent in 2000, 6.7 percent in 2001, and 9.6 percent in 2002. This decline, combined sometimes with consumer surveys[10], has often been presented by the recording industry to argue that activities related to file-sharing caused consumers to purchase less music.

Although this line of argument might be believable due to the emergence of Napster in 1999, there are two major problems. First, the current slowdown is not unprecedented. In addition to the data from the RIAA, I consider the measures of aggregate music expenditure from the *Personal Consumption Expenditure* (PCE), an aggregate time-series for U.S. consumer expenditures estimated by the Bureau of Economic Analysis. Though neither series might be precisely computed, in figure 1, both series show that the serious slump in music sales started even before 1999. Neither the slump in early 1980's (in PCE) nor the downturn around 1996 (in RIAA) appears to be less serious than the recent slowdown. However, file-sharing technology did not exist in 1980's and was not popularized before 1999. This implies that there are important reasons other than downloading music that might have caused the slump in the recording industry.

Second, total sales reported by the RIAA do not agree with those from the PCE. This is shown clearly in the figure 1. According to the PCE, declines in recording sales around 2000 are not as apparent as in the RIAA's numbers. There are several reasons for the differences between these two series[11], but I cannot rule out the possibility that either series might be incorrectly estimated.

Such discrepancies become more evident if I compare the two series with the weighted sum of recorded music expenditure from the CEX.[12] Similar differences

[8]Refer to the Yearend Statistics from the RIAA website. RIAA's total sales include CDs, cassettes, LPs, and music video.

[9]All values are in terms of millions of 1996 dollars, adjusted by CPI deflator for all items.

[10]For example, surveys by Peter Hart Research Associates commissioned by the RIAA.

[11]Based on its manual (U.S. NIPA, 1990) and conversation with a staff in BEA, the PCE is first estimated using shipments from Census of Manufacturer by Census Bureau, and then extrapolated using retail sales data. Total value in the PCE is computed using transaction price. On the contrary, the RIAA website states that the figures by the RIAA are compiled by PricewaterhouseCoopers LLP, using data collected from RIAA member companies which account for approximately 84% of music sales, as well as retail sales data from SoundScan for the remainder of the market. The staff in BEA said that the RIAA probably used average price of records to compute total value, instead of using actual transaction price.

[12]The Interview survey in the CEX contains two kinds of expenditures related to recorded music. One is expenditure on CDs, audio tapes, needles, or records purchased other than through a mail-order club. The other is expense on CDs, tapes, videos, or records purchased from a mail-order club like Columbia Record Club. I define the sum of these two expenses as recorded music expenditure. Weighted sum is computed using weights provided by the CEX. These weights are assigned to an address and are the inverse of the probability of selection of the household.

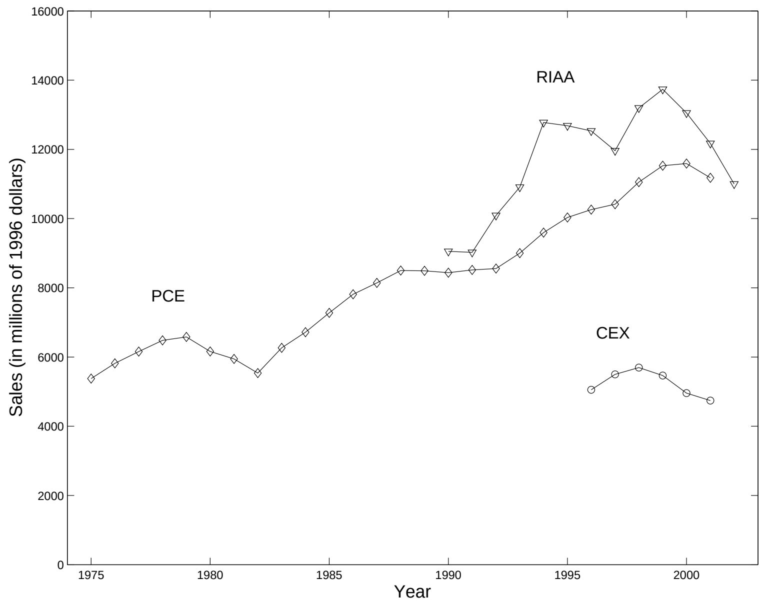

Figure 1: Total sales of recorded music (in millions of 1996 dollars)[a]

[a]Note: RIAA is total value of shipments of recorded music reported by the RIAA; PCE is aggregate expenditure of CDs, Tapes, and LPs in the Personal Consumption Expenditure; CEX is weighted sum of recorded music expenditure in the CEX.

between the CEX and national accounts such as the PCE have been noted by Battistin(2003) and references therein. One possibility is recall problem. That is, survey respondents often forgot their purchases, so that the CEX tends to underestimate the true expenditures. The other is that the CEX surveys only households, so that expenditures from institutions such as government, businesses, libraries, and radio stations are not included in the CEX. These may explain the different levels between the CEX and the other series. However, differences in trend are puzzling.[13]

Though it is unknown which series reflects true trend and level of recorded music sales, the slowdown in recording sales is likely to have occurred around 2000. Due to the following reasons, however, this aggregate trend does not provide any evidence for what caused the slowdown. First, the two preceding problems are not inconsequential. Second, and more importantly, investigation of possible reasons for this slowdown requires accounting for consumer heterogeneity and controlling for different sources of variations affecting consumption. Music consumption varies significantly across the population, and the impact of new technology on individual

[13]According to the CEX, expenditures from mail-order clubs declined significantly in the late 1990's, and RIAA (or PCE) might not include shipments from mail-order clubs, which might explain why the peak occurred in 1998 in the CEX, while it happened in 1999 in RIAA (or 2000 in PCE).

consumption, if any, differs considerably for heterogeneous consumers.

In order to verify any claim related to the effect of Napster on music consumption, therefore, one needs to focus on micro-level data of music expenditure and Internet usage with rich information on heterogeneity of consumers. Although there exist several related sets of survey data, most of them are proprietary and usually contain poor information on demographics. In contrast to these data, the CEX is the only publicly available data which meets the previous condition. Consequently, I analyze the CEX in the following sections.[14]

3. What might have caused the recent slowdown in music sales?

Recorded music sales slowed down around 2000, though the magnitude and length of the slump is still unclear. Napster enabled millions of its users to share and download copyrighted music, thereby loosening the enforcement of copyrights[15], which might then have made the recording industry's revenues less secure. In other words, is the slump due to the impact of Napster? Or are there other explanations which are not related to the fundamental premise of copyright protection?

To answer these questions, consider a slight modification of the standard consumer's problem:

$$\max_{\{x,z\}} \quad U(x_1, x_2, z; A) \qquad \text{s.t.} \quad p_{x_1} \cdot x_1 + p_{x_2} \cdot x_2 + p_z \cdot z \leq M,$$

where x_1 is a vector of recorded music consumption including CD and DM. CD denotes music consumption of existing media including CDs, LPs, and tapes, while DM denotes downloaded music. x_2 is a vector of other entertainment goods such as movies, and z is a vector of all the other goods. p means price vector, and M is total expenditure. Finally, A is a vector of demographic variables. Conditional on DM, the solution to this problem yields the following demand for CD.

$$CD = CD(p_{x_1}, p_{x_2}, p_z, M, A, DM). \tag{1}$$

Individual-level consumption of CD is thus a function of not only downloading but also prices of other goods and demographics.

Equation (1) implies that in order to isolate the effect of Napster on consumption of CD, one needs to account for the effect of consumer heterogeneity reflected in A, as well as the substitution effect of prices of other goods. These other effects have

[14]The possible problem in the CEX as discussed above could indicate that one needs to account for biased sampling. However, I do not take this into account in the present analysis.

[15]Because the RIAA filed a lawsuit against Napster in 1999, the recording industry certainly attempted to enforce copyrights strictly around 2000. Before the ruling of the Ninth Circuit in 2001 February, however, this attempt does not seem to have prevented millions of Napster users from sharing and downloading mostly copyrighted music. Moreover, copyright enforcement then was far less strict compared to the current situation where several hundreds of individual consumers are sued for copyright infringement.

nothing to do with copyright infringement and subsequent enforcement of copyrights. However, changes in prices or demographics might affect recorded music sales. It is presumable that some entertainment goods are substitutes for recorded music, and relative prices of those goods might have decreased in the late 1990's, which is then likely to shift CD demand downwards. It is also plausible that the music taste of a consumer may change over time, and she may even lose her interest in music as she gets older.[16]

In the following sections, I thus consider three approaches to take account of the aforementioned effects on CD. I first investigate the effect of Napster on household-level music expenditure, using a difference-in-difference approach. I then consider two other alternative explanations for the recent slump in music sales. One explanation is associated with prices of other goods. I estimate a household-level demand system for entertainment goods to measure substitutability between different entertainment goods and quantify the declines in music sales due to changes in relative prices. The other is related to demographics. A transition effect, briefly described in the introduction, is examined to measure the effect of demographics on music expenditure.

For a precise measurement of the effect, it would be desirable to identify different effects on CD demand. However, the CEX lacks detailed information on quantities and prices of CD. In consequence, the first and the third approaches quantify the effects on expenditure, but not on demand. The second method, on the other hand, attempts to identify the effect on CD demand by using additional data on prices and a demand system consistent with utility maximizing behavior. Nevertheless, all three approaches inform us of the effect on the recording industry's revenues. Moreover, prices for recorded music seem to have changed little according to detailed monthly CPI data plotted in figure 2 in section 5.1. Accordingly, the estimated effect on the household-level music expenditure may be able to quantify the shift in household-level demand for recorded music attributed to the appearance of Napster.

By separating different effects, therefore, these methods will help answer the preceding questions, eventually providing evidence related to the premise of copyright protection. However, such answers might not be complete in that they explain only the demand-side effect, while ignoring a supply-side effect. It is possible that musicians might have produced less songs, and less music albums might have been released than before. Moreover, it is plausible that used markets for CDs became

[16]Even without any change in p, M, A, and DM at the individual level, changes in aggregate-level demographics may affect recorded music sales. For instance, the following hypothetical situation is possible. Suppose that music consumption of individual consumers changes as they get older, but that individual-level music purchases at each age are the same for different birth cohorts. Some cohorts such as baby boomers include more people, so that aggregate-level music purchases by baby boomers may be larger than those by other cohorts even though average purchases at the same age are the same for different cohorts. It is possible that current young cohorts may include less people than the previous ones, which could explain the recent slowdown in total recording sales.

more popular thanks to the Internet (e.g. half.com or eBay). Taking the supply side into account is thus desirable. Nevertheless, this paper concentrates on the demand side not only because of the lack of supply-side information, but also because of the significance in the demand-side aspect.

4. The effect of Napster: difference-in-difference analysis

This section attempts to directly quantify the effect of Napster on household-level music expenditures. I first describe the essential idea of the difference-in-difference approach. I next describe the data set and consider comparability between the treatment group and the control group. Because of a potential bias explained below, the treatment group may not be comparable to the control group. Consequently, I use difference-in-difference kernel matching to account for this bias. I then report estimation results and discuss their implications. Finally, I present results from a robustness check and consider other potential biases.

4.1. Difference-in-difference approach

There are at least three reasons to consider the event of Napster as a natural experiment of temporarily loose enforcement of copyrights. First, Napster enabled millions of its users to easily search and download songs from other users of Napster for free, so that they were able to enjoy music inexpensively. In doing so, many of users violated copyrights.[17] Before the ruling of the U.S. appeal court for the Ninth Circuit, however, little legal action was taken against individual users to prevent them from file-sharing of copyrighted music. In other words, the introduction of Napster facilitated copyright infringement, hence loosening copyright protection.[18] Second, when Napster was popular around 2000, it was the only dominant file-sharing service. Hence, treatment of Napster, if any, can be considered reasonably homogeneous, though its effect might vary substantially for different consumers. Third, it is not unreasonable to assume that the introduction of Napster was exogenous of household-level expenditure on recorded music. Individual-level music consumption is unlikely to have affected the introduction of Napster in June 1999.

Accordingly, difference-in-difference (DID) seems to be the natural approach to quantify the effect of Napster. The idea of DID is simple. For those who had access to the Internet, any change in music expenditure is likely to have resulted from the treatment effect of Napster as well as a general time effect. Specifically, I consider $[Y_a^1 - Y_b^1]$ for this treatment group, where Y denotes expenditure on recorded music,

[17] Distributing or copying songs without permission from copyright holders is violation of copyrights. At that time, most songs were not given such permission, except a limited number of songs.

[18] Likewise, closing Napster in early 2001 can be imagined as natural experiment of strict copyright enforcement. Such interpretation, however, is confounded by one crucial problem. By the time Napster was closed, other file-sharing services such as KaZaA became popular and more people downloaded songs for free until legal online music services were popularized in 2003.

superscript 1 indicates the treatment group, and subscripts a and b, respectively, mean $after$ and $before$ Napster was introduced. For those who did not have Internet access, on the other hand, most of the change in music expenses was likely due to a time effect. Similarly to the treatment group, I consider $[Y_a^0 - Y_b^0]$ for this group, where superscript 0 denotes the control group. Accordingly, the effect of Napster can be obtained by differencing time effect out. The DID estimator is then given by

$$\text{DID} = [E(Y_a^1) - E(Y_b^1)] - [E(Y_a^0) - E(Y_b^0)]. \tag{2}$$

This simple framework, however, relies on the following assumptions. First, it is implicitly assumed that the treatment effect enters the outcome equation linearly. In terms of equation (1) in section 3,

$$Y_t^j \equiv p_t^{CD} \times CD_t^j = \phi_t^j(p_x, p_z, M, A) + \alpha DM_t^j + \epsilon_t^j, \qquad j = 0, 1, \quad t = a, b, \tag{3}$$

where DM means using Napster, $\phi_t^j(\cdot)$ is an unknown function of its argument, ϵ is an independently distributed error term, p_t^{CD} denotes price of CD, and α captures the effect of Napster. DM_t^j is a dummy variable which is equal to one if the household was using Napster, so that DM_t^0 is always zero. There is no reason to believe that DM_t^j enters the outcome equation linearly, but I assume this because otherwise, I would need to consider a different approach to measure the effect of Napster.

The second assumption is related to variables indicating the Internet connection. In the CEX, two kinds of information indicate whether the household has Internet access from home. The first is expenditure on computer information service which consists mainly of Internet service fees. The second is whether the household is living in a college dormitory. The CEX identifies students living in college dormitories as separate households from their parents. Presumably, most college dormitories were already equipped with broadband connection in the late 1990's. Consequently, I assume that households had the Internet connection, either if their expenses on computer information service were positive, or if they were living in a college dormitory.

Third, I assume that a household with an Internet connection adopted Napster immediately after it was introduced in June 1999. This might not be too unreasonable provided that Napster became popular within a surprisingly short period of time. However, both this and the second assumption are certainly not satisfactory. Given limited information contained in the CEX, I cannot do without these assumptions. Although there are a few data sets available which contain detailed information on usage of the Internet and downloading music, incorporating these data into the current DID framework is not feasible. However, there are other frameworks which might be able to overcome this problem. For example, a two-sample instrumental variable method as in Arellano and Meghir (1992) and Angrist and Krueger (1992) might enable one to combine these other data sets and the CEX. Because this method requires different assumptions on the economic model and related moment conditions, I do not consider it in this paper. Future research will investigate the effect of Napster under this framework.

4.2. Difference-in-difference kernel matching

I use 1996–2001 Interview surveys in the CEX. The CEX is composed of an Interview panel survey and a Diary survey.[19] The Interview survey interviews each household five times for every three months; the first for general demographics and the second through the fifth for household expenditures. In the Interview survey, new panels are added every month of the year, so that twenty percent of the sample that finished its final interview in the previous quarter are replaced by panels initiated newly in each quarter.[20]

This limited panel structure, however, is not used in this paper due to the following reasons. First, the period of the panel survey for the household lasts only one year. Second, many households in the sample skipped a few interviews. Third, the CEX provides different weights for different interviews of the same household. Accordingly, I consider each interview of the household as one observation. In other words, if a household performed two interviews, I consider two interviews as separate observations from two different households. The overlapping feature of the Interview survey can then be avoided in this way.

The DID estimator in (2) using the CEX, however, entails one serious bias. The following example illustrates this potential problem. Suppose that only a small number of high-income households had access to the Internet before 1999. This means that the treatment group before the introduction of Napster might spend a lot on many items including CDs, whereas the control group before Napster might consist of middle- and low-income households who spent less on most items including CDs. In other words, it is plausible that $E(Y_b^1)$ is high and $E(Y_b^0)$ is low. Suppose also that more people have the Internet connection after the emergence of Napster, so that some of middle-income households are now included in the treatment group instead of the control group. As a result, $E(Y_a^1)$ is not as high as before, while $E(Y_a^0)$ does not change much because it was already low. This yields a negative estimate of DID, or negative effect of Napster, even though there was no effect of Napster on music expenditure.

The previous example is a simplified situation, but it implies that the control

[19] The Diary survey is completed by the sample households for two consecutive one-week periods. Because the Interview survey contains more detailed information on expenditures, ownership of appliances, and demographics, I use the Interview survey. See the appendix for more detail on the CEX.

[20] Because of the overlapping feature in the Interview survey, some observations include quarterly expenditures occurred in months before and months after the introduction of Napster. Accordingly, I define an observation as *before*-group if the beginning month of three months, during which actual expenditures occurred, is earlier than April 1999. Similarly, *after*-group is defined if the beginning month of the three months is later than June 1999.

Table 1: Average expenditure on recorded music and the percentage of the treatment group in total samples in the CEX

	Year	1996	1997	1998	1999	2000	2001
Average music expenditure[a]	Treatment group	27.64	26.86	26.10	23.29	18.90	16.92
	Control group	11.16	11.50	10.58	9.74	8.74	7.98
Percentage of the treatment group[b]		8%	12%	18%	24%	29%	35%

[a]In terms of the quarterly household expenditure in 1998 dollars.

[b]The percentage of each year $= \dfrac{\text{No. of samples in the treatment group}}{\text{No. of total samples}} \times 100$

group may not be comparable to the treatment group. More precisely, the potential problem is that *before*-group, whether it is the treatment or control group, is not likely to be comparable to *after*-group, so that the treatment group is not comparable to the control group. In terms of equation (3), $\phi_b^j(p_x, p_z, M, A)$ is different from $\phi_a^j(p_x, p_z, M, A)$, so that $[E(Y_a^j) - E(Y_b^j)]$ does not difference out $\phi_t^j(\cdot)$. This problem arises not only because the ratio of Internet users to non-users has continued to increase over time, but also because the CEX is, essentially, repeated cross-sectional data.

This possible problem is indicated in table 1. First, average music expenditures have decreased over time for both groups. However, the decline for the treatment group is more substantial than that for the control group. Second, the number of samples in the treatment group has increased while that in the control group has decreased. These two observations seem to be consistent with the preceding example.

Consequently, the aforementioned bias is very likely to affect the DID estimator. In order to account for this potential bias, I need to control for the probability of having Internet access and make the control group more comparable to the treatment group. In order to control for such selection bias, I consider difference-in-difference kernel matching(DDM), as in Heckman and Smith(1999), and Smith and Todd(2003).

The procedure to estimate DDM is given as follows. I first estimate $P(A) \equiv Pr(D = 1|A)$, the propensity score of having Internet access from home given demographics. For the current results, I estimate a Probit using the CEX in 2000. To estimate $P(A)$, I use only data in 2000 due to the following reasons. First, the probability of having Internet access has increased over time. However, the probability for 2000 seems to be more relevant than those for other years because Napster became popular in 2000. Second, and more importantly, the estimated $P(A)$ using 2000 data satisfies the heuristic specification test of propensity score proposed by Shaikh, et al.(2004) better than those using data in other years or all years. The appendix presents a brief description of this test and the corresponding results with kernel density of the propensity score for treatment and control groups. The kernel density of $P(A)$ also indicates that trimming to ensure the common support for both groups is not necessary in the present analysis because I have a positive density for almost the entire region of the propensity score.

Based on the estimated $P(A)$, I then essentially match each individual in the treatment group with samples in the control group. Similarly to the repeated cross-section version of DDM estimator in Smith and Todd(2003), I specifically estimate

$$\begin{aligned}\widehat{\text{DDM}} &= \frac{1}{N_a^1}\,\Sigma_{i\in I_a^1}\{Y_{ai}^1 - \Sigma_{j\in I_a^0} W(i,j) Y_{aj}^0\} \\ &- \frac{1}{N_b^1}\,\Sigma_{i\in I_b^1}\{Y_{bi}^1 - \Sigma_{j\in I_b^0} W(i,j) Y_{bj}^0\},\end{aligned}$$

$$\text{where} \quad W(i,j) = \frac{G(\frac{P_j - P_i}{a_n})}{\Sigma_{k\in I^0} G(\frac{P_k - P_i}{a_n})}, \tag{4}$$

where $G(\cdot)$ is a kernel function, a_n is a bandwidth, and I^1 and I^0 denote the treatment group and the control group, respectively. For $G(\cdot)$, I use normal density kernel. The bandwidth is set to .001. As for number of observations, $N_a^1 = 19,647$, $N_b^1 = 10,841$, $N_a^0 = 41,507$, and $N_b^0 = 66,937$.

Because of the large number of samples, the actual implementation of the DDM estimator in (4) will require enormous computational hours. Note that I need to compute kernel-weighted averages, $\Sigma_{j\in I_t^0} W(i,j) Y_{tj}^0$ for every observation in the treatment group. In other words, for every observation among 19,647 in the treatment group after Napster, I need to compute kernel-weighted average using 41,507 observations in the *after*-control group. Because bootstrap estimation is required for a consistent standard error of DDM using the estimated propensity score, such implementation would be practically infeasible. Consequently, I instead discretize the propensity score from 0 to 1 by .001. Hence, I round the propensity score at the fourth decimal places. I compute the kernel-weighted average for each propensity score from .001 to .999, and then I match these kernel-weighted averages with each observation in the treatment group based on $P(A)$ to compute $\widehat{DDM}$.

4.3. Estimation results

I estimate both DID and DDM for all samples. The results are as follows[21]:

$$\widehat{DID} = -5.344(1.169), \quad \text{and} \quad \widehat{DDM} = -2.463(1.065).$$

[21] All numbers are in terms of the quarterly household expenditure in 1998 dollars. Bootstrap standard errors are shown in parentheses. They are based on 50 replications with 50% resampling with replacement. It takes two and half hours to compute one replication. However, if I use 100% sampling, it will take five hours for one replication. In order to reduce computational hours, therefore, 50% sampling is done.

Both estimates are precisely estimated, and DDM clearly reduces selection bias. The results from DDM imply that the quarterly music expenditure of the average U.S. household had declined by approximately three dollars as a consequence of using the Internet and, possibly, starting to use Napster.

One possible interpretation of the preceding results is as follows. If a household with Internet access started to use Napster, then the household spent roughly ten dollars less in one year than it would have without Napster. This amount is almost equivalent to two thirds of a CD album. Under the assumption that prices of CDs have changed little, the appearance of Napster then seems to have shifted demand for recorded music downwards by approximately two thirds of a CD album.

Furthermore, the results from DDM imply that 33% of the recording sales decrease in 2000 could be attributed to the emergence of Napster, which is based on the following computation. Assume that there are 100 million households in the U.S., and that the percentage of households with the Internet connection is given by 24% in 1999 as in table 1. Note that I use the percentage in 1999 instead of 2000. Because the underlying assumption for the DID approach is that households in the treatment group had access to the Internet before the introduction of Napster, more relevant percentage of households with Internet access is 24% in 1999 rather than 29% in 2000. The yearly decline in recording sales due to Napster is then calculated by

$$-\$236.448 \text{ million} = 100 \text{ million} \times 0.24 \times (-\$2.463) \times 4$$

From 1999 to 2000, the total real value of recorded music declined by 711.11 million dollars.[22] Accordingly, 33% of the decrease in total real value of recording sales in 2000 might be due to the emergence of Napster.

The previous interpretation of the results, however, requires caveats. First, the DDM method might not completely remove the aforesaid selection bias. Using DDM, I have reasonably matched the control group with the treatment group. However, one serious problem is that the size of the treatment group has increased substantially, so that *before*-treatment group may not be comparable to *after*-treatment group. Those who started to use the Internet around 2000 could be middle- or lower middle-income group, so that using the Internet substituted for their other entertainment including recorded music because their budget constraint and time constraint could then be binding. It is not clear whether DDM successfully controls for this difference between *before*-group and *after*-group.

Second, although Napster is likely to have influenced household-level music expenditure for the treatment group after the emergence of Napster, it is also highly

[22] According to the RIAA, nominal value of total shipments in 1999 is \$14,584.7 million, and that in 2000 is \$14,323.7 million. Because the estimated DDM is in terms of 1998 dollar, I divide these nominal values by CPI deflator for all items with 1998 as base year. The resulting real values are \$14,269.84 million in 1999 and \$13,558.72 million in 2000.

Table 2: DDM estimates for other entertainment goods[a]

Goods	Admission	Video	Toys	Reading	Sport
$\widehat{DDM}$	-4.869	1.294	1.287	-1.587	-0.772
S.E.[b]	2.289	0.892	2.941	1.529	1.846

[a]All numbers are in terms of the quarterly household expenditure in 1998 dollars.

[b]Bootstrap standard errors are reported. They are based on 50 replications with 50% resampling with replacement.

likely that other factors related to using the Internet might have affected music expenditure for the treatment group at the same time. Since the late 1990's, the Internet has been developed considerably, thereby inducing many consumers to spend more time and dollars on various activities on the Internet aside from downloading music. This implies that there are limitations in treating the introduction of Napster as a natural experiment, namely, identifying the treatment of Napster only by one time variation. In terms of equation (3), I use a proxy variable for DM_t^j because I do not observe whether a household used Napster or not. That is, $DM_b^1 = 0$ and $DM_a^1 = 1$. However, ϵ_t^j might capture unobserved activities on the Internet which became popular after 1999 and substituted for listening to music. As a result, DM_t^j might be correlated with ϵ_t^j.

One way of verifying the seriousness of this problem is to apply the DDM method to other entertainment expenditures. If the estimated magnitude for recorded music expenditure using DDM is to be attributed to Napster only, the same method should not find any changes in expenditures on other entertainment. Accordingly, I consider five other entertainment expenditures: Admission, Video, Toys, Reading, and Sport.[23] I apply exactly the same DDM method to these expenditures. Table 2 reports the results. Most estimates do not seem to be precisely estimated, except Admission. This indicates that the problem related to DDM might not be too serious. However, it should also be noted that the estimated magnitude for Admission is considerable and does not seem to be estimated imprecisely. Additionally, it is possible that more replications or 100% resampling might be needed to compute bootstrap standard errors.

As a results, DDM seems to reasonably quantify the effect of Napster on household-level music expenditure. Nonetheless, I cannot rule out the possibility that the estimated magnitude using DDM also captures the general effect of using the Internet that might have reduced expenditures on some entertainment goods including

[23]Admission includes expenditures on both season tickets and single tickets to movies, plays, opera, and concerts. Video consists of video cassettes, tapes, or discs both purchased and rented. Toys mean not only toys but also games including video games, computer game software, and hand-held computer game. Reading includes books and magazines. Sport is composed of both season tickets and single tickets to spectator sporting events.

recorded music or admission tickets to movies and concerts.[24]

The problems pointed out for DID in section 4.1 are also applied to DDM because both are based on the same idea, namely, differencing out a common effect. This framework relies on a linearity assumption. However, this assumption may be highly restrictive, particularly in equation (3). For example, other factors such as prices of other goods may enter the demand function nonlinearly, in which case the differencing strategy in the DID framework might fail to work.

Therefore, in the following sections, I consider two alternative methods to quantify the changes in household-level music expenditure and demand attributable to changes in prices for other entertainment goods as well as demographics. The estimated measure for such changes will indirectly inform us of the effect of Napster on music sales.

5. Other entertainment goods: demand system estimation

Other reason for the current downturn in recorded music could be the effect of changes in the prices for other goods, particularly other entertainment goods. This section investigates this possibility. To complete this objective, I estimate a household-level demand system for entertainment goods, from which I recover cross-price elasticities between these goods. I first discuss the demand system for entertainment goods, describing variables and data. I then specify an econometric model of the demand system. I next present estimation results and examine their implications.

5.1. Entertainment demand system

I begin with clarifying what I mean by entertainment goods. Entertainment usually means movies and music, but it sometimes indicates activities other than labor, namely leisure or recreation. However, I restrict its definition to the usual meaning. I thus consider movies, music, and computer games only. Hence, other leisure activities such as reading books and magazines, attending sports games, or general hobbies, like photography, are ignored.

Specifically, I use the following variables for entertainment goods. First, *CDs* denote recorded music including CDs, tapes, and LPs. Second, *Video* includes video cassettes, tapes, or discs, both purchased and rented. Third, *Admission* consists of both season tickets and single tickets to movies, plays, opera, and concerts. Lastly, *Toys* mean not only toys but also games including video games, computer game software, and hand-held computer games.

For four reasons, all the variables are aggregates of similar goods. First, the CEX defines some variables in that way. Second, these variables correspond to detailed *Consumer Price Index* (CPI) data estimated by the BLS. Third, aggregation reduces

[24] Separating this general effect of the Internet from that of Napster is desirable, but it will be attempted in the future research as described in section 4.1.

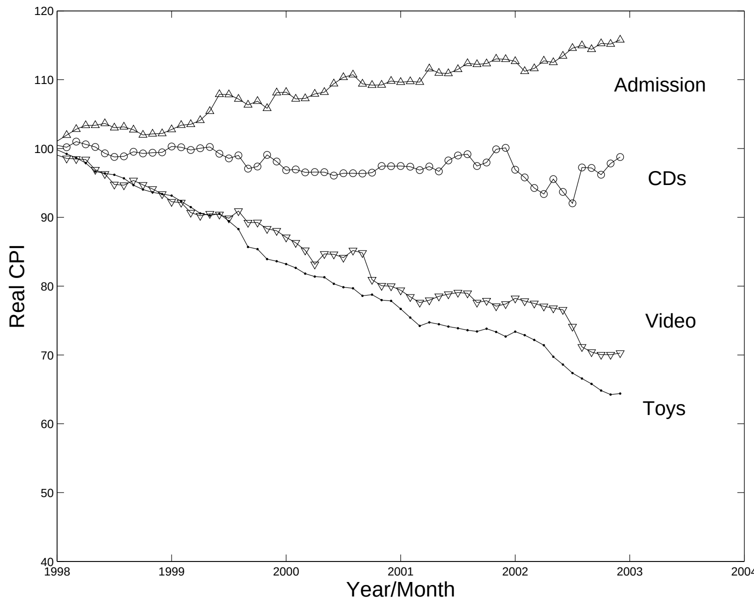

Figure 2: Monthly real CPI for entertainment goods (1998.1. – 2002.12.)

the number of goods to be considered, hence avoiding an extremely complicated demand system. Fourth, aggregation decreases observations with zero expenditure that might result from corner solutions, infrequent purchase, or measurement errors. Accounting for zero expenditure for many goods might be practically difficult.

Given these goods, I then make two crucial assumptions. First, I assume that entertainment goods are separable from labor and all the other goods in household utility. This separability assumption and the narrow definition of entertainment as above are maintained so as to avoid a complicated demand system and resulting difficulties in estimation.

The second assumption is that households take prices as given. It might be possible that prices are endogenous in the demand system since firms would charge prices differently depending on total demand. However, it is not unreasonable to assume that prices are exogenous from the perspective of an individual household, because any influence that each household would make on total demand is not significant.

For prices of the four entertainment goods, I use the monthly CPI. The CPI might not be ideal price data because it allows only time variation, however, transaction-level prices for these goods are not publicly available, and I have not succeeded in obtaining any price data with cross-sectional variation as well as relevant demographic information that can be matched with the CEX. Nonetheless, time variation and some limited cross-sectional variation[25] in the monthly CPI seem to be sufficient to identify price coefficients in the demand system.

The BLS publicly reports the CPI for CDs beginning from January 1998. Therefore, I use the 1998-2001 CEX and monthly CPI. Figure 2 plots the monthly real price indexes for entertainment goods.[26] According to this figure, prices for CDs did not change much, whereas prices for Video and Toys declined substantially, and prices for Admission relatively increased. Since all series for price indexes are normalized to 100 in January 1998, the relative magnitudes of actual level changes in prices for these goods could be either bigger or smaller than those presented in figure 2. This could cause some problem in estimation, but there is no solution to this problem given available data.

5.2. Econometric specification

In order to derive a flexible demand system on which I can impose the restrictions implied by utility maximizing behavior, I assume the translog indirect utility function for entertainment goods. Following Wolak(1996a,b), I specifically assume the translog indirect utility function with demographics defined as

$$\ln V(P,M,A) = \alpha_0 + \sum_{i=1}^{4} \alpha_i \ln(\frac{p_i}{M}) + \frac{1}{2}\sum_{i=1}^{4}\sum_{j=1}^{4} \beta_{ij} \ln(\frac{p_i}{M})(\frac{p_j}{M}) + \sum_{i=1}^{4}\sum_{k=1}^{K} \eta_{ik} \ln(\frac{p_i}{M}) A_k,$$

where P denotes a vector of prices, M is total expenditure on all the entertainment goods, and A is a vector of demographic variables.

The conditions consistent with utility maximizing behavior are (i) homogeneity of degree zero in prices and total expenditure, (ii) symmetry of the Slutsky matrix, and (iii) negative semi-definiteness of the Slutsky matrix. The translog indirect utility satisfies the first condition without imposing any further restrictions. The second condition can be easily imposed by setting $\beta_{ij} = \beta_{ji}$. Imposing the third condition, however, is not straightforward.[27]

[25] Since the CEX is quarterly data, I need to use three-month average of price indexes instead of monthly price indexes. In every month, however, approximately 10% of samples in the CEX are dropped because of the ending of all five interviews, and new samples are added to replace old ones. Because of this rolling panel nature, I compute three-month average of price indexes for each beginning month of three months, and then match it with observations in the CEX by the first month of quarterly expenditures. As a result, not every household faces the same three-month average price indexes, which introduces some additional cross-sectional variation in prices.

[26] Real CPI is computed as follows. I first divide monthly nominal CPI's for all items by the nominal CPI for all items in January 1998, where nominal CPI means the reported CPI in the BLS web site. The resulting numbers are CPI deflator. I then obtain real CPI by dividing nominal CPI for each entertainment item by the CPI deflator.

[27] The third condition should be imposed on estimation to satisfy integrability. However, the present estimation is performed without imposing this restriction, which might prevent one from interpreting the estimated results as economic primitives in indirect utility function.

Given this indirect utility function, Roy's identity implies the translog share equation as follows:

$$w_i(P, M, A) = \frac{\alpha_i + \Sigma_{j=1}^{4}\ \beta_{ij}\ln(p_j/M) + \Sigma_{k=1}^{K}\ \eta_{ik}A_k}{-1 + \Sigma_{j=1}^{4}\ \beta_{Mj}\ln(p_j/M) + \Sigma_{k=1}^{K}\ \eta_{Mk}A_k}, \tag{5}$$

$$\text{where } w_i = \frac{p_i x_i}{M}, \quad \text{and} \sum_{i=1}^{4}\ p_i x_i = M \quad (\text{or} \sum_{i=1}^{4}\ w_i = 1),$$

with the following restrictions;

$$\begin{aligned} \sum_{i=1}^{4} \alpha_i &= -1 & &\text{(normalization)} \\ \beta_{ij} &= \beta_{ji}, \quad \forall i, j & &\text{(symmetry of Slutsky matrix)} \\ \beta_{Mi} &= \Sigma_{j=1}^{4}\beta_{ij}, \quad \forall i & &\text{(budget constraint)} \\ \eta_{Mk} &= \Sigma_{j=1}^{4}\eta_{jk}, \quad \forall k & &\text{(budget constraint).} \end{aligned}$$

Notice that these restrictions allow me to recover parameters in the demand equation for any one good from those in the equations for the rest of goods, which leads me to estimate the demand system using three share equations.

As a final element to estimate the preceding demand system, I assume a mean zero error, ϵ_i, which enters the share equation linearly. ϵ_i is supposed to capture differences between the observed shares in the left-hand side and the predicted shares in the right-hand side of equation (5). In other words, ϵ_i is considered as the part of the indirect utility function that is unobservable to the econometrician. It is assumed to be independently distributed across households but possibly correlated with the errors in the other share equations for the same household.

Although the entertainment demand system as discussed seems reasonable and is likely to yield consistent estimates for parameters, I should point out potential problems that may be important but are currently not accounted for mainly due to difficulty in estimation or lack of relevant information. I plan to take these into account in the near future. First, zero expenditures in entertainment goods are likely to result from corner solutions. Lee and Pitt(1986) and Wolak(1996a) provide a framework and actual estimation that take corner solutions into account in the demand system. For present analysis, however, I do not account for this problem. Second, I do not impose the restriction of a negative semi-definite Slutsky matrix. Third, M is assumed to be exogenous although it is likely to be endogenous. Fourth, it is probable that current expenses on appliances such as a sound component system and VCR might affect current demand for entertainment and be endogenous as well, hence motivating durable good modeling. However, I only include the number of electronic appliances owned in A, while ignoring current expenditures on these appliances. Lastly, the maintained assumption regarding exogenous prices may not be innocuous.

5.3. Estimation results

Similarly to Wolak (1996a,b), I estimate the entertainment demand system in (5) by maximizing the quasi-maximum likelihood given by

$$\begin{aligned} L(\theta, \Sigma) &= -J(4-1)\ln(2\pi) - \frac{J}{2}\ln|\Sigma| \\ &\quad - \sum_{j=1}^{J} \frac{1}{2}(y_j - f_j(P^j, M^j, A^j))'\Sigma^{-1}(y_j - f_j(P^j, M^j, A^j)), \end{aligned}$$

where θ is the vector of parameters in (5), Σ means the covariance matrix for the error terms, y_j denotes 3-dimensional vector of observed expenditure shares for the jth household, $f_j(P^j, M^j, A^j)$ is 3-dimensional vector of fitted shares, and J is the total number of households.

Specifically, I estimate the system of three equations for CDs, Video, and Admission by imposing the restrictions listed above. Hence, the parameters for the Toys equation is recovered from these restrictions and the estimated parameters. As for demographic variables, I consider information on age, education, gender, family size, family composition, working, region, internet access, and appliance ownership. I include thirty-nine demographic variables which I presume would be relevant for entertainment demand. The list of these variables is attached in the appendix. I estimate this entertainment demand system using 69,915 observations from the 1998-2001 CEX.[28] Estimation results are reported in the appendix. To summarize, most of the price coefficients are estimated precisely, and many of demographic coefficients are significantly different from zero.

The goal of this exercise is to estimate own-price and cross-price elasticities between different entertainment goods in order to quantify the effect of relative prices for other entertainment goods on demand for CDs. In the case of the translog demand system, own-price and cross-price elasticities are given by

$$\begin{aligned} e_{ii} &= \frac{\frac{\beta_{ii}}{w_i} - \Sigma_k \beta_{ki}}{B} - 1, \quad \forall i \\ e_{ij} &= \frac{\frac{\beta_{ij}}{w_i} - \Sigma_k \beta_{kj}}{B}, \quad \forall i \neq j, \end{aligned}$$

where $B = -1 + \Sigma_j \beta_{Mj} \ln(\frac{p_j}{M}) + \Sigma_k \eta_{Mk} A_k$. Using the parameter estimates from the demand system estimation, I calculate these elasticities for each observation in the

[28] The number of original samples is 115,633, among which observations with positive value of total entertainment expenditure are selected. This results in 69,915 samples. Among these, 26,810 observations have positive expenditure on CDs.

Table 3: Own-price and cross-price elasticity estimates[a]

	Goods	Mean	10th%	25th%	Median	75th%	90th%
e_{cc}	CDs	-0.949	-0.999	-0.994	-0.981	-0.951	-0.888
e_{vv}	Video	-5.740	-10.604	-5.435	-3.097	-2.017	-1.747
e_{aa}	Admission	-1.096	-1.146	-1.093	-1.069	-1.056	-1.048
e_{tt}	Toys	-2.595	-4.065	-2.531	-1.881	-1.612	-1.470
e_{cv}	CDs-Video	0.233	-0.055	-0.027	0.047	0.219	0.594
e_{ca}	CDs-Adm.	-0.336	-0.636	-0.333	-0.192	-0.125	-0.093
e_{ct}	CDs-Toys	0.005	-0.0706	0.009	0.044	0.059	0.069
e_{va}	Video-Adm.	1.231	0.134	0.194	0.496	1.142	2.573
e_{vt}	Video-Toys	3.822	0.587	0.803	1.678	3.573	7.762
e_{at}	Adm.-Toys	-0.328	-0.703	-0.270	-0.089	-0.015	-0.003

[a]The numbers are summary statistics of elasticities for 69,915 samples.

sample. The histograms of these elasticities are plotted in the appendix. Table 3 summarizes all of these elasticities.

According to table 3, own-price elasticities for all entertainment goods are negative for most samples, which should be expected for a reasonable demand system. Additionally, I compute the mean of e_{cc} for each year. It is -0.950 in 1998, -0.952 in 1999, -0.949 in 2000, and -0.946 in 2001. Because CD prices have changed little according to figure 2, this implies that the recent changes in music demand might result from the shift in demand, rather than changes in its slope.

The results of interest are cross-price elasticities between CDs and other goods. Table 3 shows that e_{cv} is mostly positive, indicating that CDs and Video are substitutes. e_{ca} is negative for most observations, implying that CDs and Admission are complements. CDs and Toys seem to be substitutes, but the magnitude of substitution between CDs and Toys is relatively inconsequential. These results partly explain why recorded music sales declined in the late 1990's. Figure 2 presented earlier shows that prices for CDs have not changed much while Video prices have decreased, and Admission prices have increased. These price changes in Video and Admission clearly explicate why demand for CDs has decreased.

One possibility is that as the relative prices of video tapes and discs decreased around 2000, many households purchased or rented more video tapes and DVDs, which substituted for their demand for recorded music. Moreover, it is plausible that if the households had attended the concerts of some artists, they would have purchased CD albums by the same artists. Because of the relatively high prices of tickets, however, most households went to music concerts less often than before, which resulted in the recent decline in demand for recorded music.

To measure this effect on music demand, I consider the following relationship which is derived from the definition of cross-price elasticities.

$$\frac{\Delta CD}{CD} = \frac{\Delta P_i}{P_i} \cdot e_{ci} \qquad i = v, a.$$

This relationship explains changes in music demand due to changes in prices of other entertainment goods. The real CPI for Video is 90.11 in 1999 and 83.86 in 2000, and that for Admission is 105.07 in 1999 and 108.19 in 2000. Using the mean of cross-price elasticities, I then obtain the magnitude of changes in other prices on music demand as follows.

$$\begin{aligned}\frac{\Delta CD}{CD} &= \frac{\Delta P_v}{P_v} \cdot e_{cv} \approx -0.069 \times 0.233 = -0.016077 \\ \frac{\Delta CD}{CD} &= \frac{\Delta P_a}{P_a} \cdot e_{ca} \approx 0.0297 \times (-0.336) = -0.009979\end{aligned}$$

According to the RIAA's Yearend statistics, total unit shipments of recorded music had decreased by 7% from 1999 to 2000. Therefore, 37% of the decrease in music demand might have resulted from changes in prices of Video and Admission.[29] Under the assumption that CD prices remained the same between 1999 and 2000, one can conclude that 37% of music sales decline in 2000 might be due to price changes in other entertainment goods.

6. Transition effect: synthetic cohort analysis

6.1. Transition of the medium for recorded music

Two kinds of transition of recorded music media can be imagined and both might affect the recent slump in recording sales. The first type of transition occurred between LPs and CDs. CDs were introduced in 1982 for the first time. They became popular in the late 1980's and gradually replaced LPs and cassette tapes throughout the 1990's. Table 1 shows this transition of the medium for recorded music. In 1990, CD sales consist of less than half of total sales, whereas more than 90% of sales rely on CDs in 2000. This fact implies that sales growth during 1990's might be partly due to the effect of transition. As mentioned previously, it is plausible that adult music buyers who used to have many LPs may have purchased many CDs during the 1990's, just to update their old LPs. However, this updating process may have ended in the late 1990's, which likely explains the current slowdown.

A similar pattern may be an ongoing transition of CDs to MP3's. MP3 and similar media for recorded music will possibly become more popular, and more music buyers may download more songs. As recent success in Apple's i-Tunes music stores suggests, quite a few consumers are willing to pay to download music from legitimate web sites. This possible transition, however, does not imply that online music distribution substitutes for existing media such as CDs, hence resulting in a downturn in recording sales. It rather implies that recording sales would have increased if the recording industry had adopted new technology much earlier. In other words, the current slowdown in recording sales may partly reflect this slow adjustment of the recording industry to new technologies.

[29] $37\% \approx \frac{0.016077+0.009979}{0.07} \times 100$

Table 4: Share of CD sales in total recorded music sales[a]

Year	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001
$\frac{CDs}{Total}$	0.46	0.55	0.59	0.65	0.70	0.76	0.79	0.81	0.83	0.88	0.92	0.94

[a]Source: The RIAA's Yearend statistics.

These two types of transitions may explain the recent slump in music sales. The latter one, however, is difficult to verify given information available at present. Therefore, this section focuses on the former type of transition. In order to investigate this transition effect, I need information on music expenditures of different birth cohorts. If the transition effect is to explicate the slump in the recording industry, at least two patterns should be observed. First, adult music buyers should have spent more than, or as much as, their young counterparts did. If most of music expenses were due to young consumers, then any change in music consumption by adult consumers would not affect recording sales substantially. Second, music expenses by adult cohorts should have increased until the late 1990's and then decreased. The following section analyzes this possibility.

6.2. Synthetic cohort analysis using the CEX

The CEX provides relevant information regarding this transition effect. I thus examine 1996–2001 Interview surveys in the CEX. Interview surveys contain the quarterly expenditures of households in the United States. The CEX is the unique data set with detailed demographic information and expenditures on recorded music including CDs, tapes, and LPs.

However, it has two features that need to be accounted for. The first is that it is a household-level data. As a result, if I regard expenditures of a household as those only by a head of the household, then this will underestimate expenditures by young consumers who are usually dependents of household heads. Table 5 shows this problem. If I use the CEX as individual-level data based on ages of heads, then thirties, forties, and fifties are over-sampled, whereas people younger than thirty are under-sampled. In this regard, it is necessary to take family composition of the household into account. The second feature of the CEX is that the Interview surveys are essentially repeated cross-sectional data though they include limited panel structure as described previously. Because of the same reason as in section 4.2, I do not use this panel feature. Accordingly, I consider each interview as one observation.

This cross-sectional feature leads me to conduct synthetic cohort analysis as in Attanasio and Davis (1996) and Attanasio (1998). Hence, I first construct cohorts. A cohort is formed by households whose head was born within a ten-year interval. Households with heads' ages between 27 and 36 in 1998, for example, belong to the same cohort with those with heads whose ages were between 28 and 37 in 1999.

Table 5: U.S. population and percentage by age group in 2001

Census (individual level)[a]			CEX (household level)		
ages	population	%	ages[b]	population[c]	%
0-9	39,509,984	13.7			
10-19	41,512,600	14.4			
20-29	39,185,522	13.6	20-29	13,831,212	12.5
30-39	42,871,294	14.9	30-39	22,333,674	20.2
40-49	44,303,788	15.4	40-49	24,095,014	21.8
50-59	33,772,415	11.7	50-59	18,908,130	17.1
60-69	21,192,111	7.3	60-69	12,278,176	11.1
70+	26,020,984	9.0	other	19,064,488	17.3
total	288,368,698	100%	total	110,510,692	100%

[a]Source: National Population Estimates by Population Division, U.S. Census Bureau.
[b]Ages of household heads.
[c]The population is the sum of weights provided by the 2001 CEX.

I then compute weighted sum of expenditure on recorded music for each cohort. Figure 2 reports the results.

A crucial observation from the figure 3 is that music expenditures by adult buyers seem to be substantial compared to young cohort with ages between 18 and 27 in 1999[30]. However, this interpretation is not entirely correct because the results are based on household-level data, not on individual-level data. Consequently, music expenditures by cohort 38-47 partially reflect purchases by their children who were likely to be teens. This explains relatively high expenditures by cohort 38-47. Nevertheless, this feature is not likely to explain fairly high expenses by cohort 28-37 and cohort 48-57 because children for cohort 28-37 might be too young to be heavy music buyers, and those for cohort 48-57 might already have their own households. Music expenditures by either cohort 28-37 or cohort 48-57, therefore, do not seem to reflect music expenses by their children. As a result, adult consumers are also likely to have purchased significant amount of recorded music, which verifies the first necessary pattern for the transition effect.

The second necessary pattern is examined in figure 4, which plots the weighted mean of music expenditures for each cohort in the CEX. As for cohorts 28-37 and 48-57, average expenditures on recorded music appear to have increased until the late 1990's, and then declined. This pattern is likely to be consistent with the transition effect because expenses for recorded music for these cohorts mainly reflect household heads' expenditures. In other words, they might have spent substantial amount of dollars on purchasing CDs during the 1990's, just to update their old LPs or cassette tapes which they had purchased when they were teens or 20's and 30's. Since the decline occurred even before 1999, it is not plausible that this decrease resulted from using Napster in late 1999 and 2000.

[30]Henceforth, I use cohort α-β to denote cohort with ages between α and β in 1999.

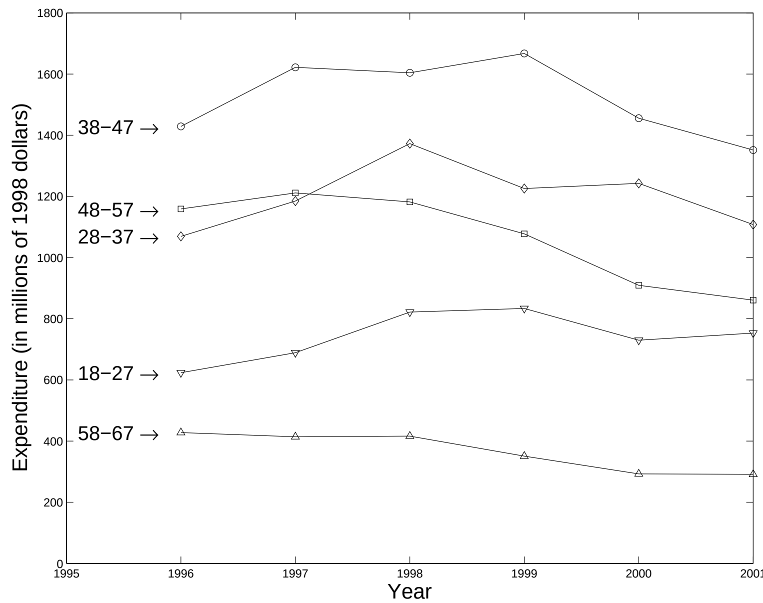

Figure 3: The weighted sum[a] of music expenditure by cohort (ages in 1999)

[a]The weights are provided by the CEX. They are assigned to an address and are the inverse of the probability of selection of the household.

Similar pattern occurred for cohort 38-47. However, the decline started from 1999. Moreover, expenditures for this cohort are likely to include their children's expenses on recorded music. Therefore, the expenditure pattern for this cohort does not rule out the possibility that downloading music might have resulted in the decline in music expenditure. In order to confirm these patterns more convincingly, I also need to examine data before 1996.

In the case of cohort 18-27, somewhat interesting pattern is observed. Average expenditure for this cohort continued to decrease from 1996 on. Note that these ages are those of household heads. Hence, households included in this cohort are mostly students or young workers living independently from their parents. The decline after 1999 could be explained by using Napster as the recording industry often blamed college students with broadband connection on campus for file-sharing activities. This effect of Napster, nevertheless, cannot explain the decrease in music purchase which occurred even before 1999.

In conclusion, the results presented in this section imply that the transition effect is consistent with the recent slowdown in the recording industry. Specifically, music expenditures by adult buyers consist of substantial proportion of total music sales, and their average expenditures seem to have increased until the late 1990's and then

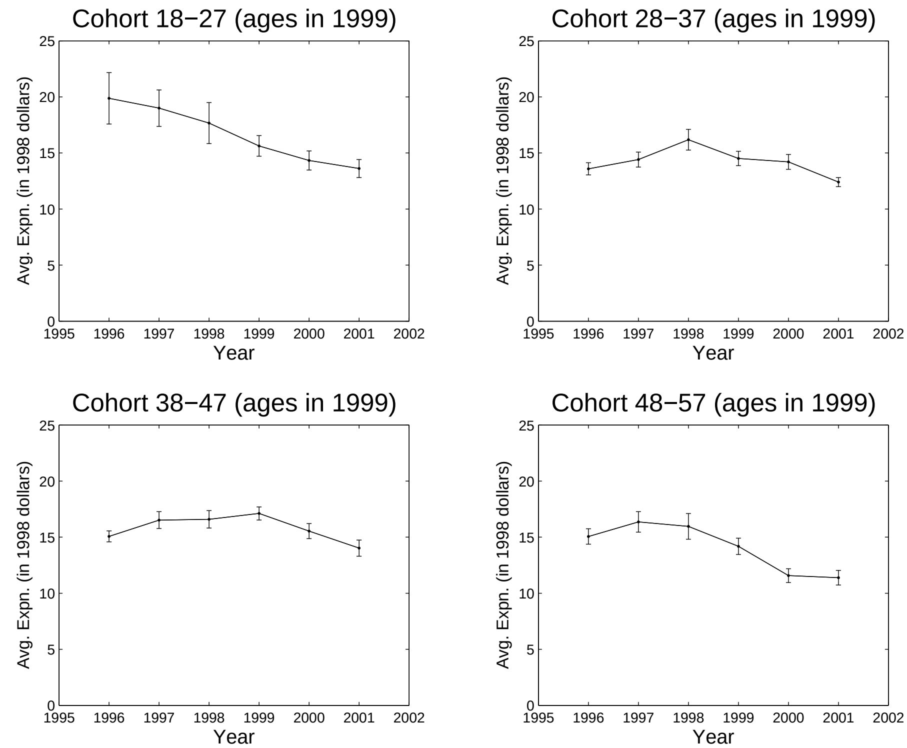

Figure 4: The weighted mean of expenditure of recorded music by cohort[a]

[a]Note: The bar stretching from the line indicates standard error for each mean.

declined. How much of the recent slump in recorded music sales is then due to this transition effect? In order to quantify this effect precisely, more work needs to be done in this direction.

Nevertheless, an approximate answer may be provided by examining changes in music expenditures by cohort, compared to total changes in recording sales. I compute the yearly changes in the weighted sum of music expenditure for each cohort. That is, I subtract the weighted sum for each cohort in each year by that in the previous year. Table 6 reports the results. The number of households in most of cohorts is fairly stable, except for cohort 18-27[31]. For most cohorts, therefore, drops in weighted sum of music expenditure are mostly due to declines in average music purchases. Assuming that decrease in music expenses for cohort 48-57 and

[31]The number of households in cohort 18-27 is increasing because more young people have their own households as they grow up.

Table 6: Level changes in weighted sum of music expenditure by cohort[a][b]

	Cohort (ages in 1999)					Total Change[c]	
Year Change	18-27	28-37	38-47	48-57	58-67	Level	%
1996 to 1997	65.51	115.55	193.25	52.40	-13.99	448.44	8.9
1997 to 1998	132.70	187.79	-17.97	-29.52	2.30	195.70	3.6
1998 to 1999	11.71	-147.01	63.32	-104.21	-65.25	-237.43	-4.2
1999 to 2000	-103.68	16.83	-211.75	-168.33	-58.32	-506.92	-9.3
2000 to 2001	23.19	-134.69	-104.04	-48.44	-1.37	-202.17	-4.1

[a]Source: author's calculation from the CEX 1996-2001.
[b]All values are in millions of 1998 dollars.
[c]Changes in weighted sum for all households in the CEX.

58-67 resulted from the transition effect, approximately 44.7% [32]of total decline in music sales from 1999 to 2000 might be then explained by the transition effect.

7. Conclusion and Future Research

As I discuss in the introduction, one fundamental premise of copyright protection is that copyrights secure revenues of copyright holders and the related industry. I attempt to verify this premise in the case of Napster by quantifying the magnitude of changes in household-level music expenditure attributed to the emergence of Napster. Exploiting the rich information contained in the Consumer Expenditure Survey, I use three approaches to measure this effect of Napster. The DDM method directly quantifies the effect. I find that the quarterly music expenditure of the average U.S. household has declined by two dollars and forty-six cents as a result of using the Internet and, plausibly, starting to use Napster. I show that this accounts for 33% of the decrease in total recording sales in 2000.

The second approach estimates a household-level demand system for entertainment goods. The estimated cross-price elasticities imply that changes in prices of other entertainment goods also explain the slump in recorded music sales. I find that in 2000, roughly 37% of the decline in recording sales is due to such changes in prices. The final method constructs synthetic cohorts. The results indicate that transition from LPs to CDs might describe the increase in music sales during the 1990's as well as the recent slowdown. This transition effect might explain 44.7% decrease of recorded music sales in 2000.

These two alternative methods indirectly measure the effect of Napster in that they explicate that more than 80% of music sales decrease in 2000 might have resulted from factors aside from Napster. This implies that the estimated magnitude using DDM may quantify changes in the household-level music expenditure due to not only Napster but also factors other than file-sharing of copyrighted music.

[32] $\frac{\text{(Changes for cohort 48-57) + (Changes for cohort 58-67)}}{\text{Total change in music expenditure from 1999 to 2000}} = \frac{-168.33-58.32}{-506.92} \approx 0.447$

Continuing in a similar direction, it would be interesting to perform the following counterfactuals: What would have happened to revenues of recorded music if copyrights had been enforced very strictly around 2000 when Napster was popular? Similarly, what if Napster had not been available? Answers to these counterfactuals will verify the preceding premise more directly. However, the current approaches would not be able to provide any direct evidence regarding these questions. This motivates a more extensive structural model. Another motivation for structural modeling is that it may be extended further to examine costs and benefit of copyright protection in the case of Napster and other file-sharing services.

There are other advantages of structural modeling. First, a more comprehensive structural model might enable me to combine the current three approaches. Second, it may allow me to link different data sets in order to exploit as much information as possible from many sources of data. Third, it will avoid some unsatisfactory assumptions in the current approaches.

These advantages, however, come with costs. Namely, a structural model relies on untestable assumptions. Nevertheless, it seems that the advantage outweighs its costs for this case. Therefore, I will continue to work in this direction.

References

Angrist, Joshua D. and Alan B. Krueger (1992) "The effect of Age at School Entry on educational Attainment: An application of Instrumental Variables with Moments from Two Samples," *Journal of the American Statistical Association*, 87: 328-336.

Arellano, Manuel and Costas Meghir (1992) "Female Labour Supply and On-the-Job Search: An Empirical Model Estimated Using Complementary Data Sets," *Review of Economic Studies*, 59: 537-557.

Attanasio, Orazio P. (1998) "Cohort Analysis of Saving Behavior by U.S. Households," *The Journal of Human Resources*, 33: 575-609.

Attanasio, Orazio P. and Steven J. Davis (1996) "Relative Wage Movements and the Distribution of Consumption," *The Journal of Political Economy* , 104: 1227-1262.

Battistin, Erich (2003) "Errors in Survey Reports of Consumption Expenditures," *The Institute for Fiscal Studies*, WP03/07.

Breyer, Stephen (1970) "The Uneasy Case for Copyright: A Study of Copyright in Books, Photocopies, and Computer Programs," *Havard Law Review*, 84: 281-351.

Heckman, James J. and Jeffrey Smith (1999) "The Pre-programme Earnings Dip and the Determinants of Participation in a Social Programme. Implications for Simple Programme Evaluation Strategies," *The Economic Journal*, 109: 303-348.

Lee, Lung-Fei and Mark M. Pitt (1986) "Microeconometirc Demand Systems with Binding Nonnegativity Constraints: The Dual Approach," *Econometrica*, 54: 1237-1242.

Shaikh, Azeem, Marianne Simonsen, Edward Vytlacil, and Nese Yildiz (2004) "A Heuristic for Identifying Misspecified Propensities for Program Evaluation," mimeo.

Smith, Jeffrey and Petra Todd (2003) "Does Matching Overcome Lalonde's Critique of Nonexperimental Estimators?," *Journal of Econometrics*, forthcoming.

Wolak, Frank A. (1996a) "The Welfare Impacts of Competitive Telecommunications Supply: A Household-Level Analysis," *Brookings Papers on Economic Activity: Microeconomics*, 269-340.

Wolak, Frank A. (1996b) "Can Universal Service survive in a Competitive Telecommunications Environment? Evidence from the United States Consumer Expenditure Survey," *Information Economics and Policy*, 8: 163-203.

Appendix A: Data description

The main source of data in this paper is the CEX. The CEX consists of an Interview panel survey and a Diary survey. Both surveys are conducted at the level of a consumer unit (CU) which is essentially a household. The CEX defines a CU as the person (or group of persons) in a household who is (or are) independent of all other persons in the household for payment of their major expenses. It is possible that more than one person (or group of persons) in a household became separate CUs. Since separate CUs report their expenditures independently, I consider them as different households. Particularly, students living in college- or university-regulated housing are considered as separate CUs from their parents' household.

Both the Interview and the Diary survey contain the similar information on demographics and expenditures, but each survey uses its own questionnaire and independent sample. In the Interview survey, each CU in the sample is interviewed once every three months over five consecutive quarters. The Diary, on the other hand, is completed by the sample CUs for two consecutive one-week periods. The Interview survey has more detailed information on expenditures relevant to the analysis in this paper than the Diary survey. Moreover, it includes data on household ownership of appliances such as computers and sound component systems, whereas the Diary does not. As a consequence, I use the Interview survey.

The Interview survey is composed of several files. My current data set is constructed by extracting data from CU characteristics and income (FMLY) file, major household appliances (APL) file, monthly expenditure (MTAB) file, and detailed expenditure (EXPN) file. In most cases, the expenditure in MTAB file is a third of the expenditure in EXPN file, but there are some cases in which the expenditure in one file is zero while it is positive in the other file. Hence, I use maximum of the expenditure in both files in a way to reduce missing expenditures. These quarterly expenditures are expenses on detailed items such as CDs, video rentals, books, and computer information service. The computer information service includes both Internet access and data services fees. The CEX started to collect this expense from the first quarter of 1996. Consequently, I use the Interview survey from 1996 to 2001. From 2001 on, the CEX began to collect more detailed data on Internet services such as expenses on DSL or ISDN, but this information is not used in this paper. Finally, I pool data for these six years to construct my data set. It contains 159,642 observations with 56,951 different CUs.

Appendix B: Definition of Variables

AGE	age of reference person in the household
WHITE	1 if the reference person is white
BLACK	1 if black
MALE	1 if male
HSGRAD	1 if highest education is high school graduate
LESSCOL	1 if some college, less than college
COLGRAD	1 if college graduate
FAM.SIZE	the number of members in the household
M.PERYT16	the number of males with age under 16
F.PERYT16	the number of females with age under 16
NTE1617	the number of members with age between 16 and 17
PERSOT64	the number of persons older than 64
NO.CH.LE11	the number of children younger than 11
NO.CH.1217	the number of children ages between 12 and 17
HW	1 if CU is family with husband and wife only
SINGLE	1 if CU is single
HW.YOUNG	1 if family only with husband and wife, and AGE under 40
HW.OLD	1 if family only with husband and wife, and AGE over 45
HW.CHILD.BF.SCH	1 if husband and wife with children before school
HW.CHILD.IN.SCH	1 if husband and wife with children in school
HW.CHILD.AF.SCH	1 if husband and wife with children after school
SP.CHILD.BF.SCH	1 if CU is single parent with children before school
SP.CHILD.IN.SCH	1 if single parent with children in school
RETIRED	1 if CU is retired
HEADWRK	1 if the head of the household is working
SPOUWRK	1 if the spouse of the household is working
INCWK1	the number of weeks in a year that head worked
INCWK2	the number of weeks in a year that spouse worked
INCHR1	the number of hours in a week that head worked
INCHR2	the number of hours in a week that spouse worked
FINCBTAX	Real final income before tax (in $10,000)
OWNER	1 if CU owns house
RENTER	1 if CU rents house
COLDORMI	1 if CU is living in college dormitory
NE	1 if household resides in Northwest Census region
MW	1 if household resides in Midwest Census region
WEST	1 if household resides in Census Western region
URBAN	1 if household resides in urban area
MSA	1 if household resides in Metropolitan Statistical Area
PS4MIL	1 if household resides in area with population size over 4 million
PS1MIL	1 if population size between 1.2 million and 4 million
PS330K	1 if population size between 330 thousand and 1.2 million
PS125K	1 if population size between 125 thousand and 330 thousand
INTNET	1 if expense on computer information service is positive
TV	Number of televisions in the household
COMP	Number of computers
SOUND	Number of sound components
VCR	Number of VCR
VEHQ	Number of vehicles

Appendix C: Probit estimation results for propensity score

Dependent variable: Dummy for having Internet access[a]

Variable	Estimate	S.E.	Pr> χ^2	Variable	Estimate	S.E.	Pr> χ^2
Intercept	-0.9008	0.1627	<.0001	spouwrk	-0.0840	0.0475	0.0773
age	0.0260	0.0040	<.0001	incweek1	-0.0032	0.0009	0.0005
$(\text{age})^2$	-0.0004	0.0000	<.0001	inchr1	-0.0019	0.0008	0.0216
white	0.2987	0.0942	0.0015	incweek2	0.0011	0.0010	0.2811
black	0.0315	0.0982	0.7485	inchr2	0.0015	0.0010	0.1430
male	0.1001	0.0187	<.0001	fincbtax	0.0562	0.0037	<.0001
hsgrad	0.3548	0.0337	<.0001	$(\text{fincbtax})^2$	-0.0013	0.0001	<.0001
lesscol	0.6056	0.0349	<.0001	owner	-1.2569	0.0710	<.0001
colgrad	0.6602	0.0336	<.0001	renter	-1.4255	0.0692	<.0001
fam.size	-0.0400	0.0166	0.0160	ne	0.0814	0.0272	0.0028
no.ch.le11	-0.0312	0.0224	0.1645	mw	0.0042	0.0236	0.8597
no.ch.1217	0.0005	0.0204	0.9820	west	-0.0258	0.0232	0.2651
persot64	0.0298	0.0256	0.2445	urban	0.0969	0.0469	0.0387
single	-0.1013	0.0403	0.0120	msa	-0.1194	0.0556	0.0319
hw.young	-0.0368	0.0864	0.6702	ps4mil	0.1180	0.0465	0.0111
hw.child.bf.sch	0.0752	0.0571	0.1876	ps1mil	0.1103	0.0453	0.0150
hw.child.in.sch	0.1566	0.0493	0.0015	ps330k	0.1332	0.0478	0.0053
hw.child.af.sch	0.0966	0.0453	0.0329	ps125k	-0.0313	0.0484	0.5179
sp.child.bf.sch	0.0037	0.0682	0.9567	tv	-0.0172	0.0086	0.0453
sp.child.in.sch	0.0698	0.0616	0.2574	comp	0.5605	0.0133	<.0001
hwold	0.0543	0.0783	0.4885	sound	-0.0009	0.0099	0.9262
retired	-0.0100	0.0450	0.8232	vcr	0.0212	0.0122	0.0830
headwrk	0.1644	0.0580	0.0046	vehq	0.0183	0.0063	0.0037

[a]1 either if the households spent positive expenditure on computer information service or if they were living in college dormitory.

Appendix D: A heuristic specification test for propensity score

A test for misspecified propensities proposed by Shaikh, et al. (2004) is based on the relationship between the density for the propensity score of the treatment group and that for the control group. Under correct specification of the propensity score,

$$\begin{aligned} f(P(A)|D=1) &= \frac{f(P(A), D=1)}{Pr(D=1)} = \frac{Pr(D=1|P(A)) \cdot f(P(A))}{Pr(D=1)} \\ &= P(A) \cdot \frac{f(P(A))}{Pr(D=1)}, \end{aligned}$$

where $P(A)$ is the propensity score, A is demographics, $f(\cdot)$ is density of $P(A)$, and D is treatment. The last equality follows from the definition of $P(A)$ and the assumption of selection on observables, $\{Y_1, Y_0 \perp D\}|A$, so that $P(A) \equiv Pr(D = 1|A) = Pr(D = 1|P(A))$. Likewise,

$$f(P(A)|D=0) = (1 - P(A)) \cdot \frac{f(P(A)}{Pr(D=0)}.$$

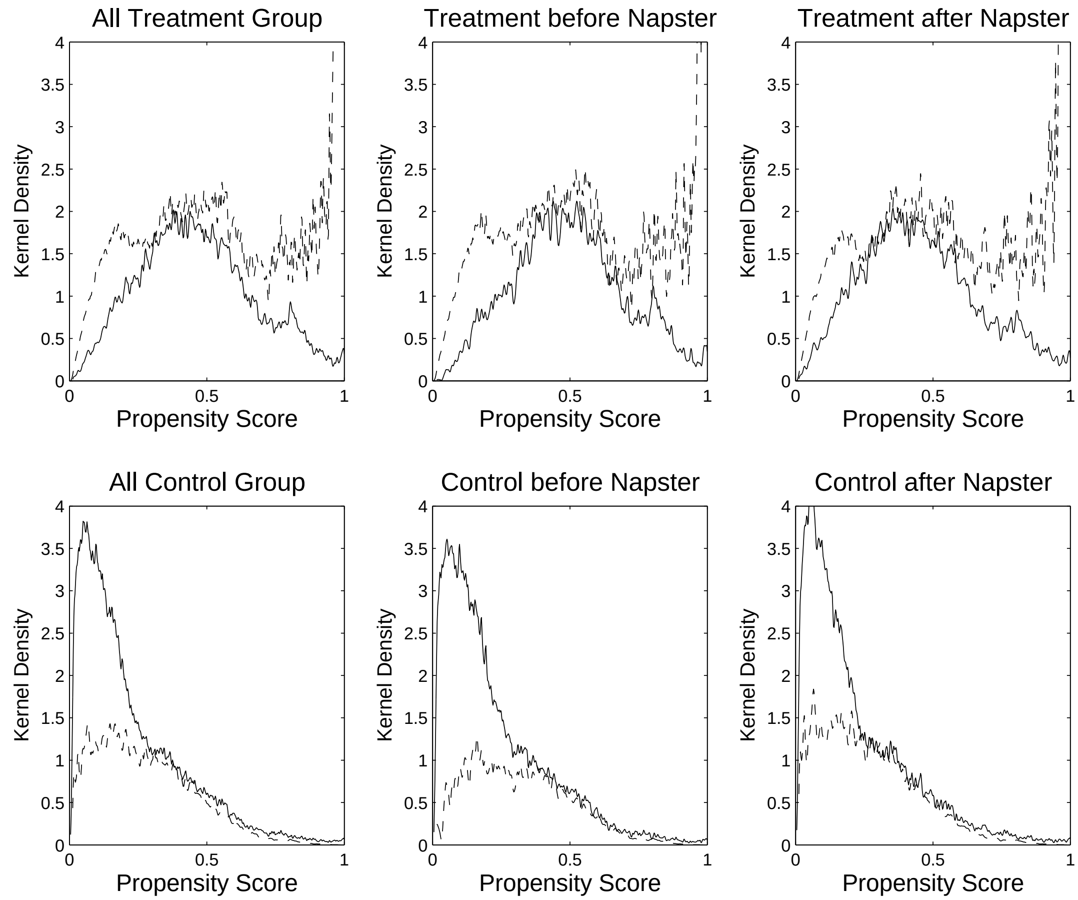

Figure D.1: Kernel density of propensity score

From the previous two equations, one can obtain the following relationship:

$$f(P(A)|D=1) = \frac{P(A)}{(1-P(A))} \cdot \frac{Pr(D=0)}{Pr(D=1)} \cdot f(P(A)|D=0).$$

Both $f(P(A)|D=1)$ and the right-hand side in the equation above are plotted in the figure D.1. The solid lines denote kernel densities of the propensity score, and the dotted lines denote the preceding right-hand side expression for the treatment group and the similar expression for the control group. Those two lines appear to be reasonably close to each other especially in the middle region. Although they are not close for the region near 0 and 1, the figure suggests that the null hypothesis of correct specification for the propensity score does not seem to be rejected.

Appendix E: Parameter estimates of entertainment demand system

	Common Denominator				CD equation				Video equation				Admission equation			
variable	param.	est.	s.e.	p-val.	param.	est.	s.e.	p-val.	param.	est.	s.e.	p-value	param.	est.	s.e.	p-val.
age	η_{M1}	-3.97E-03	1.03E-03	[.000]	η_{101}	-2.98E-04	2.05E-04	[.146]	η_{201}	6.09E-04	2.78E-04	[.029]	η_{301}	-2.72E-03	5.11E-04	[.000]
white	η_{M2}	0.209212	0.075837	[.006]	η_{102}	-7.99E-03	0.021511	[.710]	η_{202}	0.13938	0.031817	[.000]	η_{302}	-0.021292	0.029752	[.474]
black	η_{M3}	0.256717	0.075841	[.001]	η_{103}	-0.025701	0.023288	[.270]	η_{203}	0.187942	0.032547	[.000]	η_{303}	1.72E-03	0.029201	[.953]
male	η_{M4}	-9.91E-03	0.01658	[.550]	η_{104}	-3.18E-03	4.20E-03	[.449]	η_{204}	-0.015571	6.42E-03	[.015]	η_{304}	-5.21E-03	6.34E-03	[.411]
hsgrad	η_{M5}	0.090982	0.027996	[.001]	η_{105}	0.034237	7.90E-03	[.000]	η_{205}	0.035834	0.011246	[.001]	η_{305}	-0.041844	0.010957	[.000]
lesscol	η_{M6}	0.032811	0.02895	[.257]	η_{106}	0.012864	7.43E-03	[.083]	η_{206}	0.028546	0.011545	[.013]	η_{306}	-0.093966	0.016854	[.000]
colgrad	η_{M7}	0.022065	0.028381	[.437]	η_{107}	9.47E-03	7.18E-03	[.188]	η_{207}	0.044666	0.011921	[.000]	η_{307}	-0.141032	0.022133	[.000]
fam. size	η_{M8}	-0.043788	0.014718	[.003]	η_{108}	-0.011935	3.60E-03	[.001]	η_{208}	-0.023495	6.11E-03	[.000]	η_{308}	-7.66E-03	4.51E-03	[.089]
m.peryt16	η_{M9}	0.017439	0.01474	[.237]	η_{109}	0.039634	6.53E-03	[.000]	η_{209}	4.59E-03	5.52E-03	[.406]	η_{309}	0.029633	6.55E-03	[.000]
f.peryt16	η_{M10}	0.048579	0.016318	[.003]	η_{110}	0.029081	5.44E-03	[.000]	η_{210}	0.013193	5.92E-03	[.026]	η_{310}	0.022151	6.04E-03	[.000]
nte1617	η_{M11}	-0.083305	0.033439	[.013]	η_{111}	-0.036185	9.01E-03	[.000]	η_{211}	-0.020193	0.011735	[.085]	η_{311}	-0.074777	0.015163	[.000]
persot64	η_{M12}	-0.046524	0.026181	[.076]	η_{112}	-0.013139	6.18E-03	[.034]	η_{212}	0.022331	8.86E-03	[.012]	η_{312}	-0.0449	0.011806	[.000]
hw	η_{M13}	-0.095309	0.034024	[.005]	η_{113}	-0.035497	8.99E-03	[.000]	η_{213}	-4.54E-03	0.010603	[.669]	η_{313}	-0.05866	0.014293	[.000]
single	η_{M14}	-0.280498	0.054422	[.000]	η_{114}	-0.114368	0.018228	[.000]	η_{214}	8.31E-04	0.011765	[.944]	η_{314}	-0.168924	0.027493	[.000]
headwrk	η_{M15}	0.05808	0.04518	[.199]	η_{115}	0.013877	0.011557	[.230]	η_{215}	5.05E-03	0.016797	[.764]	η_{315}	3.93E-03	0.017578	[.823]
spouwrk	η_{M16}	-0.053921	0.031637	[.088]	η_{116}	-0.025878	8.42E-03	[.002]	η_{216}	-0.020449	0.011717	[.081]	η_{316}	-0.055671	0.013992	[.000]
incwk1	η_{M17}	-6.70E-04	7.37E-04	[.364]	η_{117}	-4.91E-04	1.98E-04	[.013]	η_{217}	-1.44E-04	2.80E-04	[.607]	η_{317}	1.46E-04	2.85E-04	[.608]
inchr1	η_{M18}	3.73E-04	6.94E-04	[.591]	η_{118}	9.40E-05	1.78E-04	[.598]	η_{218}	-3.79E-04	2.57E-04	[.140]	η_{318}	1.84E-04	2.67E-04	[.490]
incwk2	η_{M19}	2.28E-03	8.40E-04	[.007]	η_{119}	6.05E-04	2.09E-04	[.004]	η_{219}	1.05E-03	3.25E-04	[.001]	η_{319}	6.60E-04	2.94E-04	[.025]
inchr2	η_{M20}	-1.40E-03	8.20E-04	[.088]	η_{120}	-2.29E-04	2.03E-04	[.259]	η_{220}	-8.76E-04	3.21E-04	[.006]	η_{320}	8.20E-05	2.94E-04	[.780]
fincbtax	η_{M21}	-7.44E-03	2.13E-03	[.000]	η_{121}	-3.24E-03	6.54E-04	[.000]	η_{221}	-2.67E-03	7.20E-04	[.000]	η_{321}	-3.93E-03	9.24E-04	[.000]
owner	η_{M22}	0.024736	0.075026	[.742]	η_{122}	0.047341	0.019767	[.017]	η_{222}	-0.021989	0.028621	[.442]	η_{322}	-3.67E-04	0.028557	[.990]
renter	η_{M23}	0.011195	0.075771	[.883]	η_{123}	0.040025	0.019842	[.044]	η_{223}	-0.028174	0.029005	[.331]	η_{323}	-4.03E-05	0.028553	[.999]
ne	η_{M24}	0.034297	0.022607	[.129]	η_{124}	-5.91E-03	5.76E-03	[.305]	η_{224}	0.035356	9.21E-03	[.000]	η_{324}	-2.46E-03	8.67E-03	[.777]
mw	η_{M25}	3.47E-03	0.018524	[.852]	η_{125}	-6.66E-03	4.85E-03	[.170]	η_{225}	7.30E-03	7.01E-03	[.298]	η_{325}	-7.94E-03	7.03E-03	[.258]
west	η_{M26}	-0.157287	0.031721	[.000]	η_{126}	-0.043967	8.20E-03	[.000]	η_{226}	-0.054249	0.011163	[.000]	η_{326}	-0.078884	0.013775	[.000]
urban	η_{M27}	0.035953	0.036186	[.320]	η_{127}	0.016447	9.90E-03	[.097]	η_{227}	9.47E-04	0.014294	[.947]	η_{327}	-0.020601	0.013088	[.115]
msa	η_{M28}	0.016515	0.041805	[.693]	η_{128}	0.014788	0.011251	[.189]	η_{228}	-0.028123	0.017373	[.105]	η_{328}	0.029752	0.014984	[.047]
ps4mil	η_{M29}	-0.110147	0.03916	[.005]	η_{129}	5.37E-03	9.27E-03	[.562]	η_{229}	0.017741	0.014924	[.235]	η_{329}	-0.154924	0.024319	[.000]
ps1mil	η_{M30}	-0.073363	0.036136	[.042]	η_{130}	1.45E-03	8.90E-03	[.870]	η_{230}	0.016499	0.014512	[.256]	η_{330}	-0.108399	0.018789	[.000]
ps330k	η_{M31}	-0.079515	0.038506	[.039]	η_{131}	-0.01969	9.95E-03	[.048]	η_{231}	0.029003	0.015377	[.059]	η_{331}	-0.086091	0.017307	[.000]
ps125k	η_{M32}	-0.034171	0.038089	[.370]	η_{132}	-3.71E-03	9.66E-03	[.701]	η_{232}	0.024691	0.015738	[.117]	η_{332}	-0.07418	0.016365	[.000]
intnet	η_{M33}	-0.045207	0.018534	[.015]	η_{133}	-0.015949	4.92E-03	[.001]	η_{233}	-0.017361	6.63E-03	[.009]	η_{333}	-3.75E-03	6.66E-03	[.574]
coldormi	η_{M34}	-0.891532	0.307425	[.004]	η_{134}	-0.234973	0.080028	[.003]	η_{234}	-0.088552	0.06702	[.186]	η_{334}	-0.565284	0.149703	[.000]
tv	η_{M35}	0.023937	7.57E-03	[.002]	η_{135}	9.75E-03	2.17E-03	[.000]	η_{235}	0.012594	3.00E-03	[.000]	η_{335}	2.75E-03	2.74E-03	[.316]
comp	η_{M36}	-0.036223	0.012722	[.004]	η_{136}	-0.010497	3.32E-03	[.002]	η_{236}	-1.27E-03	4.12E-03	[.759]	η_{336}	-0.032024	6.30E-03	[.000]
sound	η_{M37}	-0.061313	0.012828	[.000]	η_{137}	-0.037232	5.66E-03	[.000]	η_{237}	-9.70E-03	3.68E-03	[.008]	η_{337}	-0.024431	4.91E-03	[.000]
vcr	η_{M38}	-0.02941	0.011056	[.008]	η_{138}	2.00E-03	2.58E-03	[.437]	η_{238}	-0.043505	7.06E-03	[.000]	η_{338}	8.25E-03	4.00E-03	[.039]
vehq	η_{M39}	-2.93E-04	5.42E-03	[.957]	η_{139}	-4.43E-03	1.52E-03	[.004]	η_{239}	-2.56E-03	2.04E-03	[.209]	η_{339}	4.09E-03	2.03E-03	[.044]
$\ln(\frac{p_{cd}}{M})$	β_{M1}	-0.014793	2.47E-03	[.000]	β_{11}	-0.015047	0.117074	[.898]	β_{22}	0.940334	0.177483	[.000]	β_{33}	0.013374	0.066622	[.841]
$\ln(\frac{p_{video}}{M})$	β_{M2}	-0.173199	0.023472	[.000]	β_{12}	-0.08976	0.090562	[.322]	β_{23}	-0.261338	0.078364	[.001]				
$\ln(\frac{p_{adm}}{M})$	β_{M3}	-0.061415	8.53E-03	[.000]	β_{13}	0.069485	0.086183	[.420]								
$\ln(\frac{p_{toy}}{M})$	β_{M4}	0.112879	0.015359	[.000]												
					α_1	-0.169521	0.025324	[.000]	α_2	-0.351694	0.030037	[.000]	α_3	0.045107	0.050315	[.370]

Appendix F: Histograms of price elasticity estimates

Figure F.1: Own-price elasticity estimates

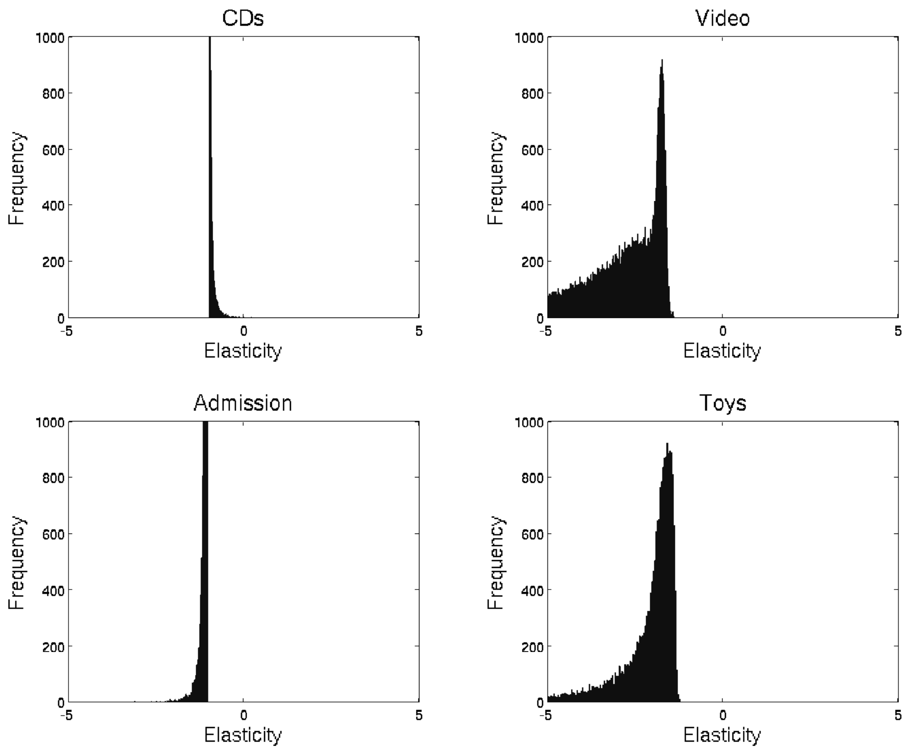

Figure F.2: Cross-price elasticity estimates