JOURNAL OF APPLIED ECONOMETRICS
J. Appl. Econ. **28**: 297–324 (2013)
Published online 5 October 2011 in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1002/jae.1269

MEASURING THE EFFECT OF NAPSTER ON RECORDED MUSIC SALES: DIFFERENCE-IN-DIFFERENCES ESTIMATES UNDER COMPOSITIONAL CHANGES

SEUNG-HYUN HONG*
Department of Economics, University of Illinois at Urbana-Champaign, IL, USA

SUMMARY

This paper measures the effect of Napster on record sales. I treat the introduction of Napster as a technological event that only Internet users experienced, and use a difference-in-differences (DD) approach. Because of potential compositional changes in Internet users, I examine identifying assumptions for the DD estimator under compositional changes and develop a test for identifying restrictions. To address potential bias due to compositional changes, I extend DD matching estimators to the case of two-variate propensity scores. I find evidence suggesting that file sharing is likely to explain 20% of total sales decline, which is driven by households with children aged 6–17. Copyright © 2011 John Wiley & Sons, Ltd.

Received 26 November 2009; Revised 6 July 2011

1. INTRODUCTION

Despite the extensive literature on the effect of file sharing on recorded music sales, the magnitude of this effect is still undetermined, partly because few studies have used representative samples of music buyers to estimate the effect, but also because previous empirical work[1] has suffered from little direct information on who downloaded music and how much their music expenditures have changed. As a result, most studies do not provide consistent estimates for the impact of file sharing on total record sales, though such estimates should be used to assess the actual damage (or benefits) of file sharing to the recorded music industry. Ideally, we wish to have panel data of representative random samples, in which individual music purchases are observed, while file sharing is exogenously available to part of the observations. Given lack of such ideal experimental data, this paper makes use of publicly available repeated cross-sectional data containing music expenditures for representative random samples of US households, and further exploits the introduction of Napster—the first file-sharing software widely used by Internet users—as a natural experiment.

In this respect, I begin with a difference-in-differences (DD) approach, treating the emergence of Napster as a technological event that only Internet users experienced. Changes in music expenditures after the introduction of Napster are attributed to a time effect and the effect of the presence of Napster, or simply, the effect of Napster.[2] Internet non-users are subject only to the time effect. The treatment group thus consists of Internet users, whereas Internet non-users belong to the control group. Using the

* Correspondence to: Seung-Hyun Hong, Department of Economics, University of Illinois at Urbana-Champaign, IL 61801, USA. E-mail: hyunhong@illinois.edu

[1] See, for example, Oberholzer-Gee and Strumpf (2007), Rob and Waldfogel (2006), and Zentner (2006). See also Liebowitz (2006a,b) and Oberholzer-Gee and Strumpf (2010) for a literature review on this topic.
[2] The 'treatment' in this paper is therefore the *presence* of Napster which includes not only *file sharing* via Napster butalso other *new online activities related to recorded music* that were present when Napster was available. Although I also attempt to isolate the effect of file sharing in this paper, I mostly focus on measuring the effect of the *presence* of Napster. See Sections 3 and 5.1 for more detailed discussion. Note also that I henceforth use 'the effect of Napster' to refer to 'the effect of the presence of Napster', because the latter term is rather cumbersome.

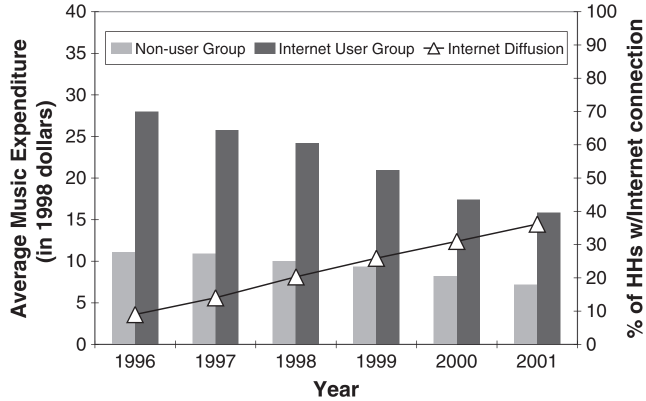

Figure 1. Internet diffusion and average quarterly music expenditure in the CEX

control group, the DD estimator then attempts to difference out the time effect and isolate the effect of the treatment.

However, an important challenge in applying the DD approach to the Napster example above is the presence of compositional changes, in that the treatment group may expand over time by including more diverse individuals. This possibility of compositional changes is shown in Figure 1, which plots the percentage of Internet users as well as changes in average music expenditures for Internet users and non-users from the Consumer Expenditure Survey (CEX). The decline in recorded music expenditures for the treatment group is accompanied by the diffusion of the Internet. This decline could result from the emergence of Napster in June 1999, but it is also plausible that more consumers with low reservation prices for music may have adoptedthe Internet over time, in which case the decline can simply reflect the compositional change in the treatment group, hence resulting in a negative bias in the conventional DD approach.

Therefore, we should take compositional changes into account in order to estimate the effect of Napster using the DD approach. However, few prior studies have attempted to address potential biases in the DD approach under compositional changes. Although various aspects of the DD methodology have been examined carefully and further improved by many studies, including several recent papers such as Abadie (2005), Athey and Imbens (2006), and Bertrand *et al.* (2004), most studies using the DD methods have implicitly or explicitly maintained the assumption of no compositional changes. Given lack of prior study, this paper thus attempts to relax this assumption, focusing on compositional changes only in terms of observables.[3]

To this end, I consider the identification restriction closely related to 'selection on observable' (Heckman and Robb, 1985) used in cross-sectional studies, and note that this restriction still allows for some forms of compositional changes.[4] For this identification restriction to be plausible, however, observed characteristics should contain rich information, suggesting conditioning on high-dimensional variables. Because high-dimensional matching is practically difficult, I instead consider the propensity

[3] The assumption of no compositional changes can be stated in terms of (i) time invariance of the distribution of observables within groups (see Assumption 2 in Section 5.2 of this paper), or (ii) time invariance of unobservables withingroups (see Assumption 3.3 in Athey and Imbens, 2006). Only the assumption (i) is relaxed in this paper. I appreciate a referee for making this subtle point.

[4] This point is not difficult to show, but it has not been discussed in the literature, presumably because compositional changes are normally not considered. See Section 3 for a simple example to illustrate this point. Section 5.2 provides more precise definition and further discussion of no compositional changes in terms of observables.

DOI: 10.1002/jae

score (PS) matching as a dimension reduction method. However, the conventional one-dimensional PS may not be sufficient to identify the treatment effect under compositional changes. Accordingly, I examine related identification issues and find evidence suggesting that at least two-dimensional PS is required to address compositional changes. To specifically estimate the treatment effect, I consider the DD matching (DDM) method developed by Heckman *et al.* (1997, 1998), and extend their estimators to the case of two-dimensional PS.

Applying the DDM method to the CEX data, I find that the presence of Napster had on average reduced the quarterly music expenditure for a household with Internet access during the Napster period[5] by \$1.45 (7.6%). Aggregating this number to the population of US households explains approximately 40% of the total record sales decline during that period. About half of this decline is driven by households with children aged 6–17, which is precisely estimated. The other half of this decline is explained by households aged 15–34, but the DDM estimate for this group is not precisely estimated. Because the effect of Napster may not reflect solely the sheer effect of file sharing, I further attempt to isolate the effect of actual music downloading, using a complementary dataset. I find that the DDM estimate for those aged 15–34 is less likely to represent the effect of music downloading, whereas the DDM estimate for those with children aged 6–17 is unlikely to be confounded with other new online activities during the Napster period. These results therefore suggest that file sharing is likely to explain about 20% of the total sales decline during the Napster period, mostly driven by downloading activities of households with children aged 6–17.

These findings contribute to the general empirical literature on the substitution of recent digital technologies and the Internet for existing offline activity,[6] as well as the literature on online piracy. Despite the vigorous debate on the effect of file sharing, few studies have measured the effect of file sharing on total record sales using nationally representative micro-level data on changes in music expenditures. In this paper, I use the CEX data to provide such estimates and further identify the demographic groups most responsible for the sales decline due to file sharing.

The rest of the paper is organized as follows. Section 2 provides a brief background on Napster and recorded music sales. Section 3 presents a simple example that illustrates a problem of compositional changes and my approach to address the problem. Section 4 describes the data. Section 5 formally defines the main parameter of interest and examines identifying assumptions. This section then discusses the DDM estimators. Section 6 presents the main estimation results. This section also reports the estimates from alternative approaches for the purpose of comparison, and further provides the results using a complementary dataset. Section 7 concludes the paper.

2. BACKGROUND ON NAPSTER AND RECORD SALES

Systematic file sharing began with Napster. After its introduction in June 1999, Napster quickly became popular among Internet users. The number of users grew extraordinarily, and numerous music files were exchanged via Napster.[7] Although other minor file-sharing programs appeared during the Napster period, Napster was undoubtedly the dominant file-sharing service until early 2001.[8] For this

[5] Throughout this paper, *pre-Napster* refers to June 1997 to May 1999, while *post-Napster* or *Napster period* refers to June 1999 to June 2001, the period in which Napster was operating.

[6] See, for example, Hong and Wolak (2008) and Sinai and Waldfogel (2004). See also Gentzkow (2007) for more references.

[7] According to *Newsbytes*, 20 July 2000, citing Napster's report on its membership, the number of Napster users grew from 1 million in November 1999 to 10 million in late April 2000, to 15 million in mid June, and to 20 million in mid July 2000. Romer (2002) cites Webnoize, a web consulting firm, estimating that Napster members exchanged 2.8 billion files in February 2001.

[8] There were a few other file-sharing programs during the Napster period. Two centralized file-sharing systems were launched in 2000: *Scour* in April and *Aimster* in August. Similarly to the Napster case, the recording industry's legal actions against these companies shut down Scour in late 2000 and Aimster in 2002. KaZaA, which took the place of Napster after it was closed, was developed in late 2000. See Chapter 3 in Fisher (2004) for more details on a variety of file-sharing systems.

DOI: 10.1002/jae

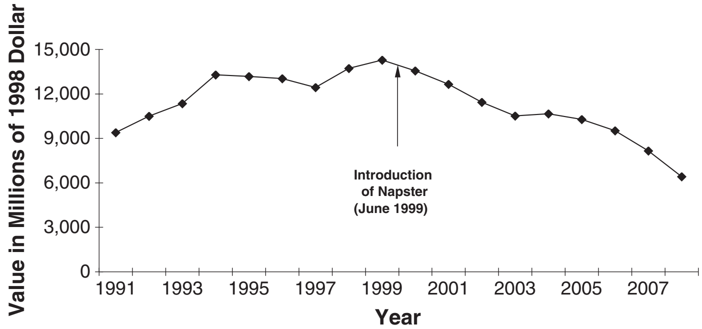

Figure 2. Total real value of record shipments in the USA. Refer to the RIAA's year-end statistics. Total sales include CDs, cassettes, LPs, and music videos. Starting from 2004, total sales also include digital formats such as legitimate download

reason, as well as lack of further information, I do not distinguish file sharing via Napster from file sharing via other programs during the Napster period.

Napster allowed its users to share a variety of individual songs, thereby providing access to unbundled songs, as opposed to bundled albums, in addition to providing songs for free. This is likely to have changed consumers' expenditures on recorded music, but the direction and magnitude of this change are unclear and are likely to differ across consumers. Zero prices could lead consumers to substitute CD purchases with music downloading, hence reducing their music expenditures. On the other hand, inexpensive access to individual songs provides consumers with better information on a variety of music, which could increase consumers' music expenditure. Moreover, considerable heterogeneity in music preferences and in the costs associated with downloading implies that consumers responded differently to the presence of Napster.

Nevertheless, this event coincided with the start of the ongoing slump in recorded music sales. According to the Recording Industry Association of America (RIAA), the total real value of shipments in the USA had reached its peak of $14,270 millionin 1999. After Napster appeared, the total real value of record sales decreased by 5% in 2000, 6.7% in 2001 and 9.6% in 2002, and continued to decline through the 2000s (see Figure 2). Accordingly, the recording industry concluded that this decline was largely a result of file sharing. Subsequent legal action by the recording industry based on these grounds succeeded in closing Napster in 2001 and other file-sharing services later in the 2000s.

This coincidence, nonetheless, does not substantiate the negative impact of Napster and file sharing on record sales, nor does it provide any magnitude of the impact. Despite the apparent negative correlation between file sharing and record sales, it is unclear to what extent the sales decline is attributable to file sharing. There are a variety of factors, other than file sharing, that can also account for the recent slump in record sales. For example, some entertainment goods might be substitutes or complements for recorded music, and changes in relative prices for these goods could have resulted in the decline in recorded music demand. Another example includes the rapid penetration of the Internet which might have led Internet users to spend more time on the Internet and less time in listening to music, thus decreasing demand for recorded music.[9]

In this regard, the literature on online piracy has vigorously debated the effect of file sharing on music sales in recent years. The majority of empirical studies find a negative effect, while a handful of

[9] Hong (2007b) reports that the consumer price index for recorded music has slightly increased between 1998 and 2004, whereas the consumer price indices for videos and toys/games have declined considerably during the same period, suggesting that consumers may have spent less on recorded music due to changes in relative prices. See Hong (2007b) for further investigationon other possibilities that might have contributed to the recent slump in record sales.

DOI: 10.1002/jae

papers find a non-negative effect of file sharing on music sales (e.g. Oberholzer-Gee and Strumpf, 2007). Nonetheless, the magnitude of this loss is still unclear, since few studies provide a consistent estimate of the effect of file sharing on total record sales based on nationally representative samples of music buyers. Moreover, most studies do not investigate underlying mechanisms based on demographic compositions, i.e. which demographic groups are responsible for how much of the sales decline due to file sharing. In contrast, I use a publicly available dataset containing music expenditures of representative random samples of the US households and further investigate underlying mechanisms based on demographic compositions.

3. EMPIRICAL STRATEGY

The basic approach in this paper is difference-in-differences. To the extent that (i) Internet users are comparable to Internet non-users and (ii) there are no compositional changes between Internet users and non-users, a time effect can be captured by changes in music expenditures of non-users, while the difference in music expenditures between Internet users and non-users during the pre-Napster period can reflect the general effect of using the Internet as well as time-invariant unobserved components of music expenditures correlated with using the Internet.

Under compositional changes, however, these usual intuitions on the validity of the DD would fail. First, if there is selection on unobservables fixed through time, we cannot difference out time-invariant unobserved components of music expenditures correlated with using the Internet. As a result, selection on unobservables cannot be allowed under compositional changes, unless more complex modeling is pursued. Second, under the assumption of selection on observables, the difference in music expenditure between Internet users and non-users during the pre-Napster period serves as an estimate of the general effect of using the Internet. We can thus recover the effect of Napster by estimating how the difference in music expenditures between Internet users and non-users changes after the introduction of Napster, i.e. by using DD.

However, even under the assumption of selection on observables, the observed composition of total Internet users can change over time as more diverse consumers adopt the Internet due to the diffusion process. For example, later adopters might include moreconsumers with older ages and lower income, in which case average Internet users in the post-Napster period might spend less on recorded music than average Internet users in the pre-Napster period, not because their willingness to pay for music has declined, but because more consumers with lower willingness to pay have adopted the Internet over time. To illustrate this issue, consider the following example with two periods: pre-Napster and post-Napster.

There are four types of consumers characterized by X, where X includes age and employment. First, 'young and employed' consumers have high propensity to adopt the Internet, and 8 out of 10 had Internet access in both periods. They also spend \$30 on CDs in each period. The second type is 'young and unemployed' consumers who have moderate propensity to adopt the Internet, and 6 out of 10 had Internet access in both periods. Because they are unemployed but still young, they spend \$20 on CDs in each period. Third, 'old and employed' consumers had low propensity to adopt the Internet in the first period, and so only 4 out of 10 had Internet access. However, the Internet became more accessible over time, so that 8 out of 10 had Internet access in the second period. In addition, these consumers spend \$10 on CDs in each period. Lastly, 'old and unemployed' consumers do not buy recorded music at all, and 4 out of 10 had Internet access in the first period,while 6 out of 10 had Internet access in the second period. This example is illustrated below, where D denotes a dummy for Internet access and T denotes a time dummy that is equal to 0 for the pre-Napster period and 1 for the post-Napster period. Additionally, Y means music expenditure, and N is the number of consumers.

DOI: 10.1002/jae

X		$T=0$		$T=1$	
		Y	N	Y	N
$D=1$	Young and employed	\$30	8	\$30	8
	Young and unemployed	\$20	6	\$20	6
	old and employed	\$10	4	\$10	8
	Old and unemployed	\$0	4	\$0	6
$D=0$	Young and employed	\$30	2	\$30	2
	Young and unemployed	\$20	4	\$20	4
	Old and employed	\$10	6	\$10	2
	Old and unemployed	\$0	6	\$0	4

In this example, no one changed their music expenditure. That is, an Internet user in the post-Napster period would continue to spend the same amount on recorded music even in the absence of Napster. Hence the effect of Napster on music expenditure should be zero. However, if we use the conventional DD approach, which assumes that the composition of Internet users remained the same in both periods, the estimated effect is given by

$$E(Y|D=1,T=1)-E(Y|D=1,T=0)-E(Y|D=0,T=1)+E(Y|D=0,T=0)\approx -\$4.69$$

which is clearly an incorrect estimate of the effect of Napster. To fix the problem, note that given X, music expenditure is the same, regardless of D and T. Hence, if we condition on X, the conditional DD estimatesare computed by

$$E(Y|D=1,T=1,X)-E(Y|D=1,T=0,X)-E(Y|D=0,T=1,X)+E(Y|D=0,T=0,X)$$

which is zero for each value of X. Calculating the weighted mean of the conditional DD estimates then yields that the estimated effect of Napster is zero.

The example shows that the conventional DD approach breaks down under compositional changes, but it also suggests that separating different types of consumers might address potential biases from compositional changes. In practice, however, separating different types would require rich information on observed characteristics. Hence, for X to be sufficient for identification, X is likely to be high-dimensional. Since matching based on high-dimensional X can be difficult, Iconsider the propensity score matching as a dimension reduction method in my application.

However, the conventional one-dimensional PS may not be sufficient to separate different types of consumers under compositional changes. To illustrate the issue as well as a potential solution, reconsider the example above and note that the PS, i.e. the probability for Internet access given X, is given as follows.

X	PS for $T=0$	PS for $T=1$	PS for both periods
Young and employed	0.8	0.8	0.8
Young and unemployed	0.6	0.6	0.6
Old and employed	0.4	0.8	0.6
Old and unemployed	0.4	0.6	0.5

If we use only the PS for $T=0$, old and employed consumers would be matched with old and unemployed consumers. Similarly, if we condition only on the PS for $T=1$ (or both periods), young and unemployed consumers would be matched with old and unemployed (or old and employed) consumers. As a result, if we condition on the one-dimensional PS, the conditional DD estimates may not be equal to zero. In contrast, if we use two propensity scores, then we can separate all four types of consumers. This idea is further formalized in Section 5.

Note that this paper focuses on identifying and estimating the effect of Napster. However, one might still wish to isolate the effect of file sharing from the effect of other new online activities during the Napster period. In this regard, after I present the main results in Section 6, I further attempt to estimate the effect of file sharing by using a complementary dataset. Section 6.4 describes my method and reports the results.

4. DATA

4.1. Data Description

The primary source of data in this paper is the 1996–2002 Interview surveys of the Consumer Expenditure Survey (CEX) by the US Bureau of Labor Statistics. The CEX is publicly available and consists of random samples of households designed to be representative of the total US population. It is a repeated cross-section with a rotating panel structure. Because of several problems, however, I do not exploit this limited panel structure.[10] The unit of my analysis is quarterly expenditures of US households. The CEX is useful for my purposes because it contains a rich set of demographic variables as well as detailed data on various expenditures, including recorded music and Internet service fees. Recorded music expenditures are defined to be the sum of expenses on CDs, tapes and LPs purchased.[11] For Internet access, I use two pieces of information. The first is computer information service expenditures which consist mainly of Internet service fees. The second is whether the household is living in acollege dormitory. The CEX identifies students living in college dormitories as separate households from their parents. It is highly likely that most college dormitories already had broadband connections in the late 1990s. Consequently, I define an Internet user group as households that either spent positive amounts on computer information service or were living in a college dormitory.

4.2. Descriptive Statistics

Descriptive statistics organized by Internet adoption and year are presented in Table I.[12] The table shows substantial differences between the Internet user group and the non-user group. Internet users are younger, richer, more educated and likely to live in urban or more populated areas. Moreover, Internet users tend to spend more money on recorded music and entertainment goods. In particular, about 80% of Internet non-users did not purchase recorded music, and most did not spend on entertainment either. Hence the non-user group does not appear to be comparable to the user group.

---

[10] See Hong (2007a) for these problems and more detailed discussion on the CEX.

[11] The weighted sum of recorded music expenditures from the CEX is about \$5.6 billion in 1997, \$5.7 billion in 1998, \$5.4 billion in 1999, \$4.7 billion in 2000 and \$4.3 billion in 2001. For comparison, the RIAA reports that the total real value of shipments was \$12.4 billion in 1997, \$13.7 billion in 1998, \$14.3 billion in 1999, \$13.5 billion in 2000 and \$12.6 billion in 2001. Thus the value of total record sales from the CEX is approximately 40% of the value reported by the RIAA. The CEX tends to underestimate the total value of expenditures compared to the national accounts such as the Personal Consumption Expenditure, an aggregate time series for US consumer expenditures estimated by the Bureau of Economic Analysis. This has been noted by Battistin (2003) and references therein. One possibility of this underestimation is a recall problem. That is, survey respondents often forgot their purchases. The other is that the CEX surveys only households, so that expenditures from institutions including the government, businesses, libraries and radiostations are not included in the CEX. Note that this problem of underestimated expenditures is not critical to my analysis, because this problem is unlikely to be related to Internet access or file sharing. Moreover, the trend of record sales from the CEX is closely related to that from the RIAA.

[12] Years in this paper refer to the period from June of the year through May of the next year, in order to conveniently separate the pre-Napster period—June 1997 to May 1999—and the post-Napster period—June 1999 to June 2001, during which Napster was operating.

DOI: 10.1002/jae

Table I. Descriptive statistics for Internet user and non-user groups

	2000		1998		1999		2000	
	Internet user	Non-user	Internet user	Non-user	Internet user	Non-user	Internet user	Non-user
Average expenditure								
Recorded music	$25.73	$10.90	$24.18	$9.97	$20.92	$9.37	$17.42	$8.22
Entertainment	$195.03	$96.71	$193.38	$84.92	$182.42	$80.19	$164.88	$71.44
Zero expenditure								
Recorded music	0.56	0.79	0.60	0.80	0.64	0.81	0.68	0.83
Entertainment	0.08	0.32	0.09	0.35	0.14	0.39	0.17	0.44
Demographics								
Age	40.2	49.0	42.3	49.0	44.1	49.4	44.3	49.9
Income	$52,887	$30,459	$51,995	$28,169	$49,970	$26,649	$47,510	$26,336
High school grad.	0.18	0.31	0.17	0.32	0.21	0.32	0.22	0.33
Some college	0.37	0.28	0.35	0.27	0.34	0.27	0.36	0.27
College grad.	0.43	0.21	0.45	0.21	0.42	0.20	0.37	0.20
Manager	0.16	0.08	0.16	0.08	0.14	0.08	0.14	0.07
Professional	0.23	0.11	0.22	0.10	0.21	0.10	0.19	0.10
Living in a dorm.	0.12	0	0.08	0	0.05	0	0.05	0
Urban	0.93	0.87	0.93	0.86	0.91	0.87	0.89	0.86
Inside an MSA	0.84	0.78	0.83	0.78	0.83	0.78	0.81	0.78
Pop. size > 4 million	0.34	0.26	0.30	0.26	0.31	0.25	0.28	0.25
Appliance ownership								
Computer	0.79	0.27	0.81	0.28	0.80	0.28	0.81	0.32
Sound system	0.81	0.57	0.79	0.58	0.78	0.56	0.76	0.56
VCR	0.83	0.72	0.86	0.74	0.86	0.72	0.85	0.72
Total households (in millions)	15	91	22	86	28	80	34	76
Observations	3,163	19,052	5,624	21,550	8,191	22,810	9,606	20,919

Note: All the statistics are weighted using the weights provided by the CEX. Years refer to the period from June of the year to May of the next year. Total households are computed by summing the CEX weights.

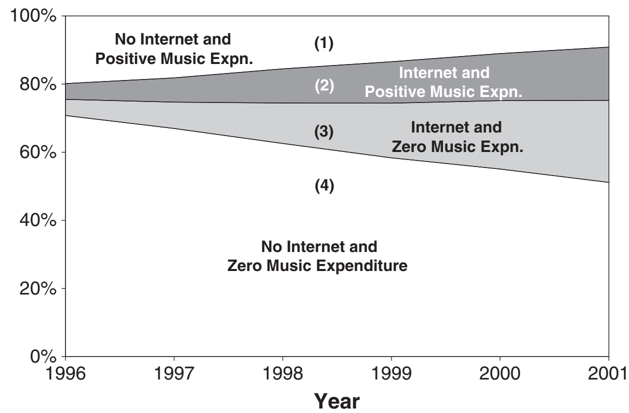

Figure 3. Percentage of households in the CEX by Internet adoption and music expenditure

Table I also documents significant compositional changes between the two groups over time. For many technologies, it is common that early adopters tend to represent the small fraction of the population that is technologically savvy, whereas later adoptershave more diverse demographic and economic characteristics. Similar patterns are also observed in the diffusion of the Internet. Internet users in 1997, for example, are likely to be younger, richer and more educated than those in 2000. Furthermore, later adopters, say in 1997, are likely to be included in the Internet user group, say in 2000. Therefore, pre-Napster Internet users are different from post-Napster users. This difference is important because later adopters may include more consumers with a lower willingness to pay for recorded music. Similar to the example in Section 3, a negative effect estimated from the conventional DD approach may simply reflect this compositional change due to the diffusion of the Internet.

The compositional change is further illustrated in Figure 3. The figure plots the percentages of households in the CEX by Internet adoption and whether they spent any money on recorded music. For example, area (3) denotes households with an Internet connection who spent nothing on recordings. More households adopted the Internet over time, so that area (2) + (3) becomes larger. However, the percentage of households who spent nothing on music (area (3) + (4)) has increased little, and more non-music buyers adopted the Internet over time. As a result, the post-Napster Internet user group includes more households with low reservation prices for recorded music than the pre-Napster user group does. This shows one reason why the decrease in the average music expenditure for the Internet user group may have nothing to do with the presence of Napster.

In the next section, I develop a methodology to address the negative bias due to compositional changes and quantify the effect of Napster on changes in music expenditures.

5. ECONOMIC FRAMEWORK

5.1. The Main Parameters of Interest

In this paper, the treatment refers to new online activities during the Napster period that affected music expenditures. Because file sharing via Napster was the most visible among all the new activities on the Internet that affected music expenditures during the Napster period, I term this treatment the presence of Napster. However, not all Internet users used file sharing, and other new online activities related to

DOI: 10.1002/jae

music expenditures might have started around the same time that Napster was introduced.[13] Hence the effect of the presence of Napster, or simply the effect of Napster, includes not only the sheer effect of music downloading via Napster but also the effect of other new online activities relatedto music expenditures. Estimating the effect of Napster, nevertheless, is still of interest because it allows us to understand the impact of new digital technology, including file sharing and other new Internet technology, on the sales of copyrighted product such as recorded music.

To formally define the effect of Napster, let D denote a dummy variable for Internet access, and let T be an indicator variable for the Napster period. That is, $T=1$ if household i is observed in period a, while $T=0$ if it is observed in period b.[14] I also use R to indicate whether observation i receives the treatment. Let $Y_{1,a,r,i}$ denote music expenditures of Internet users in period a in the presence of Napster, where subscript r indicates $R=1$. Using the subscript $\bar{r}$ to denote $R=0$, $Y_{1,a,\bar{r},i}$ then representsthe counterfactual music expenditures of Internet users in period a, had they not received the treatment. Because Internet users in period b and non-users could not receive the treatment, I drop the subscript $\bar{r}$ and use $Y_{1,b,i}$ and $Y_{0,b,i}$ (or $Y_{0,a,i}$) to respectively denote music expenditures of Internet users in period b and music expenditures of non-users in period b (or in period a). To simplify the exposition, I suppress the subscript i below.

Given the notations above, the observed music expenditure Y can be written as

$$\begin{aligned} Y &= Y_{1,a,r}DTR + Y_{1,a,\bar{r}}DT(1-R) + Y_{0,a}(1-D)T + Y_{1,b}D(1-T) + Y_{0,b}(1-D)(1-T) \\ &= \left(Y_{1,a,r} - Y_{1,a,\bar{r}}\right)DTR + \left(Y_{1,a,\bar{r}} - Y_{1,b} - \left(Y_{0,a} - Y_{0,b}\right)\right)DT + \left(Y_{0,a} - Y_{0,b}\right)T \\ &\quad + \left(Y_{1,b} - Y_{0,b}\right)D + Y_{0,b} \end{aligned}$$

where $Y_{1,a,r} - Y_{1,a,\bar{r}}$ reflects the individual treatment effect. The main parameter of interest, M, is then the average of the individual treatment effects, defined as

$$M = E\left(Y_{1,a,r} - Y_{1,a,\bar{r}} \middle| D = 1, T = 1\right) \tag{1}$$

Note that I consider the average effect as the main parameter of interest, partly because the mean effect of 'treatment on the treated'[15] is one common parameter to be estimated (see, for example, Heckman *et al.*, 1999), but also because the mean effect M in (1) can be easily translate into the effect on total record sales, since the CEX data are nationally representative random samples.

5.2. Identifying Assumption

In general, M cannot be identified without any assumption, because $Y_{1,a,\bar{r}}$ is not observed for Internet users during the Napster period. For this reason, I begin by assuming that the counterfactual music expenditure $Y_{1,a,\bar{r}}$ for household i using the Internet in period a should be equal to its music expenditure from using the Internet in period b and the time effect that it would experience without Internet access. In other words, I assume that $Y_{1,a,\bar{r}} = Y_{1,b} + \left(Y_{0,a} - Y_{0,b}\right)$. As a result, the observed music expenditure Y can be rewritten as

$$Y = \left(Y_{1,a,r} - Y_{1,b} - \left(Y_{0,a} - Y_{0,b}\right)\right)DTR + \left(Y_{0,a} - Y_{0,b}\right)T + \left(Y_{1,b} - Y_{0,b}\right)D + Y_{0,b}$$

[13] For example, in January 2000, Half.com launched a person-to-person marketplace where consumers could buy and sell various items including used CDs. Facing lowered prices from online secondary markets, Internet users could reduce their music expenditures.

[14] Period b (or a) refers to the pre-Napster (or post-Napster) period, i.e. 2 years *b*efore (or *a*fter) the introduction of Napster.

[15] One may be interested in 'treatment on the untreated', but its identification is not necessarily implied by the identifying assumptions in Section 5.2. Hence it is not considered in this paper.

Conditional on observed characteristics X, I can then rewrite M as

$$E(M|X) = E(Y|D = 1, T = 1, X) - E(Y|D = 1, T = 0, X) - [E(Y|D = 0, T = 1, X) \\ - E(Y|D = 0, T = 0, X)] + B$$

where $B = E(Y_{1,b}|D = 1, T = 0, X) + [E(Y_{0,a}|D = 0, T = 1, X) - E(Y_{0,b}|D = 0, T = 0, X)] - E(Y_{1,b} + Y_{0,a} - Y_{0,b}|D = 1, T = 1, X)$. The equation above suggests that M can be estimated by using the difference-in-differences approach, provided that $B = 0$.

Therefore, the key assumption maintained in this paper is as follows.

Assumption 1. $\mathbf{Y} \perp (D, T)|X$

where $\mathbf{Y} = (Y_{1,b}, Y_{0,a}, Y_{0,b})$. I consider Assumption 1, since it implies that $B = 0$, i.e.

$$E(Y_{1,b} + Y_{0,a} - Y_{0,b}|D = 1, T = 1, X) = E(Y_{1,b}|D = 1, T = 0, X) \\ + E(Y_{0,a}|D = 0, T = 1, X) - E(Y_{0,b}|D = 0, T = 0, X) \quad (2)$$

Note that Assumption 1 is closely related to 'selection on observables' (Heckman and Robb, 1985), or unconfoundedness (Rosenbaum and Rubin, 1983) used in cross-sectional studies, and that the assumption based on (2) is used in the economic literature on program evaluation using difference-in-differences (e.g. Heckman *et al.*, 1997, 1998; Heckman and Smith, 1999; Abadie, 2005). Assumption 1 is also motivated by the detailed demographic information in the CEX. Upon conditioning on rich demographic variables in the CEX, including age, race, education, appliance ownership (e.g. video tape or disc players, sound component system), occupation, family composition, work, housing, income, and geographic information, most determinants of the Internet adoption are unlikely to be correlated with recorded music expenditures.

Although most of my analysis requires only the condition in (2), I maintain Assumption 1 because this assumption is more commonly used in the literature than the condition in (2).[16] More importantly, Assumption 1 still allows for there to be compositional changes. Specifically, Assumption 1 does not imply no compositional changes defined as follows.

Assumption 2. No compositional changes $T \perp (X, D)$

which is equivalent to (i) $X \perp T|D$ and (ii) $D \perp T$, i.e. (i) the observed characteristics of households within a given group do not change over time, and (ii) whether a household is in the treatment group does not depend on time. One can easily check that Assumption 1 implies neither (i) nor (ii). Moreover, even if Assumption 2 does not hold, Assumption 1 can be true. Recall the example in Section 3. We can easily check that Assumption 2 is violated, while Assumption 1 can be still allowed in that example. Hence Assumption 1 can be used to address potential biases in the DD approach under compositional changes.

Three caveats are in order, however. First, I acknowledge that Assumption 1 cannot address all forms of biases under more general forms of compositional change, particularly in terms of unobservables, but investigating more general issues is beyond the scope of this paper. Similarly,

[16] Although the condition in (2) is more general than Assumption 1, this condition obscures any economic meaning, and it is unclear how much gain might be obtained from this additional degree of generality. As a result, I maintain Assumption 1 as the main assumption in this paper.

I acknowledge that my data and the method proposed in this paper cannot address potential bias due to 'selection on unobservables'.[17] Lastly, given that many non-users did not purchase recorded music (see Section 4.2), an overall downward trend in music expenditures may not be fully reflected in music expenditures of non-users,[18] in which case the DD estimates would overestimate a potential negative effect. This censoring issue is partly related to the problem of compositional changes discussed in Section 4.2, but I acknowledge that the method used in this paper cannot fully address the censoring issue.

5.3. Identification under Compositional Changes

To make use of Assumption 1, we need to match directly based on X. Because of high dimensionality in X, the standard approach is to rely on the result from Rosenbaum and Rubin (1983) and match based on a one-dimensional propensity score (PS).[19] I follow this literature and use the PS matching for a dimension reduction method. Because most studies in the literature do not consider compositional changes, however, I extend the standard PS matching to the DD models under compositional changes. Specifically, I first consider the key restrictions that allow for the PS matching in the DD models. These restrictions, which are equivalent to Assumption 3 below, are satisfied under no compositional changes, which is proved below. However, if compositional changes are allowed, it is unclear whether these restrictions would hold. Nevertheless, these restrictions do not necessarily imply the absence of compositional changes. My approach is therefore to check whether they are consistent with the data. Although we might suspect potential compositional changes, we may still use PS matching, as long as these restrictions are plausible.

Let us begin with the following assumption.

Assumption 3. $X \perp (D, T) | P$

where P denotes the probability of using the Internet conditional on X. Normally, P is the one-dimensional PS, i.e. $P \equiv \Pr(D = 1|X)$, but in this paper P can be two-dimensional as well, in which case $P \equiv (\Pr(D = 1|T = 0, X), \Pr(D = 1|T = 1, X))$. Note also that P is treated as a known variable in this section. In my application, the propensity score is parametrically estimated. Specifically, I assume the following model:

$$D = 1 \quad \text{if} \quad D^* = X\mu + v \geq 0; \quad D = 0, \quad \text{otherwise}$$

[17] As a referee suggested, after Napster emerged, those with stronger preference for free music would likely adopt the Internet earlier. Hence, under this form of selection on unobservables, the DDM estimate discussed in Section 5.5 will be larger (in the absolute value) for the earlier period than for the later period. To check this, I separate the first year and the second year during the Napster period, and estimated the DDM estimates separately for these two years, using the sample of households with children aged 6–17. I find that the DDM estimates (and their standard errors) for the earlier period and for the later period are respectively − 2.42 (0.94) and − 3.45 (0.72), suggesting that this form of selection on unobservables might not be serious. Following a suggestion from another referee, I also estimated a similar DDM estimate using only the pre-Napster period, treating the first year as period b and the second year as period a. The estimate is − 0.57 (0.81), suggesting that variation in group composition is unlikely to be driven by selection on unobservables. Nonetheless, I acknowledge that these pieces of evidence do not provide sufficient evidence against selection on unobservables. Fully addressing this issue, however, is beyond the scope of this paper.

[18] Note that this paper implicitly assumes theStable-Unit-Treatment-Value-Assumption (Rubin 1978); i.e. the presence of Napster should not affect the music expenditures of Internet non-users. This assumption might be violated if the presence of Napster led recording firms to change the prices ofCDs, hence affecting non-users' music expenditures. However, this is unlikely, because the prices of recorded music did not change much during the Napster period (see Hong, 2007b), and most non-users do not appear to be music buyers.

[19] The PS matching has been further developed theoretically by several studies, including Hahn (1998) and Hirano *et al*. (2003), and has been applied to various empirical studies (e.g. Dehejiaand Wahba, 1999; Lechner, 1999). In particular, Abadie (2005) and Heckman *et al*. (1997, 1998) extend the PS matching methods to semi-parametric and nonparametric DD models. See also Imbens and Wooldridge (2009) for anexcellent literature review on the PS matching as well as matching methods in general.

where μ is a vector of parameters, v is the error term, and D^* denotes the net benefit of adopting the Internet, so that if $D^* \geq 0$, then households would adopt the Internet.[20] Because of a potential concern that this parametric model might be misspecified, however, I also perform a specification test in Section 6.1.

Assumption 3 is equivalent to the following restrictions conditional on the PS:

$$f(X|D=1,T=1,P)=f(X|D=0,T=1,P) \tag{3}$$

$$f(X|D=1,T=1,P)=f(X|D=1,T=0,P) \tag{4}$$

$$f(X|D=1,T=0,P)=f(X|D=0,T=0,P) \tag{5}$$

where $f(\cdot)$ denotes a probability density function. As long as the restrictions (3), (4) and (5) hold, we can rewrite the condition in (2) in terms of the PS, instead of high-dimensional X, which is established by the following proposition.

Proposition 1. If Assumptions 1 and 3 hold, then it follows that

$$\begin{aligned} E\left(Y_{1,b}+Y_{0,a}-Y_{0,b}\middle|D=1,T=1,P\right) &= E\left(Y_{1,b}\middle|D=1,T=0,P\right)+E\left(Y_{0,a}\middle|D=0,T\right.\\ &\left.=1,P\right)-E\left(Y_{0,b}\middle|D=0,T=0,P\right) \end{aligned} \tag{6}$$

Proof. Let $F(\cdot)$ denote a cumulative distribution function. Assumptions 1 and 3 imply that

$$\begin{aligned} E\left(Y_{1,b}+Y_{0,a}-Y_{0,b}\middle|D=1,T=1,P\right) &= \int E\left(Y_{1,b}+Y_{0,a}-Y_{0,b}\middle|D=1,T=1,P,X\right)dF(X|D=1,T\\ &=1,P) = \int \left\{E\left(Y_{1,b}\middle|D=1,T=0,P,X\right)+E\left(Y_{0,a}\middle|D=0,T\right.\right.\\ &\left.\left.=1,P,X\right)-E\left(Y_{0,b}\middle|D=0,T=0,P,X\right)\right\}dF(X|D=1,T\\ &=1,P) = \int E\left(Y_{1,b}\middle|D=1,T=0,P,X\right)dF(X|D=1,T\\ &=0,P) + \int E\left(Y_{0,a}\middle|D=0,T=1,P,X\right)dF(X|D=0,T\\ &=1,P) - \int E\left(Y_{0,b}\middle|D=0,T=0,P,X\right)dF(X|D=0,T\\ &=0,P) = E\left(Y_{1,b}\middle|D=1,T=0,P\right)+E\left(Y_{0,a}\middle|D=0,T\right.\\ &\left.=1,P\right)-E\left(Y_{0,b}\middle|D=0,T=0,P\right) \end{aligned}$$

where Assumption 1 leads to the second equality, and Assumption 3 implies the third equality. □

Proposition 1 thus shows that we can use PS matching to identify $Y_{1,a,\bar{r}}$ for post-Napster Internet users, and that Assumption 3 provides key restrictions. However, these restrictions have not been discussed in the literature, presumably because the literature implicitly assumes no compositional changes. In particular, the restrictions in Assumption 3 are satisfied under no compositional changes. To see this, recall that Assumption 2 assumes no compositional changes, which states that $T \perp (X, D)$. Hence the propensity to adopt the Internet would not change over time. The following lemma immediately follows from Assumption 2.

[20] The error term v reflects unobserved determinants such as network effects and learning. These determinants are mostly related to technology diffusion, and thus they are unlikely to be directly related to music expenditures. As a result, even conditional on X, some consumers may adopt the Internet for a reason not related to music expenditures, while others with the same X may not adopt the Internet. Accordingly, both the treatment group and the control group are likely to be observed conditional on X. This is consistent with the common support condition, which is further examined empirically in Section 6.1.

DOI: 10.1002/jae

Lemma 1. Define $P_{\text{all}} \equiv \Pr(D=1|X), P_b \equiv \Pr(D=1|T=0,X)$ and $P_a \equiv \Pr(D=1|T=1,X)$. Suppose that Assumption 2 holds. Then, we have that $P_{\text{all}} = P_b = P_a$.

Given Lemma 1, we now can show that no compositional changes imply Assumption 3, which is established by the following lemma.

Lemma 2. Suppose that Assumption 2 holds. Then, Assumption 3 also holds.

Proof. Given Lemma 1, let $P = P_{all} = P_b = P_a$. It follows that

$$
\begin{aligned}
f(X|D=1,T=1,P) &= \frac{f(X,D=1|T=1,P)}{\Pr(D=1|T=1,P)} \\
&= \frac{\Pr(D=1|T=1,P,X)f(X|T=1,P)}{\Pr(D=1|T=1,P_a)} \\
&= \frac{P_a f(X|T=1,P)}{P_a} = f(X|T=1,P)
\end{aligned}
$$

Similarly, we obtain $f(X|D=0,T=1,P)=f(X|T=1,P)$. Hence $f(X|D=1,T=1,P)=f(X|D=0,T=1,P)$. Similar results hold for $T=0$. Therefore, $f(X|D=1,T=0,P)=f(X|D=0,T=0,P)$. Lastly, Assumption 2 means that $f(X|D,T)=f(X|D)$, which implies that $f(X|D=1,T=1)=f(X|D=1,T=0)$ and $f(P|D=1,T=1)=f(P|D=1,T=0)$. Therefore, $f(X|D=1,T=1,P) = \frac{f(X,P|D=1,T=1)}{f(P|D=1,T=1)} = \frac{f(X,P|D=1,T=0)}{f(P|D=1,T=0)} = f(X|D=1,T=0,P)$. □

Lemma 2 and Proposition 1 therefore show that if the composition of the treatment group and the control groups does not change over time, we can match based on the PS instead of X. However, if we consider the diffusion of the Internet over time, Assumption 2 is clearly violated, because as more Internet non-users adopt the Internet over time the treatment group (i.e. Internet user group) would include more diverse households with different X. That is, X is not independent of T conditional on D, and the propensity to adopt the Internet would change over time. Therefore, Lemma 1 does not hold, so that P_{all}, P_b, and P_a need not equal one another. Hence, if we let $P = P_{xall}$, then $\Pr(D=1|T,P)$ is not necessarily equal to $\Pr(D=1|T,X)$, in which case $f(X|D=1,T,P)$ is not necessarily equal to $f(X|T,P)$. As a result, (3) and (5) may not be maintained. Moreover, it is unclear whether $f(P|D=1,T=1)=f(P|D=1,T=0)$. Hence (4) may not be true either. Accordingly, (6) would not follow. One might condition on either P_b or P_a, instead of P_{all}. Nonetheless, if we use onlyP_b, only (5) is satisfied, while (3) and (4) are not necessarily true. Similarly, if we use only P_a, only (3) would hold. These results are summarized in the following lemma.

Lemma 3. If Assumption 2 is not satisfied, then conditioning on $P = P_{all}$ is not sufficient for (3), (4), and (5) to hold, whereas conditioning on $P = P_b$ (or $P = P_a$) implies only (5) (or (3)).

By contrast, if we condition on both P_b and P_a instead of the univariate PS, it is easy to show that $f(X|D=1,T,P)=f(X|T,P)$, where $P=(P_b,P_a)$. Hence we obtain the following lemma.

Lemma 4. Conditioning on both P_b and P_a is sufficient for both (3) and (5) to hold, even if Assumption 2 is not satisfied.

In this regard, using two-dimensional PS helps identification. Identification is not fully guaranteed, however, because without further information it is unclear whether (4) would hold.

Under compositional changes, we thus need to verify the conditions in (3), (4) and (5) which the identification result in Proposition 1 hinges on. Because these conditions include only observed variables, they are testable in principle, in contrast to Assumption 1 which is not directly testable. For this reason, I focus on the conditions in (3), (4) and (5), and the next section develops a nonparametric test to check whether data are consistent with these conditions.

DOI: 10.1002/jae

Nevertheless, one may wonder whether we could bypass the test and use a different approach. To that end, the following proposition suggests an alternative approach.[21]

Proposition 2. If Assumption 1 holds, then it follows that

$$\mathbf{Y} \perp (D, T) | \Pr(D = 0, T = 0|X), \Pr(D = 0, T = 1|X), \Pr(D = 1, T = 0|X)$$

Proof. Define $P_{d,t} \equiv \Pr(D = d, T = t|X)$. By the law of iterated expectation, it follows that $\Pr(D = 0, T = 0|P_{0,0}, P_{0,1}, P_{1,0}, \mathbf{Y}) = E[\Pr(D = 0, T = 0|X, P_{0,0}, P_{0,1}, P_{1,0}, \mathbf{Y})|P_{0,0}, P_{0,1}, P_{1,0}, \mathbf{Y}]$, which is equal to $E[\Pr(D = 0, T = 0|X, P_{0,0}, P_{0,1}, P_{1,0})|P_{0,0}, P_{0,1}, P_{1,0}, \mathbf{Y}]$ by Assumption 1. This conditional expectation can be rewritten as $\Pr(D = 0, T = 0|P_{0,0}, P_{0,1}, P_{1,0})$. Likewise, we can obtain that $\Pr(D = 0, T = 1|P_{0,0}, P_{0,1}, P_{1,0}, \mathbf{Y}) = \Pr(D = 0, T = 1|P_{0,0}, P_{0,1}, P_{1,0})$, and that $\Pr(D = 1, T = 0|P_{0,0}, P_{0,1}, P_{1,0}, \mathbf{Y}) = \Pr(D = 1, T = 0|P_{0,0}, P_{0,1}, P_{1,0})$. Therefore, $\mathbf{Y}$ is independent of (D, T), conditional on $P_{0,0}$, $P_{0,1}$ and $P_{1,0}$. □

Proposition 2 suggests that we can match on three propensity scores instead of two propensity scores. An advantage of this approach is that it does not require (3), (4) and (5), hence bypassing a test proposed in the next section. However, I do not use this approach in this paper, because the actual implementation of this approach requires three-dimensional nonparametric matching, which is computationally much more intensive and requires a very large dataset.

5.4. Test of Conditional Mean Independence

The conditions in (3), (4) and (5) are conditional independence restrictions. Because of high dimensionality in X, testing conditional independence is not practically feasible. Therefore, I consider conditional mean independence conditions given by

$$E(X|D = 1, T = 1, P) = E(X|D = 0, T = 1, P) \tag{7}$$

$$E(X|D = 1, T = 1, P) = E(X|D = 1, T = 0, P) \tag{8}$$

$$E(X|D = 1, T = 0, P) = E(X|D = 0, T = 0, P), \tag{9}$$

which is an important necessary condition for conditional independence in (3), (4) and (5). Consequently, I need to test the following null hypothesis:

$$H_0 \quad (7),\ (8) \text{ and } (9) \text{ are true for all values of } P \tag{10}$$

As mentioned in Section 5.3, P can be one-dimensional (i.e. P_{all}, P_b, or P_a), or two-dimensional, i.e. (P_b, P_a). I consider this null hypothesis for each case of P. For example, I consider H_0 when $P = P_{\text{all}}$, and separately consider H_0 when $P = (P_b, P_a)$.[22]

To develop a feasible test for (10), note first that $E(XDT|P) = E(XDT|D = 1, T = 1, P) \times \Pr(D = 1, T = 1|P)$. Similar results can be obtained for $(D = 1, T = 0)$; $(D = 0, T = 1)$; and $(D = 0, T = 0)$. Therefore, it follows that

$$E(X|D = 1, T = 1, P) = \frac{E[XDT|P]}{q_{1,a}(P)}, \qquad E(X|D = 1, T = 0, P) = \frac{E[XD(1 - T)|P]}{q_{1,b}(P)},$$

$$E(X|D = 0, T = 1, P) = \frac{E[X(1 - D)T|P]}{q_{0,a}(P)}, \qquad E(X|D = 0, T = 0, P) = \frac{E[X(1 - D)(1 - T)|P)}{q_{0,b}(P)}$$

[21] I appreciate a referee for suggesting this alternative approach.

[22] Given Lemma 4, if $P = (P_b, P_a)$, then we only need to check (4) or (8). Nevertheless, I consider all three restrictions in (7)–(efmean3) to examine more general cases for P.

DOI: 10.1002/jae

where $q_{1,a}(P) \equiv \Pr(D=1, T=1|P)$, $q_{1,b}(P) \equiv \Pr(D=1, T=0|P)$, $q_{0,a}(P) \equiv \Pr(D=0, T=1|P)$, and $q_{0,b}(P)$ $\Pr(D=0, T=0|P)$. Using these results, (7) can be rewritten as

$$E\left[\frac{XDT}{q_{1,a}(P)} - \frac{X(1-D)T}{q_{0,a}(P)}\middle|P\right] = 0$$

This conditional moment restriction implies the following unconditional moment:

$$E\left[\left(\frac{XDT}{q_{1,a}(P)} - \frac{X(1-D)T}{q_{0,a}(P)}\right)P\right] = 0,$$

which generates the sample moment

$$\frac{1}{N}\sum_{i=1}^{N} m_1(Z_i, q_i) = \frac{1}{N}\sum_{i=1}^{N}\left[\frac{X_iD_iT_i}{q_{1,a}(P_i)} - \frac{X_i(1-D_i)T_i}{q_{0,a}(P_i)}\right] P_i \tag{11}$$

where $Z_i = (X_i, D_i, T_i, P_i)$, $q_i = (q_{1,a}(P_i), q_{1,b}(P_i), q_{0,a}(P_i), q_{0,b}(P_i))$, and N denotes the number of observations. Similar moment restrictions for (8) and (9) can be obtained as

$$\frac{1}{N}\sum_{i=1}^{N} m_2(Z_i, q_i) = \frac{1}{N}\sum_{i=1}^{N}\left[\frac{X_iD_iT_i}{q_{1,a}(P_i)} - \frac{X_iD_i(1-T_i)}{q_{1,b}(P_i)}\right]P_i \tag{12}$$

$$\frac{1}{N}\sum_{i=1}^{N} m_3(Z_i, q_i) = \frac{1}{N}\sum_{i=1}^{N}\left[\frac{X_iD_i(1-T_i)}{q_{1,b}(P_i)} - \frac{X_i(1-D_i)(1-T_i)}{q_{0,b}(P_i)}\right]P_i \tag{13}$$

These sample moments, however, cannot be directly computed because $q_{1,a}(P_i), q_{1,b}(P_i), q_{0,a}(P_i)$ and $q_{0,b}(P_i)$ cannot be estimated using only one observation i. To overcome this problem, I use kernel estimators as follows:

$$\hat{q}_{1,a}(P_i) = \frac{\sum_{j=1}^{N} D_jT_jK_h(P_j - P_i)}{\sum_{j=1}^{N} K_h(P_j - P_i)}, \qquad \hat{q}_{1,b}(P_i) = \frac{\sum_{j=1}^{N} D_j(1-T_j)K_h(P_j - P_i)}{\sum_{j=1}^{N} K_h(P_j - P_i)},$$

$$\hat{q}_{0,a}(P_i) = \frac{\sum_{j=1}^{N} (1-D_j)T_jK_h(P_j - P_i)}{\sum_{j=1}^{N} K_h(P_j - P_i)}, \quad \hat{q}_{0,b}(P_i) = \frac{\sum_{j=1}^{N} (1-D_j)(1-T_j)K_h(P_j - P_i)}{\sum_{j=1}^{N} K_h(P_j - P_i)}$$

where $K_h(u) = h^{-1}K(u/h)$, h is a bandwidth, and $K(\cdot)$ is a kernel function.

Testing the moment restrictions requires the asymptotic distribution for the sample moments in (11), (12) and (13), where $q_{1,a}, q_{1,b}, q_{0,a}$ and $q_{0,b}$ are replaced by nonparametric estimators $q_{1,a}, q_{1,b}, q_{0,a}$ and $\hat{q}_{0,b}$. Although the presence of kernel estimators makes it difficult to derive the asymptotic distribution, I can take advantage of the similarity between these sample moments and semi-parametric M-estimators such as GMM estimators where kernel estimators are plugged into the sample moment restrictions (e.g. Newey 1994a, 1994b). The only difference is that my sample moments do not include unknown parameters, whereas thesample moments in semi-parametric M-estimators include unknown parameters which are usually estimated by minimizing the sample moment restrictions. Consequently, the asymptotic results for the sample moments in semi-parametric M-estimators should hold in my sample moments as well.

Therefore, under appropriate regularity conditions (see Newey, 1994a),

$$\sqrt{N}\hat{m}_N \xrightarrow{d} N(0, V) \tag{14}$$

where $\hat{m}_N = \frac{1}{N}\sum_{i=1}^{N} m(Z_i, \hat{q}_i)$, and $m(Z_i, \hat{q}_i) = (m_1(Z_i, \hat{q}_i)', m_2(Z_i, \hat{q}_i)', m_3(Z_i, \hat{q}_i)')'$. A Wald statistic for the null hypothesis (10) is then given by $N\hat{m}_N{}'\hat{V}^{-1}\hat{m}^N$, where $\hat{V}$ is a consistent estimator for the asymptotic variance V (see Appendix A for its formula). Note that under the null hypothesis $\sqrt{N}\hat{V}^{-\frac{1}{2}}\hat{m}_N \rightarrow^d N(0, I_J)$, where I_J is a $J \times J$ identity matrix, and J is the number of moments in $\hat{m}_N$. Hence it follows that $N\hat{m}_N{}'\hat{V}^{-1}\hat{m}_N \rightarrow^d \chi^2(J)$. Note that this test can be easily extended to the case of two-dimensional propensity scores, by using two-variate kernel estimators.

Using the CEX, I then compute test statistics $N\hat{m}_N{}'\hat{V}^{-1}\hat{m}_N$ conditional on univariate PS and two-variate PS. For P_{all}, P_b and P_a, the test statistics are respectively 182.68 (0.035), 184.53 (0.028), and 294.14 (0.000), where p-values are in parentheses. For both P_b and P_a, the test statistic is 117.68 (0.975). This test suggests that identifying restrictions are rejected if conditioned on the standard one-dimensional PS. In contrast, if I condition on two-dimensional PS, I fail to reject identifying restrictions in the CEX.[23]

5.5. DDM Estimator

The preceding test suggests that if we use the CEX data and exploit a natural experiment, at least two-dimensional propensity scores are required for identification under compositional changes. Because existing methods for PS matching are developed only for one-dimensional PS, I need to extend it to the case of two-dimensional PS. Both Abadie (2005) and Heckman *et al.* (1997, 1998) extend PS matching methods to DD models. While Abadie (2005) proposes clever weighting schemes using one-dimensional PS, these weighting schemes break down if $P \neq P_b \neq P_a$ (more specifically, Lemmas 3.1 and 3.2 in his paper do not hold). In contrast, the DD matching (DDM) estimator developed by Heckman *et al.* (1997, 1998) is based on the idea closely related to (6) in Proposition 1. As long as the conditions in (3), (4) and (5) are satisfied by using univariate or multivariate PS, the DDM method is still valid. For this reason, I extend the DDM estimator to the case of two-variate PS in what follows.

To do so, let us begin with the main parameter of interest in (1) and rewrite it as

$$\begin{aligned} M &= E\left(Y_{1,a,r} \middle| D=1, T=1\right) - E\left(Y_{1,b,\bar{r}} \middle| D=1, T=1\right) \\ &= \int \left[E\left(Y_{1,a,r} \middle| D=1, T=1, P_b, P_a\right) - E\left(Y_{1,b} \middle| D=1, T=0, P_b, P_a\right)\right. \\ &\quad \left. -E\left(Y_{0,a} \middle| D=0, T=1, P_b, P_a\right) + E\left(Y_{0,b} \middle| D=0, T=0, P_b, P_a\right)\right] \times dF(P_b, P_a | D=1, T=1) \end{aligned}$$

where the second equality follows from (6), which is likely to hold if we condition on both P_b and P_a. The equation above suggests the following estimator:

[23] As a robustness check, we can consider the conditional mean independence conditions for the second moments XX'. In principle, we can apply the same test to S, where S denotes the column vector of all diagonal elements and upper (or lower) diagonal elements in the matrix XX'. The actual implementation is computationally difficult, however, because given that there are 50 variables in X, the dimension of S is 1275×1, which means that the total number of moments is 3825. To compute the test statistics $N\hat{m}_N{}'\hat{V}^{-1}\hat{m}_N$, we need to invert $\hat{V}$. However, it is computationally very intensive to invert a 3825×3825 matrix. As a result, I instead perform the same test using only the diagonal elements of XX'. For P_{all}, P_b and P_a, the test statistics are respectively 182.76 (0.035), 185.01 (0.027), and 295.47 (0.000), while for both P_b and P_a the test statistic is 121.25 (0.959). Hence, we obtain similar results.

DOI: 10.1002/jae

$$\hat{M} = \sum_{i \in G_{1,a}} \left[Y_i - \hat{E}\left(Y_j \middle| D_j = 1, T_j = 0, P_{b,i}, P_{a,i}\right) - \hat{E}\left(Y_j \middle| D_j = 0, T_j = 1, P_{b,i}, P_{a,i}\right) + \hat{E}\left(Y_j \middle| D_j = 0, T_j = 0, P_{b,i}, P_{a,i}\right) \right] \times w_i \quad (15)$$

where $G_{1,a}$ denotes the post-Napster Internet user group, and $\hat{E}\left(Y_j \middle| D_j = 1, T_j = 0, P_{b,i}, P_{a,i}\right)$, $\hat{E}\left(Y_j \middle| D_j = 0, T_j = 1, P_{b,i}, P_{a,i}\right)$ and $\hat{E}\left(Y_j \middle| D_j = 0, T_j = 0, P_{b,i}, P_{a,i}\right)$ are the conditional expectation estimators for each group conditional on $P_{b,i}$ and $P_{a,i}$ of observation i in $G_{1,a}$, and the weight for i is given by $w_i = (\text{CEX weight})_i / \left(\sum_{k \in G_{1,a}} (\text{CEX weight})_k\right)$.

The proposed estimator in (15) is a modified version of the DDM estimator in Heckman *et al.* (1997, 1998), for which I modify the standard DDM estimator by nonparametrically matching each observation in the post-Napster Internet user group with observations in the pre-Napster Internet user group and the Internet non-user group based on $P_{b,i}$ and $P_{a,i}$ of each i in $G_{1,a}$. I use local linear matching, following Heckman *et al.* (1997, 1998), who use univariate local linear matching instead of kernel matching because local linear estimators perform better particularly at boundary points. Because I match based on two-dimensional propensity scores, I further use multivariate versions oflocal linear regression estimators developed by Ruppert and Wand (1994). Specifically, I estimate the conditional expectation in (15) by using local linear regressions as follows:

$$\hat{E}\left(Y_j \middle| D_j = d, T_j = t; P_{b,i}, P_{a,i}\right) = e_1'(X_{Pi}' W_{P_i} X_{P_i})^{-1} X_{Pi}' W_{P_i} Y,\ e_1 = (1 \quad 0 \quad 0)',\ Y = \left(Y_1, \cdots, Y_{N_{d,t}}\right)',\ X_{P_i}$$

$$= \begin{pmatrix} 1 & \cdots & 1 \\ \left(P_{b,1} - P_{b,i}\right) & \cdots & \left(P_{b,N_{d,t}} - P_{b,i}\right) \\ \left(P_{a,1} - P_{a,i}\right) & \cdots & \left(P_{a,N_{d,t}} - P_{a,i}\right) ndarray \end{pmatrix}',$$

$$WP_i = diag\left(W_1, \cdots, W_{N_{d,t}}\right), W_j = K_h\left(P_{b,j} - P_{b,i}\right) \times K_h\left(P_{a,j} - P_{a,i}\right)$$

where $N_{d,t}$ is the number of observations for the group $G_{d,t}$ characterized by d and t, $K_h(u) = h^{-1} K(u/h)$, h is a fixed bandwidth, and $K(\cdot)$ is a biweight kernel function as

$$K(s) = 15/16\left(s^2 - 1\right)^2 \times 1\{|s| < 1\}$$

For the fixed bandwidth, I use 0.07 throughout the estimations. I find that results are comparable for other fixed bandwidths within ±0.02 of 0.07. To estimate the standard errors for the DDM estimator in (15), I use the bootstrap method.[24]

6. RESULTS

6.1. The DDM Estimates of the Effect of the Presence of Napster

To estimate the main parameter of interest in (1), I first estimate two-variate PS and then apply the DDM methods developed in the previous section. I follow the standard PS-matching literature (see,

[24] Note that Abadie and Imbens (2008) prove that the bootstrap is not valid for the nearest-neighbor matching estimator. They further conjecture that asymptotic normality might not be sufficient to validate the use of the bootstrap if the estimators are not asymptotically linear. In this regard, the bootstrap is even more likely to be valid in my case because the DDM estimator uses local linear matching, and Heckman *et al*. (1998) prove that this class of estimator is not only asymptotically normal but also asymptotically linear.

DOI: 10.1002/jae

Table II. Estimates for the propensity scores

Variable	Before	After	Variable	Before	After
constant	−0.485(0.157)	−0.864(0.116)	hw.no.child	0.037(0.050)	−0.016(0.036)
age	−0.014(0.004)	0.025(0.003)	hw.child.bf.school	0.011(0.055)	0.061(0.042)
$(\text{age})^2$	0.000(0.000)	0.000(0.000)	hw.child.in.school	0.162(0.049)	0.127(0.036)
white	−0.104(0.094)	0.186(0.067)	hw.child.af.school	0.118(0.045)	0.096(0.032)
black	−0.388(0.099)	−0.052(0.070)	sp.child.bf.school	−0.323(0.077)	−0.078(0.050)
male	0.086(0.020)	0.095(0.014)	sp.child.in.school	0.029(0.060)	0.045(0.044)
hs.grad	0.379(0.041)	0.317(0.025)	retired	0.028(0.049)	0.045(0.033)
less.college	0.612(0.040)	0.545(0.025)	head.working	0.192(0.061)	0.114(0.045)
college.grad	0.705(0.042)	0.616(0.026)	spouse.working	−0.097(0.046)	−0.058(0.034)
tv	0.009(0.008)	0.015(0.006)	work.week	−0.005(0.001)	−0.002(0.001)
computer	1.113(0.021)	1.160(0.016)	work.hour	−0.001(0.001)	−0.002(0.001)
sound.system	0.022(0.023)	−0.028(0.017)	spouse.work.week	0.001(0.001)	0.000(0.001)
vcr	−0.344(0.028)	−0.376(0.021)	spouse.work.hour	0.000(0.001)	0.002(0.001)
vehicle	0.021(0.006)	0.031(0.005)	owner	−0.958(0.043)	−1.228(0.050)
manager	0.183(0.032)	0.129(0.025)	renter	−1.225(0.041)	−1.377(0.049)
teacher	−0.021(0.047)	−0.068(0.036)	income.bf.tax	0.047(0.004)	0.049(0.003)
professional	0.186(0.031)	0.108(0.024)	$(\text{income.bf.tax})^2$	−0.001(0.000)	−0.001(0.000)
admin	0.108(0.036)	0.090(0.027)	northeast	0.067(0.026)	0.079(0.020)
technician	0.210(0.042)	0.112(0.033)	midwest	−0.020(0.023)	0.026(0.017)
sales	0.044(0.036)	0.108(0.027)	west	0.065(0.022)	0.041(0.017)
services	0.048(0.037)	0.003(0.026)	urban	0.243(0.044)	0.097(0.034)
family.size	−0.016(0.019)	−0.044(0.012)	msa	−0.094(0.055)	−0.088(0.040)
pers.age.lt.11	−0.086(0.024)	−0.016(0.016)	pop>4million	0.129(0.048)	0.103(0.033)
pers.age.12-17	−0.034(0.022)	0.009(0.015)	pop>1million	0.132(0.048)	0.098(0.033)
pers.age.gt.64	−0.087(0.029)	0.000(0.019)	pop>330k	0.183(0.050)	0.093(0.035)
single	−0.041(0.043)	−0.164(0.029)	pop>125k	−0.001(0.050)	0.011(0.035)
			Observations	46,124	61,526

Note: Probit models are estimated using the CEX. The dependent variable is a dummy for Internet access. Beginning from 1999, the CEX increased its sample size by about 50%.

for example, Heckman *et al.*, 1997; Dehejia and Wahba, 1999; Lechner, 1999; Behrman *et al.*, 2004; Smith and Todd, 2005), where propensity scores are parametrically estimated by either probit or logit. Using the CEX data, I thus estimate probit models of Internet access separately for the pre-Napster and post-Napster periods. The estimates are presented in Table II. Note that most coefficient estimates for the pre-Napster period are considerably different from those for the post-Napster period, which is consistent with compositional changes between Internet users and Internet non-users over time. Using these estimates, I impute P_b and P_a for all observations. Note that for each bootstrap sample I re-estimate the probit and re-impute P_b and P_a in order to account for the first stage estimation errors.

However, one may be still concerned that the propensity score might be misspecified because it is parametrically estimated. To address this concern, I use a specification test developed by Shaikh *et al.* (2009). This test is based on a restriction between the density of the PS among the treatment group and the density among the control group, which will not be satisfied if the PS is misspecified. I perform this test and find that the test statistics for P_b and P_a are respectively 0.382 (0.702) and 0.427 (0.670), where p-values are in parentheses. Therefore, I fail to reject the null hypothesis that the PS is correctly specified.[25]

In addition to the specification test, I also examine whether there are overlaps between the estimated propensity scores for the treatment group and the control group. Figures 4 and 5 in Appendix B present the histograms of the estimated P_b and P_a for four groups: $G_{1,a}, G_{1,b}, G_{0,a}$ and $G_{0,b}$, respectively representing Internet users after and before, and non-users after and before. The figures show that

[25] I appreciate Marianne Simonsen for providing me with the Gauss code. The test statistic is computed by $\hat{V}_n/\sqrt{\hat{\Sigma}_n}$ (see equations (9) and (10) in Shaikh *et al.*, 2009), and its asymptotic distribution follows standard normal. The bandwidth of 0.001 is used to compute the test statistics. For other bandwidths, I also fail to reject the null hypothesis.

DOI: 10.1002/jae

Table III. DDM estimates using two-variate propensity scores

	Local linear matching	Kernel matching
	(1)	(2)
M	– 1.446 (0.624)	– 1.746 (0.579)

Note: Bootstrapped standard errors in parentheses. Both matching methods use a fixed bandwidth of 0.07 and biweight kernel.

for all the bins the estimated PS for $G_{1,a}$ is reasonably overlapped with those for $G_{1,b}$, $G_{0,a}$ and $G_{0,b}$. Hence the overlap condition for the PS matching is unlikely to be violated in my estimation.

Based on the estimated P_b and P_a, I then match each post-Napster Internet user i with Internet non-users and pre-Napster Internet users to construct the counterfactual $E\left(Y_{1,a,\bar{r}}\middle|D = 1, T = 1\right)$. The DDM estimate for the main parameter of interest M is reported in column 1 of Table III. I use local linear matching as described in the previous section. The estimate is precisely estimated and indicates that the presence of Napster had reduced the quarterly music expenditures for the average Internet users during the Napster period by \$1.45, which is 7.6% of their music expenditures.

Because the CEX data are nationally representative random samples and provide weights for each observation in the CEX, I can further estimate changes in total record sales among the US population. To this end, consider the following back-of-the-envelope calculation of the impact of Napster on total record sales during the Napster period. The average percentage of Internet users in the CEX is 26% during the first year of the Napster period and 31% during the second year. There were approximately 100 million households in the USA during this period. Noting that the estimate is quarterly expenditure in terms of 1998 dollars, total record sales decline in the period attributable to Napster is then given by – \$329.69 million = 100 million × (0.26 × 4 + 0.31 × 4) × (– \$1.446). According to the CEX, the decrease in total record sales from the pre-Napster period to the post-Napster period amounts to \$832.24 million. This suggests that 39.6.

6.2. The Estimates from Alternative Approaches

The main estimate in the previous section is estimated by the DDM method using local linear matching based on two-variate PS. To examine the extent to which this method addresses the negative bias from compositional changes, I further considerother approaches to estimate *M*. I begin with the conventional DD regressions, which attempt to estimate *M* by estimating the following equation:

$$Y_i = \alpha + \theta D_i T_i + \gamma D_i + \delta T_i + X_i \beta + v_i \tag{16}$$

where θ is a fixed parameter for the treatment effect and v_i is an error term. For θ in (16) to identify M, however, the DD regression requires strong assumptions that are unlikely to hold.[26] Table IV presents the coefficient estimates for θ. Without any control, the conventional DD estimate is – 4.69. Adding additional control reduces the negative bias to some extent, and the estimate becomes – 3.60. However, the implied negative bias is still large. The preceding back-of-the-envelope calculation using the estimate of – 3.6 yields – \$820.8 million and therefore one would incorrectly conclude that almost 100% of sales decline was due to the presence of Napster.

Proper weighting schemes may address imbalances in the distribution of the covariates between treated and untreated, thereby reducing the negative bias. For this reason, I next consider weighted least

[26] See Hong (2007a) for further discussion on these assumptions. Abadie (2005) also provides related discussion on similar assumptions.

DOI: 10.1002/jae

Table IV. DD regression estimates

	Unweighted		P_{all}-Weighted		P_b-Weighted		P_a-Weighted	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
θ	−4.69	−3.60	−3.27	−2.66	−2.86	−2.50	−3.42	−2.69
	(0.53)	(0.52)	(0.54)	(0.54)	(0.56)	(0.55)	(0.54)	(0.53)
Control	No	Yes	No	Yes	No	Yes	No	Yes

Note: Standard errors in parentheses. Regressions with control include various covariates such as age, education, income, appliance ownership, occupation, family composition, and region. All regressions are estimated by weighted least squaresusing the CEX weights. PS-weighted regressions use the product of the CEX weights and the univariate PS as weights.

squares of the DD regressions using the univariatePS.[27] Columns 3–8 of Table IV report three kinds of PS-weighted regressions based on the propensity scores for only the pre-Napster period, for only the post-Napster period, or for all periods: P_b, P_a, or P_{all}. The table shows that the magnitude of the negative effect is further reduced by using PS-weighted regressions. The negative bias still remains, however.

Because the conventional DD regressions are based on parametric linearity assumptions, I consider nonparametric approaches for the PS matching. Hence I use the DDM method, but I start with matching based only on the univariate PS. Specifically, I estimatethe conditional expectation in (15) by using $\hat{E}\left(Y_j \middle| D_j = d, T_j = t; P_i\right) = \sum_{j \in G_{d,t}} W(i,j) Y_j$, where $G_{d,t}$ denotes the group characterized by d and t, and P_i is the univariate PS, which can be $P_{b,i}, P_{a,i}$, or $P_{all,i}$. The weight for observation j in $G_{d,t}$ matched to observation i in the post-Napster Internet user group is denoted by $W(i,j)$. For kernel matching, it is given by

$$W(i,j) = \frac{K_h\left(P_j - P_i\right)}{\sum_{k \in G_{d,t}} K_h(P_k - P_i)}$$

and $W(i,j)$ for local linear matching is given by

$$W(i,j) = \frac{S_{ij} \sum_{k \in G_{d,t}} (P_k - P_i)^2 S_{ik} - \left(P_j - P_i\right) S_{ij} \sum_{k \in G_{d,t}} (P_k - P_i) S_{ik}}{\sum_{l \in G_{d,t}} S_{il} \sum_{k \in G_{d,t}} (P_k - P_i)^2 S_{ik} - \left(\sum_{k \in G_{d,t}} (P_k - P_i) S_{ik}\right)^2}$$

where $S_{ij} = K_h(P_j - P_i)$. The estimates are presented in Table V. The table first shows that local linear matching yields slightly smaller magnitudes than kernel matching, suggesting that local linear matching performs moderately better than kernel matching. All the estimates in Table V have smaller magnitudes than those from the DD regressions. However, compared to PS-weighted regressions with control in Table IV, the decrease in the negative bias is not substantial. A comparison of Tables III–V shows that in terms of magnitudes the DDM estimates using two-variate PS are significantly smaller than the estimates using alternative approaches, suggesting the importance of matching based on two propensity scores.

6.3. The DDM Estimates for Age and Family Groups

Using the DDM method based on two-variate PS, I further investigate which demographic groups are responsible for the sales decline attributable to the presence of Napster. To this end, I consider different age and family groups and specifically examine four mutually exclusive groups: (i) households with children

[27] Finkelstein (2004) used similar PS-weighted regressions.

DOI: 10.1002/jae

Table V. DDM estimates using univariate propensity score

	Matching based on P_{all}	Matching based on P_b	Matching based on P_a
	(1)	(2)	(3)
A. Local linear matching			
M	−2.318 (0.742)	−2.060 (0.682)	−2.485 (0.602)
B. Kernel matching			
M	−2.341 (0.584)	−2.378 (0.703)	−2.526 (0.583)

Note: Bootstrapped standard errors in parentheses. Both matching methods use a fixed bandwidth of 0.07 and biweight kernel.

aged 6–17,[28] and three groups excluding (i), which include (ii) households with heads aged 15–34, (iii) households with heads aged 35–49 and (iv) households with heads aged over 50. Note that the number of observations in each group is similar, except for those aged over 50, and total record sales for each group are comparable although the 15–34 age group spent slightly more than other groups in the CEX.

Table VI presents the DDM estimates of the main parameter of interest for these four groups. The DDM estimate for those with children aged 6–17 is – $3.26 and is precisely estimated. The estimate for those aged 15–34 is – $2.99, but its standarderror is 3.01.[29] For those aged 35–49 and those aged over 50, the estimates are small and statistically insignificant.

The DDM estimate is estimated precisely only for those with children aged 6–17. Nevertheless, the estimated magnitudes can be used for similar back-of-the-envelope calculations in Section 6.1. The CEX reports that the percentages of four groups in the post-Napster Internet user group are 21% for those aged 15–34, 23% for those aged 35–49, 31% for those aged over 50, and 24% for those with a child aged 6–17. Using these percentages and the DDM estimates, I find that the DDM estimate for households with children aged 6–17 is translated into – $196 million, which accounts for about 20% of total record sales decline during the Napster period (the 95% confidence interval is [– $286 million, – $106 million]).

For households aged 35–49 and those over 50, the DDM estimates imply – $26 million and – $31 million, respectively. The magnitudes of these estimates are small and their standard errors suggest that they are statistically indistinguishable from zero. As for those aged 15–34, the DDM estimate is translated into – $159 million, which approximately explains another 20% of total sales decline during the Napster period. However, we cannot be confident about this result because of the large standard error for this demographic group.

[28] I consider a separate group for households with children aged 6–17 to examine changes in music expenditures for those young consumers who are presumed to be heavy music buyers. Note that in the CEX young consumers' expenditures are included in their parents' expenditure unless they are financially independent. One may then worry about a potential measurement error, in that parents might not know about their children's music expenditures. Nonetheless, such a measurement error is unlikely correlated with Internet access or file sharing, and thus unlikely to affect the DD estimates. However, I acknowledge that if the measurement error for this demographic group were significant *relative to* other demographic groups, the magnitude from the back-of-the-envelope calculation would be underestimated for this group compared to other demographic groups, in which case the percentage changes could be underestimated as well. Given information available, however, this problem cannot be addressed in this paper.

[29] Note that college students living in a dormitory are also included in those aged 15–34. In this paper, I treat them as Internet users because most college dormitories were likely to have the Internet connection in the late 1990s. However, I was not able to obtain any official evidence to substantiate this, other than anecdotal evidence. Another issue related to college students living in dormitories is that their counterparts in the control group might be difficult tofind. For these reasons, I drop them from those aged 15–34, and re-estimate the DDM estimate. I find that the estimate is – $2.17, but its standard error is equal to 2.32. Hence the relatively high standard error for those aged 15–34 is not necessarily caused only by including 'college students living in a dormitory' in the treatment group.

DOI: 10.1002/jae

Table VI. DDM estimates for age and family groups

	Age 15–34 (1)	Age 35–49 (2)	Age 50+ (3)	HHs w/children aged 6–17 (4)
M	−2.992 (3.012)	−0.453 (0.988)	−0.408 (0.765)	−3.256 (0.748)

Note: Bootstrapped standard errors in parentheses. The DDM estimates are estimated from local linear matching based on two-variate PS. The propensity scores for both pre- and post-Napster periods are estimated separately for each group, excluding age and family composition variables from the probit estimation in Table II. A fixed bandwidth of 0.07 and biweight kernel are used. The group aged 15–34 accounts for 21% of total samples; those aged 35–49 include 20%; those with children aged 6–17 contain 18%; those aged over 50 include 41% of samples. In particular, the percentage of each group in Internet users during the post-Napster period is comparable.

6.4. The Effect of Music Downloading on Music Expenditure

The DDM estimates measure the effect of Napster, but they do not necessarily reflect solely the pure effect of file sharing. As a result, one may wish to further isolate the effect of actual music downloading. To this end, I consider an additional approach to decompose the effect of Napster into the effect of music downloading and the effect of other new online activities that might have affected music expenditures during the Napster period. Because the DDM estimates suggest that the effect of Napster is small and statistically insignificant for those aged 35–49 and those over 50, I focus on households with children aged 6–17 and households with heads aged 15–34 in this section.

Since the CEX does not contain any information on music downloading, I use a complementary dataset with detailed demographic variables and information on actual downloading activity. Specifically, I use annual household-level surveys on Internet usage collected by the Center for Communication Policy at the University of California, Los Angeles (henceforth, UCLA Internet Survey, or UCLAIS) for 2000–2002. The UCLAIS, however, does not include music expenditures, and thus it must be combined with the CEX. I first maintain the DDM approach and estimate nonparametric bounds using the method proposed by Cross and Manski (2002), but I find that these bounds are not informative.[30] Therefore, I consider a two-sample instrumental variable (2SIV) approach (Angrist and Krueger, 1992; Arellano and Meghir, 1992) which exploits linearity in the DD regression to allow for data combination and provides more informative results. Though there are other approaches to combine different micro-level data (see, for example, Moffitt and Ridder, 2007; Chen *et al.*, 2008; Hong and Wolak, 2008), I use the 2SIV method mainly because of its tractability.

I begin with the DD regression in (16) and decompose θ into θ_0 and θ_1 as

$$Y_i = \alpha + \theta_0 D_i T_i + \theta_1 D_i T_i \mathrm{DM}_i + \gamma D_i + \delta T_i + X_i \beta + \varepsilon_i \tag{17}$$

where ε_i is an error term, and DM_i is an indicator dummy equal to 1 if observation i downloaded music. To implement the 2SIV, I first estimate probit models of music downloading, using the UCLAIS. Based on the estimates in Table VII, I next compute the predicted value of $\Pr(\mathrm{DM}_i = 1|Z_i)$ for each observation in the CEX, where Z_i is a vector of common variables in both datasets. Finally, I replace DM_i in (17) with the predicted $\Pr(\mathrm{DM}_i = 1|Z_i)$ and estimate θ_0 and θ_1 by running standard regressions. To account for the first-stage estimation errors, I resample both the UCLAIS and the CEX, and use bootstrap procedures to estimatestandard errors.[31]

[30] For more details on the UCLAIS and the bound results, see Hong (2007a).
[31] For more details on my application of the 2SIV, see Hong (2007a).

DOI: 10.1002/jae

Table VII. Probit estimates for music downloading

	Age 15–34	HHs w/children aged 6–17
	(1)	(2)
constant	1.831 (1.444)	1.001 (0.749)
age	−0.208 (0.119)	−0.130 (0.034)
$(\text{age})^2$	0.003 (0.002)	0.001 (0.000)
male	0.328 (0.098)	0.027 (0.130)
hs.grad	−0.010 (0.195)	0.208 (0.249)
less.college	0.084 (0.195)	0.122 (0.253)
college.grad	−0.022 (0.221)	0.098 (0.263)
college.student	0.169 (0.150)	
family.size	−0.004 (0.046)	−0.040 (0.104)
single	−0.074 (0.156)	
pers.age.lt.11		0.061 (0.124)
pers.age.12-17		−0.061 (0.155)
employed	−0.153 (0.125)	0.169 (0.159)
computer	0.155 (0.051)	0.052 (0.066)
internet	0.895 (0.138)	0.994 (0.189)
high.internet	0.489 (0.138)	0.348 (0.175)
income	−0.030 (0.037)	−0.093 (0.048)
$(\text{income})^2$	0.000 (0.002)	0.003 (0.003)
observations	1312	912

Note: Standard errors are in parentheses. The dependent variable is an indicator dummy for music downloading. Probit models are estimated separately for each demographic group in the UCLAIS.

Table VIII. 2SIV estimates for age and family groups

	Age 15–34	HHs w/children Aged 6–17
	(1)	(2)
A. DD estimates		
θ	− 3.432 (1.284)	− 3.258 (1.203)
B. 2SIV estimates		
θ_0	− 2.427 (0.949)	− 0.120 (1.092)
θ_1	− 2.719 (2.079)	− 22.510 (6.889)
Mean of imputed downloading probability	0.346	0.140

Note: Standard errors in parentheses. The dependent variable is music expenditure in 1998 dollars. All regressions are estimated by weighted least squares using the CEX weights. Panel A reports the coefficient estimates for θ in the DD regression (16), which includes controls such as age, education, income, appliance, occupation, family composition, and region. Panel B reports the coefficient estimates for θ_0 and θ_1 in the regression (17), which includes various covariates. Bootstrap is used to estimated standard errors.

Table VIII presents the results. Panel A reports the DD estimates for households aged 15–34 and those with children aged 6–17. Panel B reports the 2SIV estimates for these demographic groups. For households with children aged 6–17, the 2SIV results indicate that the effect of actual downloading measured by θ_1 is considerably negative and statistically significant, while the effect of other new online activities measured by θ_0 is statistically indistinguishable from zero, which implies that the DD estimate is unlikely to be confounded by other new online activities. Note also that the DD estimate for this group is decomposed as $-3.258 \approx -0.120 + (-22.510) \times 0.140$. As a result, the back-of-the-envelope calculation using the 2SIV results is almost identical to that in Section 6.3. These results suggest that the DDM estimate for this demographic group is more likely to represent the effect of file sharing. By contrast, the 2SIV results for those aged 15–34 show that the effect of actual downloading is fairly small and statistically insignificant, whereas the effect of other new online activities is statistically significant. These results therefore suggest that we can more confidently rule out the significant negative effect of file sharing on recorded music expenditure for households aged 15–34.[32]

7. CONCLUSION

To what extent is file sharing culpable for the recent slump in record sales, and which demographic group is primarily responsible for the sales decline due to file sharing? I study changes in household-level recorded music expenditure between the periods before and after the introduction of Napster, accounting for the likely relationship between music expenditure and the propensity to adopt the Internet, as well as potential confounding factors. I find evidence suggesting that file sharing can account for about 20% of the sales decline in recorded music during the Napster period, and that this negative effect is concentrated in households with children aged 6–17.

These findings contribute in part to the large literature on copyright protection in a digital era (see, for example, Posner, 2005; Varian, 2005). Digital technologies have dramatically reduced the cost of copying, thereby increasing consumers' access to copyrighted materials. Under this new environment, however, conventional copyright protection might fail to secure revenues of copyright holders, hence reducing financial incentives to create new works. In accessing the social efficiency of more restrictive copyright protection in a digital era, one necessary (but not sufficient) piece of information is the extent to which digital technologies have reduced revenues of copyright holders. My findings provide such information in the case of the recording industry.

The approaches in this paper are not limited to the Napster case and the analysis of the recording industry. In the absence of ideal experimental data, one may have to rely on natural experiments, and it is plausible to expect that some natural experiments entail compositional changes between the treatment group and the control group. In this paper, I examine identifying assumptions for the DD estimator under compositional changes and propose a test for identifying restrictions. To further address a negative bias due to compositional changes, I extend nonparametric DD matching estimators to the case of two-variate propensity scores. These approaches can be also applied to other natural experiments to address potential compositional changes.[33]

ACKNOWLEDGEMENTS

I thank Frank Wolak, Liran Einav and Peter Reiss for their invaluable insights and encouragement. I am grateful to the editor Edward Vytlacil and four anonymous referees for their valuable comments and suggestions, which significantly improved the manuscript. I also thank Tim Bresnahan, Han Hong, Roger Koenker, Mark Jacobsen, Alan Sorensen and Azeem Shaikh for their helpful comments and suggestions. All remaining errors are my responsibility.

---

[32] Recall that for those aged 15–34 the DDM estimate is -2.99 (3.01). However, the estimate becomes -2.17 (2.32) if I drop college students (see footnote 29). This suggests that the effect of Napster might be much larger for college students. However, the fraction of college students is only 5% among those aged 15–34 in the CEX. Therefore, even if the effect of file sharing is significantly negative for college students (see Rob and Waldfogel, 2006), it is only partially reflected in the estimates for those aged 15–34. However, I do not separately consider college students in this paper, partly because the small sample size leads to a very imprecise estimate, but also because it is difficult to find a reasonable control group counterpart to college students living in a dormitory (see footnote 29).

[33] One example of a DD approach under compositional change is a potential study investigating the effect of policy changes in unemployment insurance (UI) on the duration of unemployment (or other outcomes such as changes in expenditures). For example, we could have a research design in which the treatment group consists of unemployed workers in states with policy changes in UI (e.g. extended benefits). To study the effect on the duration of unemployment, the control group willinclude unemployed workers in states without policy change (or to study the effect on changes in expenditures, the control group can also include employed workers in the same states). However, the composition of unemployed workers in each state is unlikely to remain the same over time, especially during the recession when more diverse workers become unemployed. One potential example of policy changes is the American Recovery and Reinvestment Act of 2009, which has led some states to amend their laws to temporarily make extended benefits available for unemployed workers (refer to the US Department of Labor web site on this Act).

DOI: 10.1002/jae

REFERENCES

Abadie A. 2005. Semiparametric difference-in-differences estimators. *The Review of Economic Studies* **72**: 1–19.

Abadie A, Imbens GW. 2008. On the failure of the bootstrap for matching estimators. *Econometrica* **76**: 1537–1557.

Angrist JD, Krueger AB. 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 M, Meghir C. 1992. Female labour supply and on-the-job search: an empirical model estimated using complementary data sets. *The Review of Economic Studies* **59**: 537–557.

Athey S, Imbens GW. 2006. Identification and inference in nonlinear difference-in-differences models. *Econometrica* **74**: 431–497.

Battistin E. 2003. *Errors in survey reports of consumption expenditures*. IFS working paper 03/07. Institute for Fiscal Studies, London.

Behrman J, Cheng Y, Todd P. 2004. Evaluating preschool programs when length of exposure to the program varies: a nonparametric approach. *The Review of Economics and Statistics* **86**: 108–132.

Bertrand M, Duflo E, Mullainathan S. 2004. How much should we trust differences-in-differences estimates? *Quarterly Journal of Economics* **119**: 249–275.

Chen X, Hong H, Tarozzi A. 2008. Semiparametric efficiency in GMM models with auxiliary data. *The Annals of Statistics* **36**: 808–843.

Cross PJ, Manski CF. 2002. Regressions, short and long. *Econometrica* **70**: 357–368.

Dehejia R, Wahba S. 1999. Causal effects in non-experimental studies: re-evaluating the evaluation of training programs. *Journal of the American Statistical Association* **94**: 1053–1062.

Finkelstein A. 2004. Static and dynamic effects of health policy: evidence from the vaccine industry. *Quarterly Journal of Economics* **119**: 527–564.

Fisher W. 2004. *Promises to Keep: Technology, Law, and the Future of Entertainment*. Stanford University Press: Palo Alto, CA.

Gentzkow M. 2007. Valuing new goods in a model with complementarity: online newspapers. *The American Economic Review* **97**: 713–744.

Hahn J. 1998. On the role of the propensity score in efficient semiparametric estimation of average treatment effects. *Econometrica* **66**: 315–331.

Heckman JJ, Robb R. 1985. Alternative methods for evaluating the impact of interventions. In *Longitudinal Analysis of Labor Market Data*, Heckman JJ, Singer B (eds). Cambridge University Press: Cambridge, UK; 156–245.

Heckman JJ, Smith JA. 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**: 313–348.

Heckman JJ, Ichimura H, Todd PE. 1997. Matching as an econometric evaluation estimator: evidence from evaluating a job training programme. *The Review of Economic Studies* **64**: 605–654.

Heckman JJ, Ichimura H, Todd PE. 1998. Matching as an econometric evaluation estimator. *The Review of Economic Studies* **65**: 261–294.

Heckman JJ, Lalonde RJ, Smith JA. 1999. The economics and econometrics of active labor market programs. In *Handbook of Labor Economics, Vol. 3A*, Ashenfelter OC, Card D (eds). Elsevier: Amsterdam; 1865–2097.

Hirano K, Imbens GW, Ridder G. 2003. Efficient estimation of average treatment effects using the estimated propensity score. *Econometrica* **71**: 1161–1189.

Hong S-H. 2007a. *Measuring the effect of digital technology on the sales of copyrighted goods: evidence from Napster*. Available: http://ssrn.com/abstract=1205862 [22 August 2011].

Hong S-H. 2007b. The diffusion of the Internet and changes in the household-level demand for entertainment. *Information Economics and Policy* **19**: 304–318.

Hong S-H, Wolak FA. 2008. Relative prices and electronic substitution: changes in household-level demand for postal delivery services from 1986 to 2004. *Journal of Econometrics* **145**: 226–242.

Imbens GW, Wooldridge JM. 2009. Recent developments in the econometrics of program evaluation. *Journal of Economic Literature* **47**: 5–86.

Lechner M. 1999. Earnings and employment effects of continuous off-the-job training in East Germany after Unification. *Journal of Business and Economic Statistics* **17**: 74–90.

Liebowitz SJ. 2006a. File-sharing: creative destruction or just plain destruction? *Journal of Law and Economics* **49**: 1–28.

Liebowitz SJ. 2006b. Economists examine file-sharing and music sales. In *Industrial Organization and the Digital Economy*, Illing G, Peitz M (eds). MIT Press: Cambridge, MA; 145–174.

Moffitt R, Ridder G. 2007. The econometrics of data combination. In *Handbook of Econometrics*, Vol. 6B, Heckman JJ, Leamer E (eds). Elsevier: Amsterdam; 5469–5448.
Newey WK. 1994a. The asymptotic variance of semiparametric estimators. *Econometrica* **62**: 1349–1382.
Newey WK. 1994b. Kernel estimation of partial means and a general variance estimator. *Econometric Theory* **10**: 233–253.
Oberholzer-Gee F, Strumpf K. 2007. The effect of file sharing on record sales: an empirical analysis. *Journal of Political Economy* **115**: 1–42.
Oberholzer-Gee F, Strumpf K. 2010. File sharing and copyright. In *Innovation Policy and the Economy, NBER Book Series*, Vol. **10**, Lerner J, Stern S (eds). University of Chicago Press: Chicago, IL; 19–55.
Posner RA. 2005. Intellectual property: the law and economics approach. *Journal of Economic Perspective* **19**: 57–73.
Rob R, Waldfogel J. 2006. Piracy on the high C's: music downloading, sales displacement, and social welfare in a sample of college students. *Journal of Law and Economics* **49**: 29–62.
Romer P. 2002. When should we use intellectual property rights? *The American Economic Review* **92**(2): 213–216.
Rosenbaum P, Rubin D. 1983. The central role of the propensity score in observational studies for causal effects. *Biometrika* **70**: 41–55.
Rubin D. 1978. Bayesian inference for causal effects: the role of randomization. *The Annals of Statistics* **6**(1): 34–58.
Ruppert D, Wand MP. 1994. Multivariate locally weighted least squares regression. *The Annals of Statistics* **22**: 1346–1370.
Shaikh A, Simonsen M, Vytlacil E, Yildiz N. 2009. A specification test for the propensity score using its distribution conditional on participation. *Journal of Econometrics* **151**: 33–46.
Sinai T, Waldfogel J. 2004. Geography and the Internet: is the Internet a substitute or a complement for cities? *Journal of Urban Economics* **56**: 1–24.
Smith J, Todd P. 2005. Does matching overcome Lalonde's critique of nonexperimental estimators? *Journal of Econometrics* **125**: 305–353.
Varian HR. 2005. Copying and copyright. *Journal of Economic Perspective* **19**: 121–138.
Zentner A. 2006. Measuring the effect of file sharing on music purchases. *Journal of Law and Economics* **49**: 63–90.

APPENDIX: A CONSISTENT ESTIMATOR FOR THE ASYMPTOTIC VARIANCE V

Following Newey (1994b), I obtain a consistent estimator for V in (14) as

$$\hat{V} = \frac{1}{N}\sum_{i=1}^{N}\left(m(Z_i,\hat{q}_i)+\hat{\alpha}_i-\sum_{j=1}^{N}\frac{\hat{\alpha}_j}{N}\right)\left(m(Z_i,\hat{q}_i)+\hat{\alpha}_i-\sum_{j=1}^{N}\frac{\hat{\alpha}_j}{N}\right)'$$

where $\hat{\alpha}_i$ is the correction term accounting for $\hat{q} = \left(\hat{q}_{1,a},\hat{q}_{1,b},\hat{q}_{0,a},\hat{q}_{0,b}\right)$ in $\hat{m}_N$ and is given by

$$\hat{\alpha}_i = \sum_{k=1}^{N}\begin{pmatrix} \dfrac{X_kD_kT_kP_kK_h(P_k-P_i)\left(1-\dfrac{D_iT_i}{\hat{q}_{1,a}(P_k)}\right)}{\sum_{j=1}^{N}D_jT_jK_h\left(P_j-P_k\right)} - \dfrac{X_k(1-D_k)T_kP_kK_h(P_k-P_i)\left(1-\dfrac{(1-D_i)T_i}{\hat{q}_{1,b}(P_k)}\right)}{\sum_{j=1}^{N}\left(1-D_j\right)T_jK_h\left(P_j-P_k\right)} \\ \dfrac{X_kD_kT_kP_kK_h(P_k-P_i)\left(1-\dfrac{D_iT_i}{\hat{q}_{1,a}(P_k)}\right)}{\sum_{j=1}^{N}D_jT_jK_h\left(P_j-P_k\right)} - \dfrac{X_kD_k(1-T_k)P_kK_h(P_k-P_i)\left(1-\dfrac{D_i(1-T_i)}{\hat{q}_{1,b}(P_k)}\right)}{\sum_{j=1}^{N}D_j\left(1-T_j\right)K_h\left(P_j-P_k\right)} \\ \dfrac{X_kD_k(1-T_k)P_kK_h(P_k-P_i)\left(1-\dfrac{D_i(1-T_i)}{\hat{q}_{1,b}(P_k)}\right)}{\sum_{j=1}^{N}D_j\left(1-T_j\right)K_h\left(P_j-P_k\right)} - \dfrac{X_k(1-D_k)(1-T_k)P_kK_h(P_k-P_i)\left(1-\dfrac{(1-D_i)(1-T_i)}{\hat{q}_{1,b}(P_k)}\right)}{\sum_{j=1}^{N}\left(1-D_j\right)\left(1-T_j\right)K_h\left(P_j-P_k\right)} \end{pmatrix}$$

DOI: 10.1002/jae

APPENDIX: B HISTOGRAM OF ESTIMATED PROPENSITY SCORES

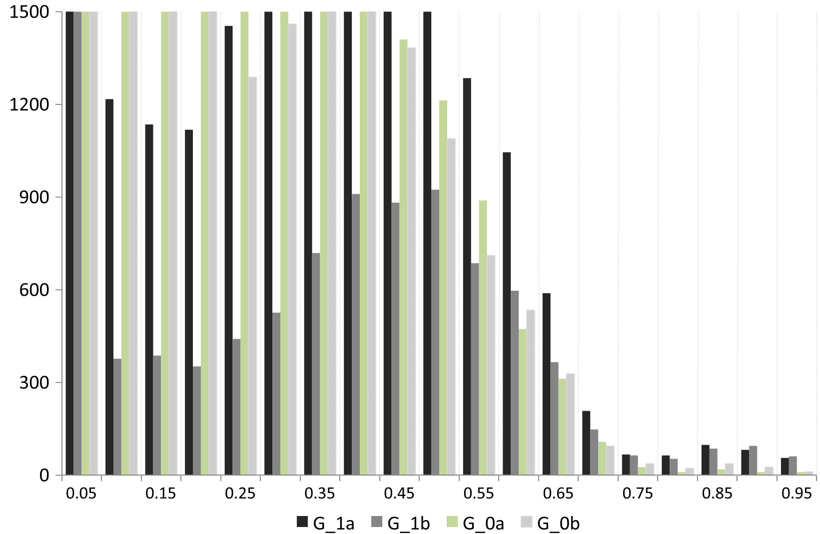

Figure 4. Histogram of estimated propensity scores. The x-axis indicates each bin of estimated P_b near the value, and the y-axis represents the frequency at each bin. To show the bin with a small number of observations, the histogram is cut off at 1500. G_{1a} denotes Internet users in period a

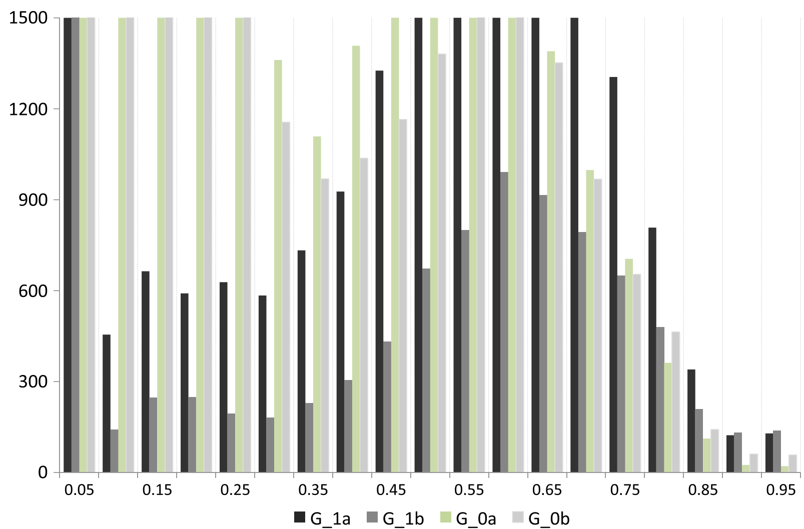

Figure 5. Histogram of estimated propensity scores P_a. The x-axis indicates each bin of estimated P_a near the value, and the y-axis represents the frequency at each bin. To show the bin with a small number of observations, the histogram is cut off at 1500. G_{1a} denotes Internet users in period a

DOI: 10.1002/jae