Measuring Network Effects in a Dynamic Environment *

Gautam Gowrisankaran
University of Arizona
HEC Montréal
NBER

Minsoo Park
Chung-Ang University

Marc Rysman
Department of Economics
Boston University

March 7, 2014

Abstract

This paper proposes methods for identifying indirect network effects with dynamically optimizing consumers purchasing a durable hardware good and associated software. We apply this model to data drawn from the DVD player and titles markets. We observe model-level prices, sales and characteristics of DVD players and sales and availability of DVDs monthly for 10 years. We augment these aggregate data with household survey data on player holdings. In our model, forward looking consumers buy possibly multiple DVD players over time and benefit from the evolution of the titles market. We provide a framework for addressing a series of econometric problems which have not been systematically addressed before.

1 Introduction

This paper proposes methods for identifying complementarity between a durable hardware good and associated software in the context of dynamically

*We with to thank the NPD Group, ICR-Centris, Home Media Magazine, and the DVD Release Report for help in securing data. Da Young Sin, Danny Chan, David Rapson, Kathleen Nosal and Jonathan Smith provided excellent research assistance. We thank the NET Institute for financial support, as well as NSF Grant SES-0551348.

optimizing consumers. We apply these methods to data drawn from the DVD player and titles markets. We observe model-level prices, sales and characteristics of DVD players and sales and availability of DVDs at the level of the month for 10 years. We augment these aggregate data with household survey data on player holdings. In our model, forward looking consumers buy possibly multiple DVD players over time and benefit from the evolution of the titles market. We provide a framework for addressing a series of econometric problems which have not been systematically recognized before.

Network effects have been an important theoretical development in economics and create numerous policy issues. For a recent discussion, see Farrell & Klemperer (2007). Our work builds on the literature that has considered the estimation of network effects. The most successful of these papers focus on static environments and exploit cross-sectional variation in data (see Saloner & Shepard, 1995; Rysman, 2004; Ackerberg & Gowrisankaran, 2006; Gowrisankaran & Stavins, 2004). [1] However, much of the motivation and impetus for studying network effects has been with regards to product diffusion over time, usually with high technology products, as in the early theory literature such as Farrell & Saloner (1986). Addressing estimation in this environment is the goal of this paper. A number of papers have taken on these issues before us, studying for example the diffusion of video cassette recorders, compact disc players, and video games. Early examples are Park (2004) and Ohashi (2003) for VCRs. [2] These papers do not address several issues that we view as important. Typically, these papers use static demand models even though the goods in question are durable. While a few papers have a dynamic interpretation, they do not address the time series feature of the data or do not account for the mismatch between a dependent variable that exhibits panel variation (such model-level sales) and an independent variable that varies only in the time-series (such as the number of software titles).

Arguably, the closest paper to ours is Lee (2013), which like us specifies a dynamic model of consumer demand for hardware (in this case, video game consoles). Lee differs from us in that he also specifies a structural

[1]More recent examples of explicitly static demand systems with an element of positive feedback loops are Fan (2013) for newspapers, Jeziorski (2013) for radio stations and Rysman (2007) for payment cards. An important early citation on newspapers is Rosse (1970).

[2]There is now a relatively large literature estimating models of diffusion in markets with indirect network effects. A partial list of more recent examples is Clements & Ohashi (2005), Derdenger (2013), and Corts & Lederman (2009) for video games, and Nair, Chintagunta & Dube (2004) for personal digital assistants.

model of demand for the complementary good, video games. This setup is appropriate for the questions of interest in the paper, which center around exclusive dealing. However, the strong assumptions on the software side of the market preclude flexibly studying the time series structure of the data in the way we envision, and Lee does not directly address the endogeneity of the two markets (although Lee argues that endogeneity is not important in his context). [3]

Formally, network effects exist when the value of a product depends on how many other consumers adopt or use the product. An indirect network effects exists when the value comes through some complementary good. For instance, the value of hardware depends on the provision of software, which typically depends on consumer hardware adoption. In this paper, we focus only on the effect of the DVD titles market on adoption of DVD players, and not the reverse effect. Thus, we identify a complementarity rather than the full "positive feedback loop" associated with network effects. Clearly, these issues are related, which is why we orient our paper around the contribution to the network effects literature.

2 Overview

We identify four important econometric problems with estimating network effects in a dynamic durable-goods environment, and then we propose methods for addressing these problems. The issues are as follows:

1. Dynamics: An appropriate model recognizes the dynamic nature of consumer decision-making. Consumer choice is affected by the durability of the goods and the fact that consumers can wait until a future date, and typically obtain similar quality for a lower price and realize higher values of the complementary good.

2. Hierarchical variation: In most econometric analysis of markets with network effects, we observe a panel of hardware products but we have only limited variation in the measure of the complementary good. For instance, we observe sales for several hundred DVD players in a month

[3] Two other important citations are and Gandal, Kende & Rob (2000), who also specify a structural model the effect of software on hardware adoption in the case of compact disks. They model the CD player market as perfectly competitive and so eliminate many of the modeling issue we face. Also, Dube, Hitsch & Chintagunta (2010) focus on tipping in the video game market. Also, Inceoglu & Park (2004) and Park (2008) provide earlier attempts to address time series issues in DVD diffusion.

but the number of titles varies only in the time series. If we are comparing DVD sales to VCR sales, perhaps we observe two measures of titles per period but the issue remains largely the same.

3. Spurious correlation: Under almost any diffusion process, we would expect sales of DVD players and DVD titles to increase over time even if they did not have a causal relationship. Since sales of both are correlated in time, a naive regression of one on the other will find a positive coefficient and falsely conclude a causal relationship. There is a second sense in which spurious correlation may be an issue. As is well-known in the time series literature, regressing one series on another may find spurious correlation if both series contain unit roots.

4. Endogeneity: Since sales of DVD titles and players are determined simultaneously and endogenously, we expect any regression to exhibit problems of endogeneity. For instance, an unobserved shock to the demand for DVD players may lead movie producers to introduce more DVD titles, creating reverse causality in our estimation equation. [4]

We propose a method that addresses these four issues. In order to address the first problem, we use a structural dynamic model of consumer behavior. In particular, we adapt the model of Gowrisankaran & Rysman (2012) to our context. Gowrisankaran & Rysman (2012) allows for persistently heterogeneous consumers to purchase one of the available products or wait based on rational expectations about the future evolution of market characteristics. The model is designed to be applied to aggregated data such as ours and allows for endogenous prices and changes in the number of products over time. We adapt the model to allow for a complementary good and importantly for our purposes, to allow consumers to hold multiple products, whereas Gowrisankaran & Rysman (2012) requires consumers to hold no more than one unit of a product at a time.

To address the second problem (hierarchical variation), we recognize that it is akin to the problem confronted in the treatment effects literature, in which researchers often employ panels with thousands of households to study policy changes that vary only across states. State-time shocks make the proper construction of standard errors challenging in this context. Moulton

[4]To clarify, we view spurious correlation and endogeneity as separate and distinct problems. For instance, sales of Commodore 64 computers and mini-vans exhibit spurious correlation. They were introduced at similar times and exhibited growing sales over time, although there was no endogeneity between them. In contrast, endogeneity could be realized in a purely cross-sectional data set, but not spurious correlation.

(1990) argues that clustering standard errors addresses this problem. However, Donald & Lang (2007) argue that clustering is not sufficient. They argue that we must consider asymptotics at the level of our policy variation. For instance, if we observe only 4 combinations of state and time, we should make inference on the policy effect as if we had only 4 observations. Donald & Lang (2007) recommend estimating with state-time dummies in a first stage and then regressing the dummies on the policy variables in the second stage. Although the second stage has many fewer observations than the first, it actually gives the correct standard errors.

Donald & Lang (2007) address only the treatment effects literature and focus on asymptotics with small numbers of observations. In our context, the "policy" variables are outcomes from the titles market. Since we observe more than 100 periods of data, we do not have the "small-numbers" problem associated with Donald and Lang. However, since the variable is a time series, we have a separate problem: asymptotic inference must account for the issues raised in time series econometrics.

Following Donald & Lang (2007) in the treatment effects context, we introduce time dummy variables into our structural model of demand for DVD players. As we show formally below, the month dummies can be interpreted as the expected current and future network benefits to a consumer at a given time, plus any other features that vary only in time. Importantly, we construct our structural model so that the addition of month dummies does not significantly increase the computational time of estimating our model. Because titles, and hence expectations about current and future titles, vary only over time and not cross-sectionally, the time dummies in our model capture the complementary goods part of utility. [5]

The structural model is a "first stage" in our estimation procedure, designed to provide us with a set of coefficients on time dummies. In our "second stage", we regress this sequence of dummy variable coefficients on variables from the titles market using standard time series techniques. Doing so allows us to deal with the third problem, spurious correlation. The second stage is a purely time series regression so we can incorporate standard tools from time series econometrics to address spurious correlation. We test for integration and heterogeneity of various orders and in particular, cointegration between the time dummy coefficients and titles variables. There is a growing recognition that time series issues should play an important

[5] The idea of using dummies in a first stage that are treated as endogenous variables in a second stage clearly predates Donald & Lang (2007). Their contribution is to recognize the implications for asymptotic approximations in the treatment effects literature, which we extend to our case.

role in microeconometric studies, and this paper contributes to that stream of research (see Bertrand, Duflo & Mullainathan, 2004; Angrist & Pischke, 2009).

The fourth issue is endogeneity. We turn to the feature film market to provide instruments. At least early in the product life when DVD sales were relatively small, activity in the film market can be characterized as exogenous to the DVD market. Chiou (2008) shows that the time period between a film's introduction and the release of a DVD varied between 5 months over our time period, and we have independent data to study this. Hence, intuitively our instrumenting assumption is that if we see sales of DVD players shifting up 4 to 6 months after a big weekend at the box office, we assume that this is happening through the titles market and is evidence of a network effect. That is, the box office affects titles but is otherwise excluded from affecting the player market.

Our goal is to be very flexible about the way that the titles market might affect the player market. An advantage of our approach is that the only computationally expensive step is the structural first stage. The step we wish to be flexible in, the relationship between the time dummy coefficients and the variables capturing the titles market, is computationally cheap. Not only can we try many forms of time series processes, but we can also experiment with different summary statistics from the titles market. A common question in this literature is about the appropriate measure of activity in the titles market: Is what matters sales of titles, the number of titles, the presence of a big hit or multiple big hits? Since we have data on each of these variables and specifications are computationally cheap – and yet consistent with dynamic optimization – we can explore all of these.

Overall, estimation of network effects models in the canonical dynamic, durable goods setting presents serious econometric challenges. We propose a polyglot method, drawing on ideas from structural micro-econometrics, treatment effects, time series and instrumental variables to address these problems. Our method addresses each of the important problems that we have identified and allows the researcher a great deal of flexibility in studying the role of network effects.

3 Data

Our data set is drawn from a variety of sources. The centerpiece comes from the NPD Group and contains monthly level observations on price and sales for DVD players, at the level of the model. We have data from March 1997 to October 2006, a long panel that reaches back to what was essentially the start of the industry. These data are drawn from relationships that NPD has with a large set of consumer electronics retailers, but unfortunately does not include WalMart or on-line sales.[6] For each model, we collected characteristics by hand based on web searches. For DVD players, characteristics are typically dummy variables for features, such as progressive scan or DTS audio capability. We also collected volume and weight although we restrict ourselves to console DVD players, as opposed to portable DVD players so this should be less important. We do not have data on other items, such as personal computers, that also have DVD capability.

Figure 1 shows the number of models that appear in our data each month over time. The growth in the number of products is startling. We observe 9 products in the first month of the data, which increases almost monotonically to show more than 350 products throughout 2006. Figure 2 shows sales (in units) by month. Like many consumer electronics products, a great deal of sales of DVD players takes place during the holiday shopping season in the fourth quarter. Conditional on that, DVD sales climb from 1997 to 2004. Interestingly, sales level off and begin to decline after 2004. That is, we observe a sort of maturation of the DVD market in our data set. Prices make a dramatic decline. Figure 1 graphs the sales-weighted average price normalized to 2000 dollars. It reaches a high in the third month of data at $766.30 and drops just below $100 in the final year of data.

DVD players are only useful with DVD titles. We have obtained a monthly time series from January 2001 to September 2008 on sales of prerecorded DVDs from the Research Department of Home Media Magazine, which uses information from Nielson VideoScan. Like NPD, Nielson's information comes from relationships with retailers, but does not include WalMart. In this case, Home Media infers WalMart sales based on their research. Figure 3 graphs this time series. Similar to DVD players, the series exhibits exaggerated holiday sales and a leveling off of sales growth around 2004. The titles data overlaps with our player data from six years, from 2001 to 2006.

We have also obtained data on counts of the number of available titles. In fact, we have a comprehensive data set on the release date of each title, as well as some characteristics such as genre. Hence, we know not only the number of DVD titles but their identity. In this paper, we focus on

[6]To be specific, NPD imputes sales at retailers that are not part of its survey, but does not attempt to impute sales at WalMart or wholesale clubs such as Costco, or on-line sales.

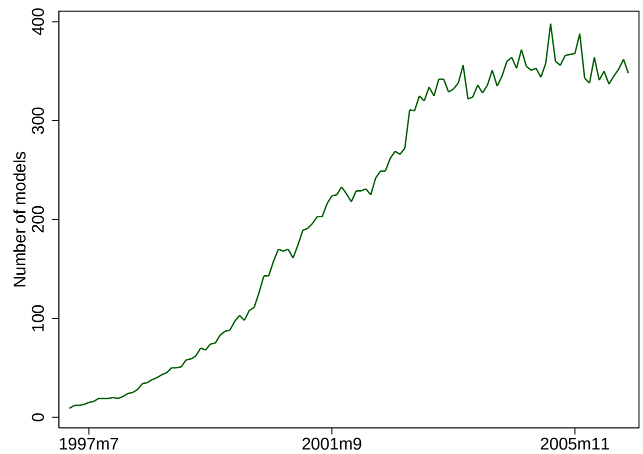

Figure 1: Number of models by month

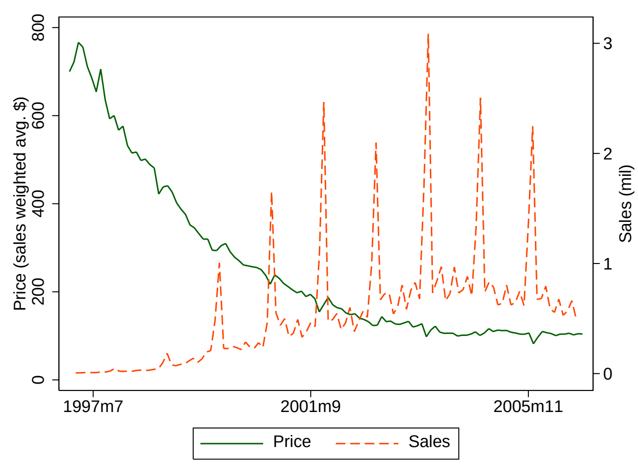

Figure 2: Number of units sold by month and average price

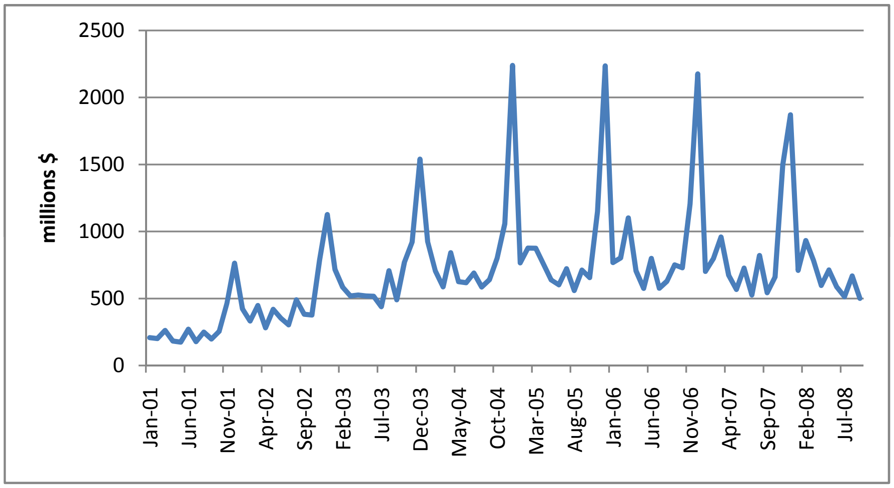

Figure 3: Sales of DVD titles by month

DVD titles associated with recent movies by restricting ourselves to DVD titles that are released within one year of the associated theatrical releases. [7] This has the effect of ignoring other types of DVDs, such as exercise videos, television shows, and releases of older movies on DVD. As mentioned below, we experiment with other measures of the DVD titles market, but we find this one to be most important. It is displayed in Figure 4. In practice, we work with the series after deseasoning according to the X11 procedure developed by the U.S. Bureau of the Census. That is the dotted line.

Sales of DVD titles are likely to be endogenous to sales of DVD players. As an instrument, we use outcomes from the cinema release market. We have obtained box office revenue and the number of movies released from Box Office Guru, an on-line source of movie information. We observe weekly data from the last week of 1995 to the 7th week of 2008. Figure 5 displays this variable. It is highly variable from week to week and displays less seasonal variation than the other variables.

Finally, households that make multiple purchases play an important role in our model. However, it is questionable whether one can infer the preva-

[7] NPD collects sales data using an Atkins formula, in which sales of the first four weeks of a quarter are allocated to the first month, the next four weeks go to the second month and next five make up the third month, repeated for each quarter of the year. An advantage of the Atkins approach is that each month contains the same number of weekends across years, making them more comparable. We follow this approach in constructing the number of DVD titles released each month.

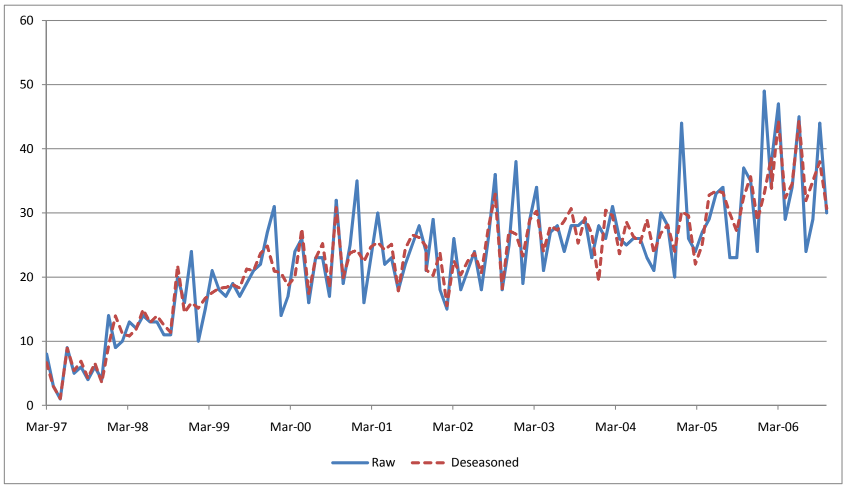

Figure 4: Number of DVD titles associated with a recent movie

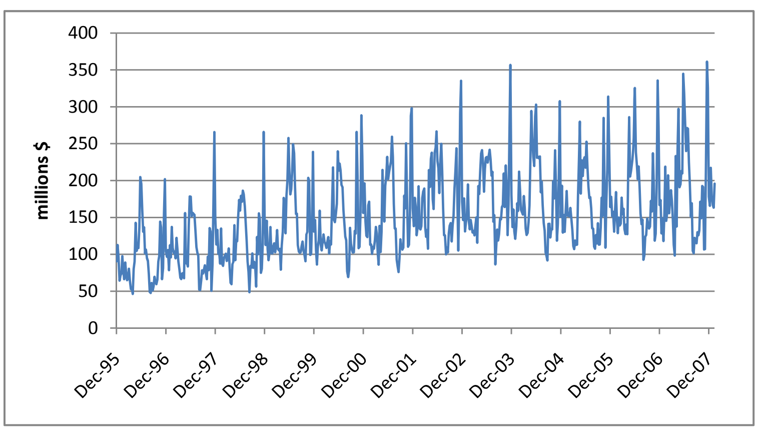

Figure 5: Box office revenue for films by week

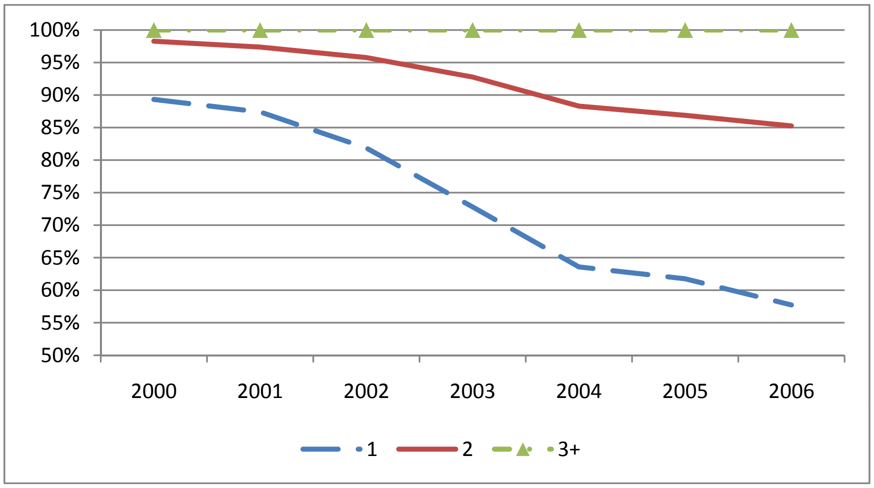

Figure 6: Number of DVD players in a household among those that have at least one

lence of multiple purchases from aggregate data on sales. To make progress, we use a data set from Centris of ICR, a market research firm. Centris performs a telephone survey based on random digit dialing of consumer holdings of consumer electronics. They complete about 4,000 surveys per month. They specifically ask each household how many console DVD players they hold, and they report the percentage of households that hold 0, 1, 2 or more than 2 console players. We obtained data for the third quarter of each year from 2000 to 2006. That data appears as a stacked line chart in Figure 6. Among households that have at least one DVD player, 87.9% have only one in 2000. This number drops to 56% by 2006, with the number reporting that they have more than 2 climbing from less than 2% to 14.3%.

4 Structural Model

Here, we present our model of consumer demand that allows us to account for the issues we describe above. The model builds on Gowrisankaran & Rysman (2012) by extending it to allow for complementary goods and for households to hold multiple products. Our model starts with the introduction of a new consumer durable good at time $t = 0$. The unit of observation is a month. We begin with a single consumer and add heterogeneity in the next subsection. Consumers have infinite horizons and discount the future with a common factor β. The consumer chooses one of among J_t products in each period t or chooses to purchase no product in the current period. From these $J_t + 1$ choices, the consumer chooses the option that maximizes the sum of the expected discounted value of future expected utilities conditional on her information at time t.

Product j at time t is characterized by a flow utility f_{jt} (which depends on its observable and unobservable characteristics), price p_{jt}, and an "environmental variable" N_t. For DVD players, observed characteristics include the presence of advanced sound and display features such as Dolby audio and progressive scan. In our context, the environmental variable describes the title market, for instance the number of titles available at time t. The environmental variable that the consumer obtains from purchase is allowed to change over time. We assume that consumers and firms know all time t information when making their time t decisions. The additional flow utility to purchasing product j in period t is:

$$u_{jt} = f_{jt} + g(N_t, \theta) + \varepsilon_{jt}$$

Here, ε_{jt} is distributed independently across consumers, products and time according to the Type I Extreme Value distribution, creating the familiar logit demand system. The consumer knows only the current set of ε_{jt}, not future values. The function $g(N_t, \theta)$ captures the value of the complementary goods market.

Now we turn to multiple holdings. We assume that products are infinitely durable. That is, a product j purchased in t delivers stand-alone utility f_{jt} every period thereafter. We refer to f_{jt} as the "stand-alone" flow utility for two reasons: it does not depend on N_t and it is the amount the consumer would get if the consumer held only that DVD player. In combination with other DVD players, the contribution of the product may differ. Let F_{0t} represent the vector of values f_{jt} for all products the consumer has purchased up to time t (as in GR, we use the "0" notation to indicate holdings). The consumer derives flow utility $u(F_{0t})$ from these holdings. When the consumer purchases product j in period t, consumer holdings become $F_{jt} = \{f_{jt}, F_{0t}\}$.

Thus, purchase product $\{j, t\}$ when holding F_{0t} delivers flow utility of:

$$u_{jt} = u(F_{jt}) - p_{jt} + (n+1)g(N_t, \theta) + \varepsilon_{ijt}.$$

Here, n_t is the number of DVD players that the consumer holds (the number of elements in F_{0t}. No purchase earns:

$$u_{0t} = u(F_{0t}) + ng(N_t, \theta) + \varepsilon_{i0t}$$

That is, we assume that the household captures $g(N_t, \theta)$ for each DVD player that the consumer holds. Note that the value of the environmental variable can change over time whereas the value of f_{jt} cannot. That is, a consumer who purchases product $\{j, t\}$ gets f_{jt} in F_{0t} and receives an extra $g(N_\tau, \theta)$ in all periods $\tau > t$.

We now turn to dynamics. Let Ω_t denote the state of the market, which is made up of current product attributes and any other factors that influence future product attributes. We assume that Ω_{t+1} evolves according to some Markov process $P(\Omega_{t+1}|\Omega_t)$ that accounts for firm optimizing behavior. We specify a stationary problem, so we write the Bellman equation without the time subscript, using the prime symbol to represent values in the next period. The relevant state variables for the consumer are holdings F_0, the vector of draws $\overrightarrow{\varepsilon}$ and the state of the market, Ω.

$$\widetilde{V}(F_0, \Omega, \overrightarrow{\varepsilon}) = \max$$

$$\max_{j=1,\ldots,J} u(F_j) + (n+1)g(N, \theta) - p_j + \varepsilon_j + \beta E\left[\widetilde{V}\left(F_j, \Omega', \overrightarrow{\varepsilon}'\right)|F_j, \Omega, \overrightarrow{\varepsilon}\right],$$

$$u(F_0) + ng(N, \theta) + \varepsilon_0 + \beta E\left[\widetilde{V}\left(F_0, \Omega', \overrightarrow{\varepsilon}'\right)|F_0, \Omega, \overrightarrow{\varepsilon}\right] \qquad (1)$$

Line 2 represents the value of buying and line 3 represents the value of not buying.

The value function $\widetilde{V}(F_0, \Omega, \overrightarrow{\varepsilon})$ is too large for us to work with numerically, so we use various techniques to simplify it. We do so in three ways, addressing $\overrightarrow{\varepsilon}$, then F_0 and then Ω. First, because $\overrightarrow{\varepsilon}$ is *iid*, it satisfies the assumption of conditional independence (Rust, 1987) and may be integrated out. We work with $\widetilde{EV}(F_0, \Omega) = \int_\varepsilon \widetilde{V}(F_,\Omega, \overrightarrow{\varepsilon})$.

Next, we make a simplifying assumption on $u(.)$. We assume:

$$u(F_0) = \sum_{j=1}^{n} f_{0j} - \psi(n)$$

where f_{0j} are the elements of F_0 and n is the number of products held in that period. Thus, the value of holding multiple goods is the sum of their stand-alone flow utilities minus a "penalty": we interpret $\psi(n))$ as capturing the declining marginal utility of holding multiple goods. We expect that $\psi(n)$ will positive, and increasing and convex in n. There is an important assumption in this statement, which is that the decline does not depend on the flow utility of the products that the consumer holds. We return to this point below. Perhaps abusing notation, we let $f_0 = \sum_{j=1}^{n} f_{0j}$.

Thus, we can write $\widetilde{EV}(f_0, n, \Omega)$ instead of $\widetilde{EV}(F_0, \Omega)$. In fact, the assumption on $u(.)$ allows for further simplification because f_0 affects both value of purchase and no purchase, and so does not affect future decision making. Past purchases affect future decisions only through n. We can eliminate f_0 to work with a simpler value function that generates the same choices. Formally, we define:

$$EV\left(n, \Omega\right) = \widetilde{EV}\left(f_0, n, \Omega\right) - \sum_{\tau=0}^{\infty} \beta^{\tau} \left(f_0 + nE[g(N_\tau, \theta)|\Omega]\right) \tag{2}$$

The function $EV(n, \Omega)$ is equal to the original value function minus the present discounted value of the stream of flow utilities and network benefits associated with past purchases.

The appendix provides a more formal derivation, but to see why $EV(n, \Omega)$ does not depend on f^0, note that the value of purchasing product j can be written as:

$$\sum_{\tau=0}^{\infty} \beta^{\tau} \left(f_j + f_0 + (n+1)E[g(N_\tau, \theta)|\Omega]\right) - p_j - \psi(n+1) + E\left[EV\left(n+1, \Omega'\right)|n, \Omega\right]$$

and the value of no purchase is:

$$\sum_{\tau=0}^{\infty} \beta^{\tau} \left(f^0 + nE[g(N_\tau, \theta)|\Omega]\right) - \psi(n) + E\left[EV\left(n, \Omega'\right)|n, \Omega\right]$$

Comparing these two values shows that $\sum_{\tau=0}^{\infty} \beta^{\tau}(f_0 + nE[g(N_\tau, \theta)|\Omega])$ enters both the value of purchase and no purchase in the same way and so does not affect choices. Thus, we can specify a dynamic problem based on $EV(n, \Omega)$ rather than $\widetilde{EV}(n, f^0, \Omega)$ that generates the same choice outcomes. This helpful simplification is a result of our assumption that $\psi(n)$ depends only on n and not on f_0. If the decreasing marginal return of holding multiple DVD players depended on their flow utility, we would have to maintain the accumulated flow utility as a state variable.

In this set-up, purchases affect future decision-making only through the number of products in the household. Conditional on buying a product, the choice of which product to buy is static. It is as if the consumer gets all of the product-specific and network benefits "up front." Hendel & Nevo (2006) use a similar approach to simplify the choice problem in a consumer-inventory context. It is now helpful to define this up front value: the mean expected flow and network utility from purchasing j:

$$v_j = \sum_{\tau=0}^{\infty} \beta^{\tau} \left(f_j + \theta E\left[g(N_\tau, \theta)|\Omega\right]\right) - p_j \tag{3}$$

Note that $EV(n, \Omega)$ satisfies its own Bellman equation (see the Appendix for on this). Subtracting $\sum_{\tau=0}^{\infty} \beta^{\tau} \left(f_0 + n\theta E\left[N_{\tau}|\Omega\right]\right)$ from each side of Equation 1, and using the notation of v_j:

$$\begin{aligned} EV\left(n, \Omega\right) \;\; &= \;\; \max\left\{ \max_{j=1,\ldots,J} \left(v_j - \psi(n+1) + \varepsilon_j + E\left[EV\left(n+1, \Omega'\right)|n, \Omega\right]\right), \right. \\ &\quad \left. -\psi(n) + \varepsilon_0 + E\left[EV\left(n, \Omega'\right)|n, \Omega\right] \right\} \end{aligned} \tag{4}$$

Thus each period, the consumer chooses between the flow utility derived from past purchases -$\psi(n)$ or buying a good and getting v_j, and thus obtaining -$\psi(n+1)$ in the future. Assuming that $\psi(n) > 0$, $\psi'(n) > 0$ and $\psi''(n) > 0$ implies that consumers facing the same market options will be less likely to buy if the hold more goods.

Finally, we turn towards simplifying Ω_t. Our goal is to make a simplifying assumption such that we can replace Ω with a sensible scalar value. In order to do so, we separate our problem into a dynamic problem of when to purchase and a static decision of which product to purchase, conditional on having chosen to purchase. Define δ to be the expected value of the static part of the choice problem, that is, from choosing which product conditional on buying a product.

$$\delta = E_{\varepsilon}\left[\max_{j=1,\ldots,J} v_j + \varepsilon_j\right] = \ln\left(\sum_{j=1}^{J} \exp(v_j)\right) \tag{5}$$

The second equation follows from the logit assumption on ε_j and is widely used in implementing discrete choice models.

Defining δ is useful because if a consumer knew current and future values of δ, the consumer would have enough information to optimally choose when to make her next purchase. The consumer does not need to know Ω_t. That is, n, δ and the contingent path of δ are sufficient statistics to define $EV(n, \Omega_t)$. Formally,

$$EV\left(n, \Omega\right) = EV\left(n, \delta, P\left[\delta|\Omega_{\tau}\right] \forall\tau\right). \tag{6}$$

Gowrisankaran & Rysman (2012) and Melnikov (2001) prove this point formally. This result follows from assuming the logit functional form for ε_j.

Unfortunately, Equation 6 does not generate a numerical simplification since consumers still predict future values of δ using all of Ω. In order to make progress, we make an important simplifying assumption on how consumers make predictions. In particular, we assume that consumers use only the current value of δ to predict future values of δ. Following Gowrisankaran

& Rysman (2012), we refer to this as the assumption of Inclusive Value Sufficiency.

Assumption 1 *Inclusive Value Sufficiency (IVS)*

If two states Ω *and* $\hat{\Omega}$ *generate the same value of* δ*, then* $P(\delta'|\Omega) = P\left(\delta'|\hat{\Omega}\right)$ *for all future values of* δ*, for all* Ω_t*,* $\hat{\Omega}$*.*

The assumption of IVS implies that all states with the same n and δ have the same continuation value, and so consumers do not have to track Ω. Thus, the state space is reduced to two dimensions. The IVS assumption can be interpreted as an assumption that consumers are boundedly rational. They use only a subset of the data potentially available to them in forming their predictions. The assumption is restrictive. For example, δ could be high either because there are many products in the market all with high prices or because there is a single product in the market with a low price. While dynamic profit maximization might lead these two states to have different patterns of industry evolution, consumers in our model will lump them into the same state.[8]

For our specifications we assume the consumer perceives $P(\delta'|\delta)$ as its actual empirical density fitted to a simple functional form and use a simple quadratic autoregressive specification,

$$\delta' = \gamma_1 + \gamma_2\delta' + \gamma_3(\delta')^2 + \nu, \tag{7}$$

where ν is normally distributed with mean 0 and $\{\gamma_1, \gamma_2, \gamma_3\}$ are incidental parameters. By assuming that consumers make predictions based on the parameters from (7) derived from the realized values of δ, we are assuming that consumers have rational expectations, conditional on the restriction in (7).

Note that our IVS assumption and Equation 7 are statements only about exogenous items such as the current numbers of products, prices, and characteristics. In this sense, our assumptions are more similar to those in Melnikov (2001) and Hendel & Nevo (2006) than Gowrisankaran & Rysman (2012). Gowrisankaran & Rysman (2012) specify δ to include the (endogenous) continuation value as well. Gowrisankaran & Rysman (2013) provides an overview of these different approaches and their relative benefits.

The fact that δ is a function of flow utilities only follows from our model in which only the number of products has dynamic content, not the characteristics of those products. This assumption also makes computation much

[8] Hendel & Nevo (2006) and Gowrisankaran & Rysman (2012) provide a similar discussion of the implications of Assumption 1.

easier. Without this assumption, the characteristics of the product would affect not only which product the consumer chooses today but also future decision-making, and so must play a role in the value function. Thus, N_t, the features of the titles market, would affect the state of the consumers, which means we could not write the time dummies as a linear function of mean utilities and we would have to search over time coefficients non-linearly (for more on this, see Section 5), which would be infeasible. Hence, the assumption that the value function depends on the number of products held and not their characteristics generates a major computational savings. We also believe it is a reasonable assumption, but we discuss this more later.

An implication of (7) is that for low enough values of γ_2, δ converges over time to an asymptote that is approached from below. This asymptote is important in our model since it represents a steady state in the evolution of product characteristics that the consumer expects to approach. The eventual arrival of a steady state is what allows us to treat the consumer as facing a stationary environment, even though observed product offerings are evolving quickly.

The logit assumption on ε_j generates a convenient closed for solution for the Bellman equation in Equation 4. Including IVS, the Bellman equation is now:

$$\begin{aligned} EV(n,\delta) \;=\; & \ln\left(\exp\left(\delta - \psi(n+1) + \beta E\left[EV(n+1,\delta')|n,\delta\right]\right)\right. \\ & \left. + \exp\left(-\psi(n) + \beta E\left[EV\left(n,\delta'\right)\middle| n,\delta\right]\right)\right) + \gamma. \end{aligned} \tag{8}$$

5 Inference

This section discusses the parametrization and estimation of the model. Our methods for estimating the model follow closely those in Gowrisankaran & Rysman (2012) and so we cover them only briefly here.

5.1 Parametrization

We begin with a discussion of our parametrization. We allow for consumer heterogeneity, indexing a continuum of consumers by i. We also keep track of t in this section, so the variables from the previous section are indexed accordingly, $f_{ijt}, p_{ijt}, v_{ijt}$, δ_{it}, EV_i, γ_i, etc. We parameterize $f_{ijt} = x_{jt}\widetilde{\alpha}_i^x + \widetilde{\xi}_{jt}$, where x_{jt} are observable characteristics such as having progressive scan or Dolby stereo, and $\widetilde{\xi}_{jt}$ captures unobserved characteristics. The price $p_{ijt} = \alpha_i^p \ln(P_{jt})$, where P_{jt} is the observed price in the data.

In practice, we work with $\alpha^x = \widetilde{\alpha}^x/(1-\beta)$ and $\xi = \widetilde{\xi}/(1-\beta)$. This notation allows us to rewrite v_{jt} from Equation 3 as:

$$v_{ijt} = x_{jt}\alpha_i^x - \alpha_i^p \ln(p_{jt}) + \sum_{\tau=t}^{\infty} \beta^{\tau-t} E[g(N_\tau, \theta)|\delta_{it}] + \xi_{jt}$$

The variable ξ_{jt} plays the role of our econometric error term. Consumers are characterized by their demand parameters $\alpha_i = \{\alpha_i^x, \alpha_i^p\}$, which stay constant over time. Integrating over consumers i does not generate a closed-form solution for the market shares for products. Hence, we simulate consumers by drawing consumer deviations. In practice, we assume that $\alpha_i \sim \mathbb{N}(\alpha, \Sigma)$, where Σ is non-zero only on the diagonal of the matrix. We draw from the standard normal to represent consumer deviation from the mean and estimate α and Σ to create each consumer's α_i.

We do not attempt to estimate β because it is widely understood to be unidentified in dynamic decision models (see Magnac & Thesmar, 2002). We set $\beta = .95$ annually, which is $0.95^{(1/12)}$ for our monthly data.

As has been alluded to in the introduction, we do not attempt to estimate the function g in the structural model. Rather, we replace $\sum_{\tau=t}^{\infty} \beta^{\tau-t} E[g(N_\tau, \theta)|\delta_{it}]$ with a sequence of time dummy variables θ_t. Each parameter θ_t has the interpretation as the full present discounted value of the expected stream of future benefits from the titles market. As time dummies, they will also naturally pick up other time specific effects.

We experiment with several versions of $\psi(n)$. This choice is important, because it is the only source of dynamics in the model. [9] Our base specification is $\psi(n) = \psi_1 n^2$, where we estimate ψ_1. We also estimate a more flexible Box-Cox specification, in which $\psi(n) = \psi_1 n^{\psi_2 - 1}/\psi_2$, where we estimate ψ_1 and ψ_2. In this specification, $\psi(n)$ can be convex, linear or concave if ψ_2 is greater, equal or less than one, and can be upward or downward sloping depending on the sign of ψ_1. We also experiment with non-parametric representations of $\psi(n)$, where it can freely take on different values for each n.

5.2 Estimation

Thus, the parameters to estimate are $\lambda = \{\alpha, \theta, \Sigma, \psi\}$. We now turn to estimation. Central to our computational procedure is $\bar{v}_{jt}$, the mean over i

[9] If $\psi(n)$ was a constant term, consumers could make decisions statically. Also, if $\psi(n)$ was linear in n, it would have the same effect as the constant term in x_{jt} and thus be equivalent to static decision-making.

of the static utility terms:

$$\bar{v}_{jt} = E_i[v_{ijt}] = x_{jt}\alpha^x - \alpha^p p_{jt} + \theta_t + \xi_{jt} \tag{9}$$

Here, α^x and α^p are the means over i of α_i^x and α_i^p. That is, $\alpha = \{\alpha^x, \alpha^p\}$.

Our computational strategy is, for any guess of parameters, to compute the vector of $\bar{v}_{jt}$ that rationalizes the observed market shares. Based on this solution, we compute ξ_{jt} and generate a GMM objective function. The solution for $\bar{v}_{jt}$ depends on the demand of each consumer i in each period, and thus requires a solution to the dynamic programming problem.

Formally, our algorithm takes a set of parameters λ and a starting guess of $\bar{v}_{jt}$. Based on these elements and our simulation draws, we construct α_i and thus v_{ijt}. We can then construct δ_{it} based on Equation 5. Then, we perform the AR(1) regression of Equation 7 for each consumer i separately, thereby recovering belief parameters γ_i. Because we discretize δ_{it} for determining the value function, we convert the parameters γ_i to a transition matrix following Tauchen (1986). Then, for each consumer separately, we guess a starting value for the value function and solve the Bellman equation (Equation 8) by successive approximations.

Once we have value function $EV_i(n_{it}, \delta_{it})$, we are ready to solve for conditional and unconditional probabilities of purchase. Conditional probabilities of purchase are as follows. For consumer i in period t who holds n_{it} products and faces a market evaluated at δ_{it}, the probability of purchase is:

$$P_{it}(n_{it}, \delta_{it}) = \frac{e^{\delta_{it} + \beta E[EV_i(n_{it}+1, \delta_{it+1})|n_{it}, \delta_{it}]}}{e^{\delta_{it} + \beta E[EV_i(n_{it}+1, \delta_{it+1})|n_{it}, \delta_{it}]} + e^{\psi(n_{it}) + \beta E[\,EV_i(n_{it}, \delta_{i,t+1})|n_{it}, \delta_{it}]}}.$$

Conditional on purchasing in period t, consumer i picks product j with probability:

$$P_{ij|t} = \frac{v_{ijt}}{\sum_{k=1}^{J_t} \exp(v_{ikt})}.$$

In order to compute the unconditional probabilities, the market shares, define the $(T+1) \times (\overline{n}+1)$ matrix s_i for each consumer i. Here, T is the number of periods in the data set, $\overline{n}$ is the maximum number of products a consumer may hold, and s_i is the share of consumers of type i holding each number of products at each period. We index the matrix s_i from 0 to T and from 0 to $\overline{n}$. We assume that for each i, the first element row is a vector of zeros, with the first element being 1. That is, everyone holds zero products in period 0. [10] Then, we can use P_{it} to successively fill out each row of s_i.

[10] This is reasonable because our data set reaches back to the onset of the industry. For an alternative approach, see Schiraldi (2011) who estimates an initial distribution in the used car market.

For instance, $s_i[1,0] = 1 - P_{it}(0, \delta_{i1})$ and $s_i[1,1] = P_{it}(0, \delta_{i1})$. [11] Because consumers cannot buy more than one product in a period, $s_{it}[1,n] = 0$ for $n > 1$. Element t,n of s_{it} is $P_{it}(n-1, \delta_{i,t-1}) s_{it}[n-1, t-1] + (1 - P_{it}(n, \delta_{i,t-1}) s_{it}[n, t-1]$, the sum of purchasers who held $n-1$ products and non-purchasers who held n products in period $t-1$.

With these elements, we can compute market shares. The market share predicted by the model of product j in period t is:

$$\widehat{s_{jt}} = \sum_{i=1}^{ns} P_{ij|t} \left(\sum_{n=0}^{\overline{n}} P_{it}(n, \delta_{it}) s_i[t-1, n] \right).$$

Here, ns is the number of consumer types that we sample. That is, we sum over each consumer type the set of consumers holding each number of products in the previous period multiplied by the probability of choosing product j.

We use the fixed point equation of Berry, Levinsohn & Pakes (1995) to generate a new guess for $\bar{v}_{jt}$. In vectors, where s^0 is the observed data, $\bar{v}$ is the vector of elements $\bar{v}_{jt}$ and $\widehat{s(\bar{v}, \lambda)}$ is the resulting market shares:

$$\bar{v}' = \bar{v} + \ln(s^0) - \ln(\widehat{s(\bar{v}, \lambda)})$$

Thus, we iteratively compute $\bar{v}$ that satisfies this equation, and thus generates the observed market shares. Although we cannot prove that there is a unique solution, we have not had any problems with convergence. Gowrisankaran & Rysman (2012) discusses this issue further. In practice, it is not necessary to solve the Bellman equation to convergence before computing a new value of $\bar{v}$, and we have found computational advantages to switching between these equations more often.

Based on the resulting vector $\bar{v}$, we compute $\xi_{jt}(\lambda)$ according to Equation 9. We form moments with the resulting vector ξ_{jt} using instruments z_{jt}. Thus, our objective function is:

$$\hat{\lambda} = \arg\min_{\lambda} \left(z'\xi(\lambda) \right)' W \left(z'\xi(\lambda) \right). \tag{10}$$

As is standard, we obtain GMM estimates in two steps. We first set $W = z'z$, which is efficient under the assumption of homoskedastic errors and generates consistent estimates, and then we construct the efficient weighting matrix allowing for arbitrary heteroskedasticity.

[11] We use the notation $s_i[l,m]$ to denote the element in row l and column m of matrix s_i.

We also incorporate micro-moments in the spirit of Petrin (2002) and Berry, Levinsohn & Pakes (2004). We do not use the survey data to establish how many households have purchased, as we are concerned that because our data set does not cover all retailers, it may mismatch in this dimension. Instead, we use survey data to determine holdings among households that hold at least one player. We use the ICR survey mentioned above to identify 2 moments at 7 time periods for 14 moments: the percentage of households holding one console DVD player amongst those holding a console DVD player annually from March 2000 to March 2006, and the percentage holding two. The remaining households hold three or more. We compute the equivalent moments by summing over consumer types with the appropriate row of s_i. We include the difference between the model's predictions and the ICR data as moments, vertically concatenated onto $z'\xi$ in Equation 10. We expand the weighting matrix by 14 elements in each dimension. The diagonal elements of the weighting matrix should be the inverse of the variance of the moment. For variance, we use $(p)(1-p)/4000$, where p is the value of the moment in the data, and 4000 is the approximate number of households sampled in each period. As this variance is very small, our weighting matrix puts a high weight on the micromoments so our estimation algorithm attempts to match these very closely.

The vector $\bar{v}$ is a function of Σ and ψ only, and based on $\bar{v}$, we can solve for the optimal values of α and θ using linear techniques. Thus, we perform non-linear search only over Σ and ψ, similar to the recommendation of Nevo (2000) in the static case. Being able to solve for the time dummies coefficients in this way, as opposed to searching for them non-linearly, is important since there are a great number of them. Two assumptions that are important in generating this result are that the consumer obtains the benefit θN_t for each DVD player that the consumer buys and that there is no heterogeneity in θ. While these assumptions are restrictive, we find them defensible, as well as necessary to take our two-stage approach to identifying the network effect. [12]

In practice, we draw 48 consumers ($ns = 48$). We discretize δ_{it} into 50 bins stretching from -40 to 0, which is much greater than the span of what we observe in our model. We set the maximum number of products a household can hold to 4 ($\overline{n} = 4$). In our results, less than 2% of households hold four

[12] This issue is related to the work of Hendel & Nevo (2006), who assume there is no heterogeneity in the value of any product characteristics, which allows for a numerically attractive two-step estimation procedure for the structural model. This approach is not possible for our paper since we still allow heterogeneity on characteristics (particularly price), but the lack of heterogeneity in θ is what allows for our two-step procedure.

DVD players at the end of the sample. We use importance sampling as described in Gowrisankaran & Rysman (2012) to reduce sampling error. We assume there are 100 million households in the United States during this time period, although in practice this changes from about 95 million to 105 million. Incorporating a growing market is straightforward but we have not done this yet.

We search using non-derivative methods such as the Nelder-Mead algorithm and direct search techniques. All programs are available on request.

5.3 Second Stage

Now consider the complementarity, represented by the function g above. Rather than identify g from the structural model, we identify complementarity from correlation between the month dummy coefficients (θ) and the exogenous variables representing the DVD titles and feature film markets. We have little prior knowledge of how the titles markets affect player markets, and hence we propose a model that allows allows a great deal of flexibility and low-cost specification searching over this issue while still capturing dynamic consumer behavior appropriately. In theory, whatever specification we find to be superior could be imposed in the structural model and we could estimate θ in the context of the structural model.

We approach our estimation of the second stage from the perspective of the literature on structural vector autoregressions. Let θ_t be the time dummy coefficients arising from the structural model, and let N_t be the variable describing the DVD titles market, for instance, the number of new titles appearing in month t. We consider two simultaneous equations. This papers seeks only to estimate the first equation, but the second is useful for expositional purposes:

$$\theta_t = \beta_0 + \beta_1\theta_{t-1} + \beta_2 N_t + \beta_3 t + u_t \tag{11}$$

$$N_t = \gamma_0 + \gamma_1 N_{t-1} + \gamma_2\theta_t + \gamma_3 t + \gamma_4 z_t + v_t \tag{12}$$

The parameters β and γ are not meant to refer to the parameters in the previous section. Note that Equation 12 may be too simple in the sense that if N_t is the number of new DVD titles appearing in period t, it may depend on outcomes from the DVD player market (θ_t) from many periods ago. However, since we do not estimate Equation 12, it is not necessary to explore this issue further here.

In the VAR literature, the fact that θ_t and N_t enter contemporaneously into the determination of the other (via β_2 and γ_2) is what makes the system “structural”. The “reduced-form” system would depend only on lagged values, which are taken as exogenous. In order to accept lagged values as exogenous, we must have that u_t and v_t do not exhibit autocorrelation. We test for this feature below. In fact, the research in the structural VAR literature usually achieves identification by restricting the correlation structure between u_t and v_t. In contrast, we introduce an excluded variable z_t, that provides identification of the first equation. As mentioned above, z_t is drawn from the movie market, particularly box office outcomes. We use them with a 5 month lag. As mentioned above, Chiou (2008) shows that there is about a 5 month lag between the release of a movie and the release of a DVD.

In addition to testing for autocorrelation, we test that u_t does not exhibit a unit root, which is analagous to testing that θ_t and N_t are co-integrated. Co-integration implies that we can use standard asymptotic approximations to make inference about the β parameters.

Finally, note that the autoregressive structure of Equation 11 is particularly appealing in our context. While including $\beta_1\theta_{t-1}$ in Equation 11 is standard in the VAR literature in order to achieve good fit, it also has a natural interpretation in the DVD market. We can interpret θ_t as the ”accumulated capital” from the addition of new titles to the DVD market. Finding that $\beta_1 < 1$ (as we find below) implies that consumers value current titles more than past titles. Then $1 - \beta_1$ can be interpreted as a depreciation rate on the value of past titles. This depreciation may arise either because current movies are more inherently valuable, or because consumers have seen older movies and no longer value them.

6 Results

In this section, we present the results of our model. Our results are preliminary. We discuss these issues and propose some possible problems in our approach so far.

We estimate the model described above. At our estimated parameters, our model predicts that less than 2% of consumers hold 4 DVD players so we do not view this as a binding constraint.

We include brand dummies in our model. In practice, we aggregate brands with less than 70 observations (for instance, 5 models for one year would be 60 observations) into a single brand. This aggregate brand still accounts for less than 5% of the observations.

Results appear in Table 1. We provide results from two specifications. In column 1, we define $\psi(n) = \psi_1 \ln(n)$. We find $\psi = 0.108$, which is precisely estimated to be different from zero. In column 2, we estimate the Box-Cox specification. Recall that column 1 can be seen as a special cost of column 2 with ψ_2 restricted to be zero. Results are similar across the two specifications, although the estimate of $\psi_2 = 0.82$ suggests that the $\psi(n)$ is not overly concave. In what follows, we focus on column 1.

We include the price in logs. Note that it is difficult to justify log price in a utility function and most other similar papers use price in levels, for instance Berry et al. (1995). However, logit based models can be interpreted as log-linear models (see Berry, 1994), so log right-hand side variables seem natural, and we find that price in logs fit the data better. We plan to experiment with price in levels as well. Logged price is negative and significant, with a coefficient of -1.539. We allow for two random coefficients. The first is on price. We find a find a coefficient of 0.883, which is significant and particularly large relative to the price coefficient.

The standard deviation in the constant term is estimated to be 6.638, which indicates substantial heterogeneity in consumer valuation for DVD players as a class of products. Note that in a static discrete choice demand model, heterogeneity in the constant term cannot be separately identified from time dummies. In a static model, heterogeneity in the constant term is identified by consumers switching from the outside product to the inside group of products, but this is precisely what is captured by time dummies. However, these effects can be separated in our dynamic model. Assuming that the flow utility of holding DVD players is constant across consumers, heterogeneity in the constant term explains the spread of consumers across different holding states. For instance, a large parameter on the constant term could generate a bimodal distribution of holdings, where consumers either hold many players or none, but time dummies could not generate this outcome. Thus, the micromoments are crucial for identifying not only ψ but also σ_1.

We include a series of dummy variables to capture observable quality. For example, we include indicators for whether the DVD can play Dolby Digital audio, whether it can play MP3 files, and whether it can hold multiple discs simultaneously. All of these coefficients should be positive since they each indicate quality. In practice, we find four of twelve coefficients to be negative, three significantly so. Interestingly, whereas most quality characteristics become more prevalent over time in our data, the four characteristics with negative coefficients either stay constant or become less prevalent. This is because all four of these variables are associated with the use of the DVD player for playing music. For example, the indicator "multi-disk" means that the player holds multiple disks at once that the user can

Linear Parameters						
constant	-26.152	(0.972)	*	-37.850	(6.589)	*
ln price	-2.419	(0.155)	*	-1.593	(0.089)	*
S-video output	0.565	(0.057)	*	0.573	(0.057)	*
Composite video output	0.009	(0.053)		0.055	(0.052)	
optical digital audio output	-0.361	(0.044)	*	-0.442	(0.042)	*
coaxial digital audio output	-0.076	(0.047)		-0.074	(0.046)	
Built-in Dolby Digital audio decoder	-0.306	(0.055)	*	-0.265	(0.054)	*
Built-in Digital Theater Systems decoder	0.286	(0.055)	*	0.245	(0.055)	*
Plays CD R/RW	0.150	(0.030)	*	0.136	(0.029)	*
Plays MP3 files	0.602	(0.064)	*	0.552	(0.063)	*
plays VHS	0.869	(0.052)	*	0.758	(0.048)	*
progressive scan (higher picture quality)	0.359	(0.045)	*	0.331	(0.044)	*
Records to DVD	0.152	(0.057)	*	0.078	(0.054)	
multi-disk	-3.01E-01	(3.9E-2)	*	-3.71E-01	(3.8E-2)	*
Non-linear parameters						
constant std. dev.	3.044	(0.850)	*	0.755	(1.616)	
ln price std. dev	1.329	(0.089)	*	0.830	(0.068)	*
psi 1	0.110	(0.007)	*	0.140	(0.030)	*
psi 2	0			0.820	(0.060)	*
11,534 observations. Includes time and brand fixed effects.						

Table 1: Results from structural estimation

select among, and the others indicate sound quality. These features were valuable early in the sample period when DVD players were popular as a substitute for CD players. However, later in the time period, CDs fell out of favor relative to MP3 players such as the Apple iPod. Table 2 displays the sales weighted average of three quality indicators over time. Progressive scan refers to picture quality, it climbs over time and we estimate a positive coefficient. However, for multi-disk and digital optical audio, we obtain different results. While perhaps a positive coefficient should be measurable from cross-sectional variation, it is interesting that our model struggles to match characteristics with these time trends. In the future, we plan to interact characteristics with time in order to capture these changing preferences.

Finally, we estimate time dummies. These appear in Figure 7. The time dummies slope up and then level off. Interestingly, they do not turn down, as does the underlying sales data. This result is a feature of our dynamic durable goods model. By the end of the sample, many consumers already hold the good, so it requires level time dummies just to maintain the falling sales we see in the data. This sequence of time dummy coefficients corresponds with our prior about the DVD titles market, which was improving

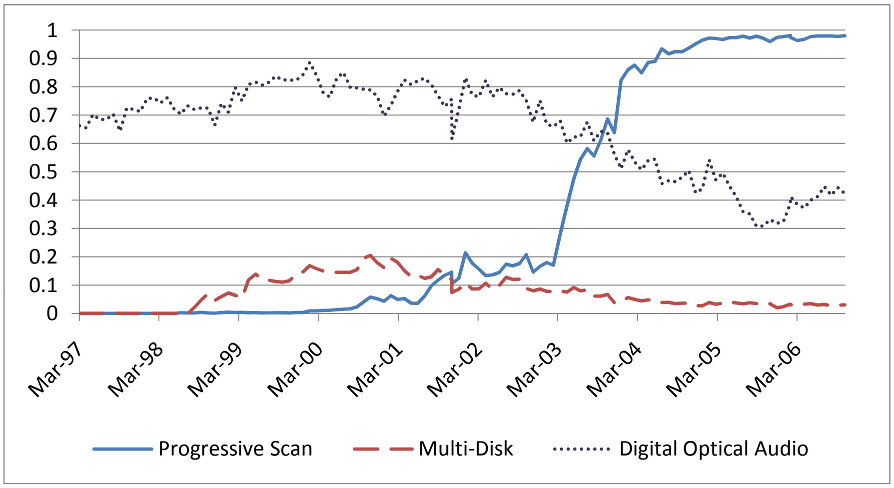

Table 2: Weighted average of three characteristic variables over time

over the early part of the product diffusion, but probably leveled off in the minds of most consumers by the end.

The fit of the model appears good in several dimensions. First, we graph the predicted ownership distribution over time. Figure 8 graphs the share of consumers with 1, 2, 3 or 4 DVD players as a share of consumers owning any DVD players over time, as predicted by the model. This evolution appears natural, and matches Figure 6 well.

Another interesting issue to evaluate is the relationship of the asymptote that results from 7 to the realizations of δ_{it}. We graph these two elements for 6 agents in Figure 9. To pick agents, we order agents by their willingness to purchase and graph numbers 1, 10, 20, 30, 40 and 48. In each case, the agent's δ_{it} converges almost exactly with the asymptote that corresponds to the agent's transition matrix. Thus, by the end of the sample, agents perceive that the DVD player market has reached a steady state and they do not expect future improvement. This result distinguishes the DVD market from the camcorder market, in which Gowrisankaran & Rysman (2012) find that the asymptotes are substantially above the realizations of δ_{it}, in a market that appears to be still improving by the end of the sample.

Now we turn to our estimation of the second stage. First, we remove seasonality from each series that we refer to below using the X11 procedure developed by the U.S. Bureau of the Census (as implemented by the PROC X11 command in SAS). A puzzle for us is that the time dummy coefficient exhibits a reverse seasonality (i.e. lower values in December), although

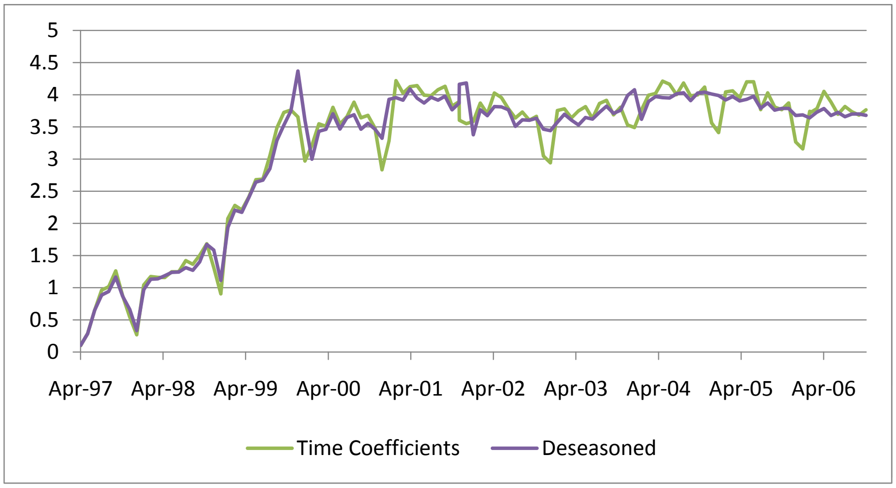

Figure 7: Time dummy coefficients

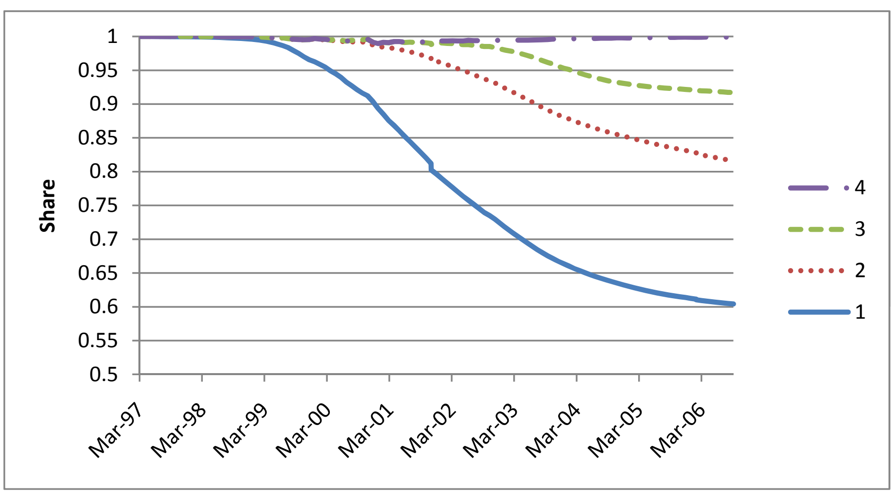

Figure 8: Share of DVD player holdings among households that hold at least one player, as predicted by the model.

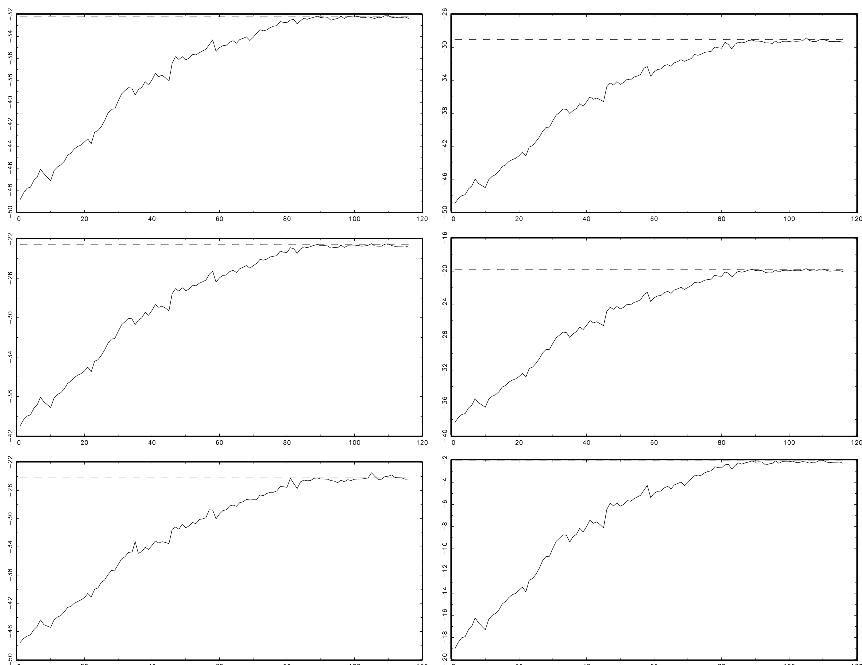

Figure 9: Asymptotes and δ_{it} for 6 representative consumers

the sales data were deseasoned before hand. This is perhaps due to the slight reverse seasonality in the price sequence (which were not deseasoned beforehand). Thus, we deseason the time dummy coefficients. That line is displayed in Figure 7.

Before moving to regressions, we run some tests on the structure of our time series data. First, we test whether the time coefficients can be characterized as a unit root process. We run the Augmented Dickey-Fuller test allowing for a drift term and a linear time trend (which has 114 observations). We find a test statistic $Z(t) = -2.252$, above the 10% critical value of -3.148. Thus, we fail to reject the null hypothesis of a unit root. Including up to 4 lags does not change this result. The Phillips-Perron test finds a similar result. A time series for which parameters change during the sample (a "structural break" in the language of this literature) can lead to a false acceptance of a unit root. However, we implement a Zivot-Andrews test for a unit root allowing for a break in both the constant and trend, where the procedure finds the break point, and we still fail to reject the null of a unit root. Running these tests on the early part of the data set (before October 2001, as we focus on below) finds similar result.

We perform a similar set of tests on our primary independent variable of interest, the number of new DVD titles associated with recent movies (that is, DVDs released within one year of an associated theatrical film release). Because we have a longer series on DVD titles, this procedure has 135 observations. In contrast to the time series coefficients, we can reject a unit root for this process. The Phillips-Perron test provides similar result.

Next, we estimate the structural VAR specified in Equation 11. For x_t, we use the number of new DVD titles associated with a recent theatrical release movie (a flow variable, not a stock). In Table 3, we report the results ignoring the endogeneity of the DVD variable. In column 1, we compute the parameters of Equation 11 via OLS using the entire time series data set, 114 observations. The variable of interest is the effect of the number of new DVD movie titles, but it turns out to be small and insignificant. Thinking about the industry and looking at the time coefficients we find (in Figure 7), it seems likely that network effects were more important early in the sample rather than later. In column 2, we confirm this hypothesis by including an interaction of the number of new DVD movie titles with time. In this case, we find that the effect of new DVD movie titles is larger and significant at the beginning, and declines in importance over time. Finally, we drop the interaction term and we try cutting the data at 4.5 years (54 observations). Estimating Equation 11 on the early data (column 3) finds similar results to the specification with the interaction term (column 2). While a coefficient

Constant	0.140	-0.077	-0.016
	(0.077)	(0.097)	(0.105)
Lag Y	0.913 *	0.782 *	0.742 *
	(0.031)	(0.049)	(0.086)
New DVD movie titles	0.010	0.035 *	0.032 *
	(0.005)	(0.009)	(0.011)
time	-0.001	0.010 *	0.006
	(0.001)	(0.004)	(0.007)
New titles X time		-0.0004 *	
		(0.00011)	
Cut-off date	none	none	Oct-01
Observations	114	114	54

Table 3: Second stage results: Time dummy coefficients as the dependent variable

of 0.032 seems small, we argue below that the cumulative effect of titles in the autorgressive process is in fact economically significant.

Recall that we can interpret the autoregressive term as a depreciation rate for the value of past releases in the current value of the titles market. We find a value of 0.782, which would be an enormous monthly rate in a capital investment context, but is probably reasonable in the movie market which emphasizes new releases. In addition, we test the error term in column 3 for autocorrelation (using the modified Durbin-Watson test in Stata) and a unit root (using the augmented Dicky-Fuller test) and reject them, facilitating our interpretation of the parameters. The unit root test implies that the variables are co-integrated, which addresses the spurious correlation problem (in the sense of time series econometrics). In practice, since we did not find that the titles variable exhibited a unit root, this is not a crucial issue for us, but we test for it for completeness.

A primary goal of the paper is to explore different ways to represent the titles market with data. Thus, we also experimented with letting x_t equal the number of new titles, including other genres than recent theatrical releases. In addition, we try different lag structures, for instance lettering the time coefficient depend on the first or second lag of x_t, rather than the contemporaneous x_t. We find that these alternative specifications fit the data worse in the sense of the sum of squared residuals. Although the difference is often not large, we prefer to define x_t as the number of new releases associated with recent theatrical films. We were initially interested in defining x_t with our data on the revenue from sales of DVD titles. However, these data only reach back to the year 2000. Since the role of network effects appears to be limited to the period before 2001, we cannot obtain results with the revenue data.

We are also concerned with the possible of endogeneity of x_t and y_t in Equation 11. We introduce the number of movies released to the box office five months previous, and the interaction of this variable with time. Results appear in Table 4. Column 1 reports the "first stage" regression of the number of titles on the exogenous variables. We see that the movie variable is positive but the interaction with time is negative. This is a surprising result: if only some movie titles are released on DVD early in the sample but all are later, we would expect the interaction with time to be positive. One explanation for our result may be provided by Chiou (2008). She shows that the five-month delay between movie and DVD release frays over time, as film producers move to distribute movies earlier. While a five month lag may still be the mode of the distribution, the average is falling over time, which may explain our results.

The main results appear in Column 2 of Table 4. We provide results only for the early part of the sample, before October 2001. The parameter on the number of new DVD movie titles is positive and significant. It is actually substantially larger than the parameter estimated in Table 3. While one might argue that the parameter should be smaller without the endogeneity issues, the standard error is much larger than before, and it would be difficult to reject many plausible outcomes.

One striking feature of this data is the rapid rise in the number of products, exhibited in Figure 1. This phenomena can also create problems for estimation. Ackerberg & Rysman (2005) argue that discrete choice demand systems make restrictive assumptions on unobserved heterogeneity (the distribution of ε_{ijt}) as choice sets expand, and one solution they recommend is including the log of the number of products as an explanatory variable in the structural model. That effect is not identified in a structural model with time dummies, but we can include it at this stage. In unreported results, we introduce the log of the number of products as an explanatory variable in Equation 11. We find that the parameter is large and negative, suggesting substantial crowding. However, the other results remain the same, and we do not pursue this issue further here.

While the parameters on DVD titles in Tables 3 and 4 are small numbers, we argue that they are economically important. In order to do so, we ask how much a consumer values the titles market, and how different that would be

	New DVD Movie Titles		Time Coefficients	
Constant	-3.435		-0.506	
	(5.021)		(0.349)	
Lag Time Coefficients	1.516		0.600	*
	(0.968)		(0.143)	
New DVD movie titles			0.109	*
			(0.054)	
time	0.478		-0.013	
	(0.169)		(0.016)	
New Movies (5 months ago)	0.380			#
	(0.200)			
New Movies X Time	-0.009	*		
	(0.005)			
Cut-off date	Oct-01		Oct-01	
Observations	54		54	

Table 4: Second stage results: Time dummy coefficients as the dependent variable

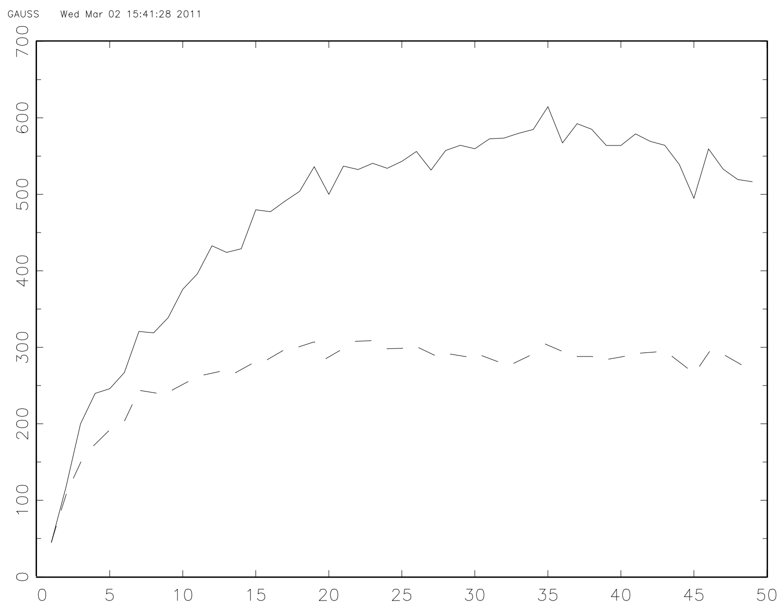

Figure 10: Welfare from the titles market

if no movie titles were introduced. We compute the dollar value by dividing the time coefficient by the derivative of utility with respect to price, which is α_p/p_t, where α_p is the average of the price coefficient. We use the sales-weighted average price in period t for p_t. Thus, the welfare could fall over time even if the time coefficients remain constant since price is falling, and price ultimately enters in the numerator.

The result appears in Figure 10. The solid line reports the result we see in the data. The effect starts close to zero and climbs to over \$500. The dashed line reports the result with $\beta_2 = 0$, so the effect of titles is shut down. The effect is large, with the welfare leveling off earlier and ending at less than \$300. One might imagine that without new titles, the value should be zero. However, there are other uses for DVD players besides recent theatrical releases, in particular other types of movies. Also, our computation is not a perfect measure of what would happen if we eliminated the titles market since we do not adjust expectations in this exercise. However, the result does suggest that titles are an important determinant of welfare.

7 Conclusion

This paper proposes methods for estimating a network effect in a dynamic environment. We address a series econometric issues that have not been well-documented in the previous literature. Our preliminary results find an important network during the time period of our data.

References

Ackerberg, D. A. & Gowrisankaran, G. (2006). Quantifying equilibrium network externalities in the ach banking industry. *RAND Journal of Economics*, *37*, 738–761.

Ackerberg, D. A. & Rysman, M. (2005). Unobservable product differentiation in discrete choice models: Estimating price elasticities and welfare effects. *RAND Journal of Economics*, *36*, 771–788.

Angrist, J. D. & Pischke, J.-S. (2009). *Mostly Harmless Econometrics*. Princeton University Press.

Berry, S. (1994). Estimating discrete choice models of product differentiation. *RAND Journal of Economics*, *25*, 242–262.

Berry, S., Levinsohn, J., & Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica*, *63*, 841–890.

Berry, S., Levinsohn, J., & Pakes, A. (2004). Estimating differentiated product demand systems from a combination of micro and macro data: The market for new vehicles. *Journal of Political Economy*, *112*, 68–105.

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

Chiou, L. (2008). The timing of movie releases: Evidence from the home video industry. *International Journal of Industrial Organization*, *26*, 1059–1073.

Clements, M. & Ohashi, H. (2005). Indirect network effects and the product cycle: Video games in the U.S., 1994-2002. *Journal of Industrial Economics*, *53*, 515–542.

Corts, K. S. & Lederman, M. (2009). Software exclusivity and the scope of indirect network effects in the u.s. home video game market. *International Journal of Industrial Organization*, *27*, 121–136.

Derdenger, T. (2013). Technological tying and the intensity of competition: Empirical analysis of the video game industry. Unpublished Manuscript, Tepper School of Business, Carnegie Mellon University.

Donald, S. G. & Lang, K. (2007). Inference with difference-in-differences and other panel data. *Review of Economics and Statistics*, *89*, 221–233.

Dube, J.-P. H., Hitsch, G. J., & Chintagunta, P. K. (2010). Tipping and concentration in markets with indirect network effects. *Marketing Science*, *29*, 216–249.

Fan, Y. (2013). Ownership consolidation and product characteristics: A study of the U.S. daily newspaper market. *American Economic Review*, *103*, 1598–1628.

Farrell, J. & Klemperer, P. (2007). Coordination and lock-in: Competition with switching costs and network effects. In M. Armstrong & R. Porter (Eds.), *Handbook of Industrial Organization, Vol. 3*. Elsevier.

Farrell, J. & Saloner, G. (1986). Installed base and compatibility: Innovation, product preannouncements, and predation. *American Economic Review*, *76*, 940–955.

Gandal, N., Kende, M., & Rob, R. (2000). The dynamics of technological adoption in hardware/software systems: The case of compact disc players. *RAND Journal of Economics*, *31*, 43–61.

Gowrisankaran, G. & Rysman, M. (2012). Dynamics of consumer demand for new durable goods. *Journal of Political Economy*, *120*, 1173–1219.

Gowrisankaran, G. & Rysman, M. (2013). A framework for modeling industry evolution in dynamic demand models. Unpublished Manuscript, Boston University.

Gowrisankaran, G. & Stavins, J. (2004). Network externalities and technology adoption: Lessons from electronic payments. *RAND Journal of Economics*, *35*, 260–276.

Hendel, I. & Nevo, A. (2006). Measuring the implications of sales and consumer stockpiling behavior. *Econometrica*, *74*, 1637–1673.

Inceoglu, F. & Park, M. (2004). Diffusion of new products under network effects. Unpublished manuscript, Sabanci University, Turkey.

Jeziorski, P. (2013). Effects of mergers in two-sided markets: Examination of the U.S. radio industry. Unpublished Manuscript, Haas School of Business.

Lee, R. (2013). Vertical integration and exclusivity in platform and two-sided markets. *American Economic Review*, *103*, 2960–3000.

Magnac, T. & Thesmar, D. (2002). Identifying dynamic discrete decision processes. *Econometrica*, *70*, 801–816.

Melnikov, O. (2001). Demand for differentiated products: The case of the U.S. computer printer market. Unpublished manuscript, Cornell University.

Moulton, B. R. (1990). An illustration of a pitfall in estimating the effects of aggregate variables on micro unit. *Review of Economics and Statistics*, *72*, 334–338.

Nair, H., Chintagunta, P., & Dube, J.-P. (2004). Empirical analysis of indirect network effects in the market for personal digital assistants. *Quarterly Marketing and Economics*, *2*, 23–58.

Nevo, A. (2000). A practitioner's guide to estimation of random coefficients logit models of demand. *Journal of Economics & Management Strategy*, *9*, 513–548.

Ohashi, H. (2003). The role of network effects in the U.S. VCR market, 1978-86. *Journal of Economics and Management Strategy*, *12*, 447–494.

Park, M. (2008). Estimation of dynamic demand with heterogeneous consumers under network effects. *Korean Journal of Industrial Organization*, *16*, 1–38.

Park, S. (2004). Quantitative analysis of network externalities in competing technologies: The VCR case. *Review of Economics and Statistics*, *86*, 937–945.

Petrin, A. (2002). Quantifying the benefits of new products: The case of the minivan. *Journal of Political Economy*, *110*, 705–729.

Rosse, J. N. (1970). Estimating cost function parameters without using cost data: Illustrated methodology. *Econometrica*, *38*, 256–275.

Rust, J. (1987). Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher. *Econometrica*, *55*, 999–1033.

Rysman, M. (2004). Competition between networks: A study of the market for Yellow Pages. *Review of Economic Studies*, *71*(2), 483–512.

Rysman, M. (2007). Empirical analysis of payment card usage. *Journal of Industrial Economics*, *60*, 1–36.

Saloner, G. & Shepard, A. (1995). Adoption of technologies with network effects: An empirical examination of the adoption of Automated Teller Machines. *RAND Journal of Economics*, *26*, 479–501.

Schiraldi, P. (2011). Autobmobile replacement: A dynamic structural approach. *RAND Journal of Economics*, *42*, 266–291.

Tauchen, G. (1986). Finite state markov chain approximations to univariate and vector autoregressions. *Economic Letters*, *20*, 177–181.

Appendix

This section derives our simpler Bellman equation which, in particular, does not depend on f_0. Use the assumption that $u(F_0) = f_0 - \psi(n)$ to rewrite the original Bellman equation (Equation 1) as:

$$\widetilde{E}V(f_0, n, \Omega) = E\,[\max$$

$$\max_{j=1,\dots,J} f_j + f_0 - \psi(n+1) + (n+1)g(N,\theta) - p_j + \varepsilon_j + \beta E\left[\widetilde{E}V\left(f_j + f_0, n, \Omega'\right)|f_0 + f_j, n, \Omega\right],$$

$$f_0 - \psi(n) + ng(N,\theta) + \varepsilon_0 + \beta E\left[\widetilde{E}V\left(f_0, n, \Omega'\right)|f_0, n, \Omega\right]]$$

where the first expectation is over ε. Use the notation for $EV\,(n, \Omega)$ from Equation 2:

$$EV\,(n, \Omega) = \widetilde{EV}\,(f_0, n, \Omega) - \sum_{\tau=0}^{\infty} \beta^{\tau}\,(f_0 + nE[g(N_\tau, \theta)|\Omega])$$

Plug this equation into the Bellman equation on both sides:

$$EV\,(n, \Omega) + \sum_{\tau=0}^{\infty} \beta^{\tau}\,(f_0 + nE[g(N_\tau, \theta)|\Omega]) = E\,[\max$$

$$\max_{j=1,\dots,J} f_j + f_0 - \psi(n+1) + (n+1)g(N,\theta) - p_j + \varepsilon_j +$$

$$\beta E\left[EV\,(n, \Omega) + \sum_{\tau=0}^{\infty} \beta^{\tau}\,(f_j + f_0 + (n+1)E[g(N_\tau, \theta)|\Omega])\,|f_0 + f_j, n, \Omega\right],$$

$$f_0 - \psi(n) + ng(N,\theta) + \varepsilon_0 + \beta E\left[EV\,(n, \Omega) + \sum_{\tau=0}^{\infty} \beta^{\tau}\,(f_0 + nE[g(N_\tau, \theta)|\Omega])\,|f_0, n, \Omega\right]]$$

We can take the term $\sum_{\tau=0}^{\infty} \beta^{\tau} f_0 + nE[g(N_\tau, \theta)|\Omega]$ out of the continuation value (in which case, we should start τ at 1 to account for discounting). We combine this term with elements in the flow utility (so we again start τ at 0). We arrange them as follows:

$$EV\,(n, \Omega) + \sum_{\tau=0}^{\infty} \beta^{\tau}\,(f_0 + nE[g(N_\tau, \theta)|\Omega]) = E\,[\max$$

$$\max_{j=1,\ldots,J} \sum_{\tau=0}^{\infty} \beta^{\tau} \left(f_j + f_0 + (n+1)E[g(N_\tau,\theta)|\Omega]\right) - \psi(n+1) - p_j + \varepsilon_j + \beta E\left[EV\left(n,\Omega\right)|f_0 + f_j, n, \Omega\right],$$

$$\left.\sum_{\tau=0}^{\infty} \beta^{\tau} \left(f_0 + ng(N_\tau,\theta)\right) - \psi(n) + \varepsilon_0 + \beta E\left[EV\left(n,\Omega\right)|f_0, n, \Omega\right]\right]$$

Now subtract the term $\sum_{\tau=0}^{\infty} \beta^{\tau} \left(f_0 + nE[g(N_\tau,\theta)|\Omega]\right)$ from both sides. Also, eliminate the condition on f_0 in the expectation of the value function in the continuation value, since f_0 no longer appears. In addition, substitute v_j for $\sum_{\tau=0}^{\infty} \beta^{\tau} \left(f_j + \theta E\left[g(N_\tau,\theta)|\Omega\right]\right) - p_j$ as in Equation 3. We fine the new Bellman equation:

$$EV\left(n,\Omega\right) = E\left[\max \max_{j=1,\ldots,J} v_j - \psi(n+1) + \varepsilon_j + \beta E\left[EV\left(n+1,\Omega'\right)|n,\Omega\right],\right.$$

$$\left.-\psi(n) + \varepsilon_0 + \beta E\left[EV\left(n,\Omega'\right)|n,\Omega\right]\right]$$

Thus, we see that $EV(n,\Omega)$ satisfies a Bellman equation. The choice problem described by $EV\left(n,\Omega\right)$ generates the same sequence of choices as $\widetilde{EV}(f_0,n,\Omega)$, and f_0 drops out since it affects the value of each choice in the same way.