Indirect Network Effects and the Quality Dimension: A Look at the Gaming Industry

Jin-Hyuk Kim, Jeffrey Prince, and Calvin Qiu[1]

May, 2013

Abstract

We present theoretical and empirical analyses of indirect network effects for a hardware market with vertically differentiated complementary goods. We demonstrate that the heretofore typical use of aggregate software counts can mis-measure the presence and/or magnitude of indirect network effects. We show this is true when there is correlation along the quality dimension between the marginal utility of software and either 1) the response of software supply to an increase in installed base, or 2) conditional variation in software availability. We illustrate this idea using a simple monopolistic competition model, and through empirical analysis of the 7th-generation console market.

Keywords: indirect network effects, vertical differentiation, video game industry

JEL Classification: L14, L82

[1] Jin-Hyuk Kim is at the Department of Economics, University of Colorado at Boulder. Jeffrey Prince is at Business Economics and Public Policy in the Kelley School at Indiana University. Calvin Qui is at the Faculty of Economics, University of Cambridge. They can be reached at jinhyuk.kim@colorado.edu, jeffprin@indiana.edu, and yq211@cam.ac.uk, respectively.

Electronic copy available at: http://ssrn.com/abstract=2194662

1. Introduction

Economists have long recognized that the benefit from consuming a particular good often depends on the number of consumers who purchase the same good. A common manifestation of such a benefit is through indirect network effects, where a larger base of adopters of a primary product (i.e., "hardware") creates a larger market for complementary goods (i.e., "software"), which then increases the value of the primary good.[2] Such effects can play a major role in the demand for, and profitability of, a product when they exist, and can even help determine its survival (e.g., HD DVD vs. Blu-ray). Consequently, the ability to identify and accurately measure indirect network effects for products in a market is important when assessing that market's current and long-run structure.

A major branch of recent research in this area uses counts of complementary goods to measure the impact of indirect network effects in the gaming industry (see, e.g., Clements and Ohashi, 2005; Corts and Lederman, 2009; and Dubé et al., 2010). In contrast, anecdotal evidence suggests that the effects of software offerings are heterogeneous across quality levels. For instance, Sony's PlayStation dominated the 5th- and the 6th-generation video game console markets, even though Sony was an entrant in the mid-1990s. One reason often quoted for Sony's success is that it had developed a large number of 'killer' applications. In the 7th-generation console market, Nintendo's Wii has dominated Microsoft's Xbox 360 and Sony's PlayStation 3, and critics say that the motion-sensor games played on Wii were the 'game-changers' and an important reason for Wii's success (Gaudiosi, 2007).

The purpose of this paper is to re-examine the indirect network effects in the home video game market, allowing the two sides of the market to have differential effects along the quality dimension. We believe that this provides an additional check on the network effects theory while also providing updated evidence on the existence and magnitude of these effects. To be specific, hardware sales may only increase low-quality software offerings, and it may only be high-quality software offerings that impact hardware sales. Therefore, the multiplicative effect, or positive feedback loop—often used in the literature as measurement of network effects—is not

[2] In an influential paper, Katz and Shapiro (1995) draw a distinction between direct and indirect network effects. Direct network effect refers to "a direct physical effect of the number of purchasers on the quality of the product" whereas indirect network effect is characterized by the lack of such direct physical effect. They mention that indirect network effects induce positive feedback between primary and complementary markets in self-fulfilling equilibria.

Electronic copy available at: http://ssrn.com/abstract=2194662

straightforward when the quality dimension is explicitly taken in account.

Our analysis begins with a straightforward extension of the static monopolistic competition model used by Chou and Shy (1990) and Church and Gandal (1992, 1993) to include vertically differentiated complementary goods. We then derive the marginal effect of installed base on the supply of games as well as the marginal utility of software availability, where the indirect network effect is defined as the product of these two effects across the quality dimension. The static model we analyze serves as a relatively simple template to illustrate the importance of accounting for vertically differentiated software in identifying indirect network effects. However, the basic insights we highlight are not dependent on a static assumption, and thus can apply to dynamic models as well (e.g., Derdenger 2012, Lee 2013)[3].

We detail market and data conditions that will generate differing measures of indirect network effects when allowing for quality-differentiated software versus aggregate counts. Specifically, we show that the use of aggregate software counts can mis-measure indirect network effects when there is correlation along the quality dimension between the marginal utility of software, and either 1) the response of software supply to an increase in installed base, or 2) conditional variation in software availability in the data being utilized. Using aggregate counts will overestimate indirect network effects when the former correlation is negative, and vice versa; and using aggregate counts will underestimate indirect network effects when the latter is negative, and vice versa.

In our data set, we show that both of the above correlations exist and both turn out to be negative. However, the direction of mis-measurement using aggregate counts is dominated by the latter. In particular, we find that, in the 7th-generation console market, the use of aggregate measures underestimates the indirect network effects by approximately 30 percent. In terms of the implied percentage increase in market share of a console, we find that using aggregate measures underestimates the predicted increase in market share over a two-year period by roughly one third. We note that our methodology is general and can be applied to other two-sided markets. The direction and magnitude of these biases would obviously be specific to the data set used by researchers.

In addition to these primary findings, our results indicating a larger effect for high-quality

[3] The econometric model we ultimately estimate can also serve as a complement to these more structural models. We discuss this further below.

software on hardware sales serve to complement dynamic structural papers analyzing this industry, e.g., Lee (2013) and Derdenger (2012). Lee (2013)'s hardware model uses a single variable representing the software side, the expected utility of software, which is derived from a structural model of game purchases. In the model, high game sales lead to inference of high game quality, and high game quality affects expected utility essentially by construction. In contrast, this paper directly analyzes the relationship between hardware sales and game quality. In this sense, our estimation results could be seen as a test of whether the specification in these structural models is reasonable, and our results indeed support this.

We are not the first to consider the effects of quality-differentiated complementary products in the home video game industry. Corts and Lederman (2009) define three levels of hit games based on dollar sales during the first 12 months after release and use them in their hardware demand estimation. However, they allow for this differentiation just to check the sensitivity of their estimates to these alternative software measures, so that only one measure is used at a time. The difference is that in our analysis, we focus on the differential effects of software groups based on game ratings and we also separately identify the effects of increased installed base on vertically differentiated software groups in order to find a true estimate of the indirect network effects.[4]

There are a number of other studies that examine indirect network effects. For instance, Gandal et al. (2000), Nair et al. (2004), and Clements and Ohashi (2005) all measure the software variety by counting the total complements available, where the focus is mostly on the tradeoff between lowering hardware price and increasing software supply. Prieger and Hu (2012) consider exclusive contracts between the platform and software developers and find that there is little evidence that exclusivity contributes to hardware demand. Basu et al. (2003) and Gretz (2010) show that there are also significant interactions between software availability and hardware attributes.

Finally, another related literature is the literature on platform competition (e.g., Rysman, 2004; Ackerberg and Gowrisankaran, 2006; and Zhu and Iansiti, 2012), where much of the focus

[4] Binken and Stremersch (2009) use information on game quality ratings to identify 'superstar' games and estimate their impact on hardware sales. They have not, however, considered the full spectrum of game groups or the software provision side. Hogendorn and Yuen (2009) show theoretically that the platform that has access to 'must-have' components would experience higher sales.

is on structurally estimating the effect of platform pricing, entry and exit, and exclusive contracts. The paper in this literature closest to ours is Rysman (2004) where the amount that consumers use a directory increases in the directory's level of advertising, and the retailer's willingness-to-pay for advertising increases in the amount of usage. Our basic framework is similar except in our model the two sides of the market interact through quality-differentiated channels. Also, we endogenize the number of complementary good producers.

The outline for the paper is as follows. Section 2 sets forth our theoretical model and its predictions. Section 3 discusses data and measurement issues. Section 4 contains the empirical methodology, and Section 5 presents the main results. Section 6 provides brief conclusions.

2. Theoretical Model

In this section we extend the monopolistic competition model used by Chou and Shy (1990) and Church and Gandal (1992, 1993) to explicitly incorporate quality-differentiated complementary goods, and given the hump-shaped distribution of game titles derive predictions regarding the consumer's marginal utility with respect to each game group as well as the supply responses from a unit change in installed base. In our model, we abstract from the hardware firm's profit maximization problem in order to highlight the effects we expect to find and use as a guide for our formal estimation.

Consider a primary goods market (i.e., consoles) with L types of complementary products (i.e., game titles), where $l = 1, \dots, L$ denotes the quality of games in an ascending order. For instance, using a continuous quality measure, we can sort game titles into deciles, where L represents the highest decile group. In each quality group, game developers may produce a number of different varieties. The console hardware provides no stand-alone benefit. Following Dixit and Stiglitz (1977), we model consumer preferences over software variety by adopting a CES utility function,

$$(1) \quad u = \mu_1 log\left[\sum_{i=1}^{N_1} y_1(i)^{\alpha_1}\right]^{\frac{1}{\alpha_1}} + \cdots + \mu_L log\left[\sum_{i=1}^{N_L} y_L(i)^{\alpha_L}\right]^{\frac{1}{\alpha_L}},$$

where $y_l(i)$ is consumption of a game i of type l; N_l is the number of game varieties of type l; the parameter μ_l, where $\mu_1 + \cdots + \mu_L = 1$, gives the share of expenditure that a representative consumer devotes to type l games; and the parameter $\alpha_l \in (0,1)$ measures the degree of

horizontal product differentiation among type l products, where a greater α_l means more substitutability or less differentiation. There is a unit measure of consumers who differ only in their outside option (non-consumption) value, and we normalize the consumer's expenditure on games to one.

It is well known that the above preferences yield demand functions faced by the producer of variety i of the form,

$$(2) \quad y_l(i) = A_l p_l(i)^{-\frac{1}{1-\alpha_l}},$$

where $p_l(i)$ is the price of good i of type l, and A_l is the demand level,

$$(3) \quad A_l = \frac{\mu_l E}{\sum_{i=1}^{N_l} p_l(i)^{-\frac{\alpha_l}{1-\alpha_l}}},$$

where E denotes the size of inside consumers (i.e., installed base), each of whom spends one unit of income on software. Notice that in a monopolistic competition model the producer of variety i of type l treats A_l as a constant.

The production of any variety of type l requires a fixed development cost, $k_l > 0$, plus a constant variable cost, $c > 0$, per unit. It is assumed that the development costs increase with the product quality, so that $k_1 < \cdots < k_L$. Without serious loss of generality the variable costs are the same across types. Developers are required to enter into licensing agreements with console manufacturers, and then submit finished games to console manufacturers for approval. Hence, a game title is of no use unless it is approved by the console manufacturer. Since complementary goods are often tailored to the primary good (due to, e.g., different languages and capacities), once the development cost is sunk, this creates a potential holdup problem.

Following Grossman and Helpman (2002), we assume that, once the development costs are sunk, the game developer and the console manufacturer bargain over the split of the joint surplus.[5] To be more precise, a developer of type l game will get a share $\omega_l \in (0,1)$ of the potential sales revenue, $p_l(i)y_l(i)$, where ω_l is exogenous. Thus, the complementary share

[5] In the video game market, the console manufacturer may make deals with developers specifying a different proportion of sales revenue to be paid to manufacturers as royalties for licensing. Special rates may be given to popular developers.

$(1-\omega_l)$ that accrues to the console manufacturer represents the per-unit royalty rate for type l games. The game developers foresee potential revenue of $\omega_l p_l(i) y_l(i)$ from producing $y_l(i)$ units, and the profit maximization leads to the following mark-up pricing,

$$(4) \quad p_l(i) = \frac{c}{\omega_l \alpha_l}.$$

Hence, for each type l the price of all games is the same in a symmetric equilibrium. Free entry ensures zero expected profits for game developers of each type that enter the market, where the profits are given by

$$(5) \quad \pi_l(i) = (\omega_l p_l(i) - c) y_l(i) - k_l$$

$$= (1-\alpha_l)\omega_l \frac{\mu_l E}{N_l} - k_l \,,$$

where the second line uses (2), (3), and (4). Therefore, the equilibrium variety of type l games is determined by setting $\pi_l(i)$ equal to zero:

$$(6) \quad N_1^* = \frac{(1-\alpha_1)\omega_1 \mu_1 E}{k_1}, \ldots, N_L^* = \frac{(1-\alpha_L)\omega_L \mu_L E}{k_L}.$$

The model can explain the observed, hump-shaped pattern of software distribution across quality dimension. Notice that the development costs, k_l's, are increasing in the quality of games; however, the other parameter values are of importance as well. The degree of product differentiation is arguably higher for higher quality games than for lower quality games meaning $\alpha_1 > \cdots > \alpha_L$.[6] Also, the share of consumption expenditure may be increasing in the quality dimension, which implies $\mu_1 < \cdots < \mu_L$. The licensing fee payable includes two components: the cost of disc production, and a royalty fee. It depends on a combination of projected game pricing and sales revenue, and is negotiable. Some anecdotal evidence suggests that it is generally around 15 percent across the board.[7]

[6] With regard to game pricing in (4), this assumption implies that higher quality games are more expensive that lower quality games.

[7] New games are typically priced at $60, where the publishers pay a $6-10 licensing fee to console manufacturers. (See, e.g., http://latimesblogs.latimes.com/entertainmentnewsbuzz/2010/02/anatomy-of-a-60-dollar-video-game.html).

Notice that holding constant other parameter values a larger development cost, k_l, tends to decrease the number of software variety. This can reconcile the fact that in our dataset the number of top-quality games (e.g., ratings in the range 85+) is smaller than the number of middle-quality games (e.g., ratings in the range 60-85). On the other hand, the assumption of decreasing α_l as well as increasing μ_l tends to increase the number of higher quality games. This can explain the pattern in our dataset that the number of titles is increasing in quality except at the very top levels, where the effect of development costs seems to dominate smaller variations in other parameter values.[8] Thus, the model can trivially match the observed, hump-shaped distribution of game titles, which we take as given.

In our formal empirical analysis we are interested in estimating the software supply response to changes in installed base (i.e., dN_l^*/dE), and the marginal utility from an increase in software variety (i.e., dU/dN_l) across the quality dimension. If both effects are significant and positive, then it means that there is a positive feedback loop between installed base and software availability, and the indirect network effects would be the product of these two across all groups. We analyze these two components of indirect network effects in turn.

First, from (6) it follows that N_l^* is a simple linear function of the size of installed base, E. Thus, the impact of a one unit change in installed base on the number of software variety of type l is proportional to the number of software titles themselves. Consequently, if the distribution of game titles is hump-shaped across the quality dimension, then we expect that dN_l^*/dE has also a hump-shaped pattern in our estimation, as the effect of development costs k_l dominates at the high end of the distribution.

Second, let U denote the indirect utility of a representative consumer in the symmetric equilibrium. Then, the marginal utility with respect to software variety can be written as follows:

$$(7) \quad \frac{dU}{dN_l} = \frac{\mu_l}{\alpha_l} log((N_l^* + 1)y_l(i)^{\alpha_l}) - \frac{\mu_l}{\alpha_l} log(N_l^* y_l(i)^{\alpha_l})$$

$$= \frac{\mu_l}{\alpha_l} log \frac{N_l^*+1}{N_l^*},$$

[8] The average development costs for a 7th-generation console game is around $15 million with high-end games often costing more than $50 million. (See, e.g., http://www.notenoughshaders.com/2012/07/02/the-rise-of-costs-the-fall-of-gaming/).

holding constant the demand level A_l that game developers take as given.

Notice that the term in the logarithm is decreasing in N_l^*. Further, the assumptions that $\alpha_1 > \cdots > \alpha_L$ and $\mu_1 < \cdots < \mu_L$ imply that dU/dN_l should be increasing in quality if the number of varieties in each group l were the same. Given that the distribution of N_l is hump-shaped, these conditions imply the following: Because the number of high-quality game titles is smaller than that of middle-quality titles, the marginal utility of software variety must be higher for the top-quality games than that for middle-quality games. On the other hand, because the number of middle quality game titles is larger than that of low-quality titles, how the marginal utilities compare between low- and middle-quality games is ambiguous because the relative sizes of software varieties and the assumptions on α_l's and μ_l's have offsetting effects.

From the above discussions, we can derive the following testable hypotheses:

Prediction 1: Supply response from a one unit change in installed base will be the largest in middle-quality groups relative to low- or top-quality groups.

Prediction 2: Marginal utility of a change in software variety will be the highest from top-quality games relative to middle- or low-quality games.

A couple of points on these predictions are worth mentioning. First, these predictions are essentially comparative statics results. As we noted in the Introduction, this approach assumes that consumers have short-term adaptive expectation rather than perfect foresight. To be more precise, the utility gain from additional software availability in a period would increase the consumption value of a console and thus the number of inside consumers in a static equilibrium. Once the size of installed base increases, a new equilibrium will be reached in the next period with a larger number of software varieties (entries) to satisfy the zero expected profit condition in equilibrium. That is, the positive feedback loop in our model is not due to the far-sighted rational expectations on the consumers, but rather resembles the traditional tatonnement process. As we mentioned earlier, the two expectation assumptions form complementary approaches.

Second, while we admit that our model abstracts from a number of potential complications, we do not think our predictions are significantly affected by them. For instance, while a game can be developed in-house (i.e., when the software developer is employed by the console

manufacturer), this would not distort the supply incentives as long as there is no vertical exclusion. Also, we did not distinguish console-level software supply from overall supply. In the former case, a developed game may only need to be ported to another platform to increase supply whereas in the latter, an increase in supply means a game changes from not existing to being developed. We believe that as long these multi-homing do not disproportionately occur for high-quality games, our main predictions would remain unchanged.[9] We support these claims by including separate estimation results for exclusive and non-exclusive titles in Section 5.

3. Data

In our empirical analysis, we focus on the 7th-generation home video game console market in the United States. The seventh generation began on November 22, 2005 with the release of Microsoft's Xbox 360 and continued with the release of Sony's PlayStation 3 on November 11, 2006 and Nintendo's Wii on November 19, 2006. We hand-collected weekly console sales data from Vgchartz (www.vgchartz.com) and game title availability for each console from Vgchartz and a number of gaming publications, including Gamespot and IGN. The aggregate number of game titles in our dataset (as of the end of April 2012) is 879 for Xbox 360, 898 for Wii, and 723 for PlayStation 3.

We also collected information about console technical specifications and model availability from multiple online sources.[10] The quality metrics for each game title came from Metacritic. Metacritic gives an aggregate score to the majority of games published in the U.S. To do so, it weighs review scores from a large number of game publications.[11] For our purposes, a game's Metacritic score is taken to be indicative of its quality. Using the scores on Metacritic (ranging 0-100), we collected the number of titles in each 5-point range (e.g., 80-84, 85-89, etc.). We then constructed the following quality groupings: 90-99, 85-89, 80-84, 75-79, 70-74, 65-69,

[9] The percentage of exclusive titles is relatively low, except for the case of Wii. Exclusives compose 20 percent for Xbox 360 titles, 16 percent for PlayStation3 titles, and 49 percent for Wii titles.

[10] These includes a wide range of websites, too many to list here. However, references are available upon request.

[11] See http://www.metacritic.com/about-metascores for more details on Metascore. One might wonder whether these review scores are simply reflective of critics' tastes, but Metacritic aggregates at least three underlying reviews. Critics' review scores correlate with sales and are gaining more acceptance as an alternative measure of quality in the literature (e.g., Waldfogel, 2011).

60-64, 55-59, 50-54, and < 50.[12]

As for console prices, we constructed our console prices by using manufacturer's price-cut announcements, which are public announcements. Most of the existing studies use the data originated from the market research firm NPD. Both approaches have pros and cons. If using NPD data, then researchers have to construct the average retail price of a console by dividing total revenues by quantities sold. For our purposes, this creates some obvious endogeneity problem. On the other hand, console manufacturers engage in price discrimination by offering multiple versions. For each console, we used the price of the version considered to be the 'basic' model at the time as the console price.

The rationale for using the basic model in our approach is that the capability of a console mainly depends on its processor (where all calculations are performed) and short-term memory (or RAM, which holds all bits of data in place when performing calculations). These are the main components that determine the console power, and they are constant across all models. The premium models are mainly differentiated with larger hard drive sizes, and bundled games for collectors' versions. Therefore, the console model chosen when constructing our price data does not affect the number of current-generation games that can be played on the console.

We determined the size of the home video gaming market based on the number of households with a television in the United States. This information is publicly available each year from Nielsen, and we interpolated the yearly numbers across each week of each year. Using this information, we constructed the size of the gaming market in any given week as the difference in the number of households with a television in the U.S. minus the number of households that had purchased a 7th-generation gaming console. This construction implicitly assumes that households only purchase one gaming console from a given generation.

Finally, we used data on the number of competing consoles from Vgchartz, where we "exited" the 6th-generation consoles when no more sales data were available. We constructed market shares for each console each week. This is done in two ways: 1) calculating the market share of each console and the "no purchase" option, and 2) calculating the market share for each console among purchases. Both market share calculations will be relevant in our formal

[12] We aggregated data above a score of 90 and below 50, since these were rather obvious cutoffs for "very good" and "very bad" games. Our findings are robust to splitting these groups into smaller subgroups.

estimation described below. We provide summary statistics for all relevant variables in Table 1. As can be seen in the Table, the number of games at each quality level is hump-shaped.

[Table 1 about here]

4. Econometric Specification

To empirically test the hypotheses specified in Section 2 above, we must specify two econometric equations: one for the demand for consoles (hardware), and one for the supply of games (software).

4.1. Demand for Consoles

Our model of demand for consoles follows the discrete-choice literature (e.g., Berry, 1994; Berry et al., 1995; Nevo, 2000), along with the recent literature analyzing the gaming industry (e.g., Clements and Ohashi, 2005; Corts and Ledermand, 2009). We begin by specifying the utility for individual i from buying console j at time t as follows:

$$(8) \quad u_{ijt} = x_{jt}\beta + \alpha p_{jt} + \lambda_1 SW_{1jt} + \cdots + \lambda_L SW_{Ljt} + \xi_{jt} + \varepsilon_{ijt},$$

where x contains observable characteristics of the console, p is the price of the console, and SW_l is the number of software titles at quality level l. The two error terms represent console-level unobserved quality and an idiosyncratic shock to utility, respectively. Notice that this utility specification is a simplified version (with respect to software availability) of the utility specified in our model in Section 2, and follows the standard specification in the literature[13]. It also allows for simple calculations of marginal utilities, and comparisons to prior analyses. In Section 5, we will conduct a simple check of the robustness of our findings to this specification.

We assume that each consumer chooses the console that maximizes her utility (including

[13] Note that we assume that consumers have period-by-period adaptive expectation rather than perfect foresight into the future, as in, e.g., Lee (2013). While our static approach may fail to capture forward-looking behavior, dynamic models typically must make assumptions such as perfect foresight, which can be tenuous in industries with disruptive innovations, to circumvent the possibility of multiple equilibria.

the outside option of no purchase). Following Berry (1994), we define the mean valuation of a console across consumers to be $\delta_{jt} = x_{jt}\beta + \alpha p_{jt} + \lambda_1 SW_{1jt} + \cdots + \lambda_L SW_{Ljt} + \xi_{jt}$. As is standard, we normalize this mean valuation to zero for the outside option. Following Corts and Lederman (2009), we adopt a nested logit framework, where we group all consoles into one nest and the outside option into another. This allows for correlation in idiosyncratic tastes (ε_{ijt}) across consoles, allowing them to be closer substitutes with each other, relative to the outside option.

With these assumptions, Berry (1994) shows the following linear formulation using market shares can be derived:

$$(9) \quad ln(s_{jt}) - ln(s_{0t}) = x_{jt}\beta + \alpha p_{jt} + \lambda_1 SW_{1jt} + \cdots + \lambda_L SW_{Ljt} + \sigma\, ln(s_{j|g,t}) + \xi_{jt},$$

where s_{jt} is the market share of console j at time t, s_{0t} is the market share of the outside option at time t, and $s_{j|g,t}$ is the market share of console j within group g at time t (i.e., the market share of console j among all console purchases at time t).

In taking this model to the data, we must address two key issues. The first is choosing the components of x. There is very little variation in the primary console characteristics (e.g., RAM) over time for a given generation, so there is little hope in identifying their impact on utility. The net effect of such time-invariant characteristics at the console level is captured by our inclusion of console fixed effects. In addition, we include a highly flexible function of console age, varying at the console level. Specifically, we include a fifth-order polynomial in console age, allowing for differing parameters across consoles.[14] This set of controls allows for different lifetime trends in purchasing for each console, essentially mapping out an "expected diffusion process," ceteris paribus, for each console. Lastly, we include console-level indicator variables for the Holiday Season, which equal one during the period of November 20th to December 31st.

The second issue we must address is potential endogeneity of several of our independent variables. Two variables that pose clear endogeneity concerns are price and the inside market share. We address these concerns by using standard instrumental variables (IVs), as established

[14] We cannot include a higher-order polynomial without encountering collinearity issues; however, there is little difference in our results between a fourth-order and fifth-order polynomial. As discussed in Section 5, we also use console-year fixed effects, but these (expectedly) result in a poorer fit, and provide similar results.

in the literature (e.g., Bresnahan et al., 1997; Gowrisankaran and Stavins, 2004). In particular, we instrument using information about competing consoles: number of 6th- and 7th-generation competing consoles, sum of competitors' processor speed, sum of competitors' memory, sum of competitors' word length, number of competitors of same generation, and number of competing consoles from the same manufacturer (i.e., presence of 6th-generation console by the same manufacturer).[15] As these IVs vary based on entry and exit decisions by all other competing consoles, the identifying assumption is that these decisions are not correlated with a given console's unobserved quality level.[16]

Other variables that may pose an endogeneity concern include our measures of available games (software variety). However, as we discuss below in our software supply model, software supply decisions for a game console have a substantial amount of lead time (typically at least one year). Consequently, if a console experiences an unobserved quality shock (e.g., a good review in a gaming magazine), this shock must persist for a very long time in order to generate a correlation between the number of games supplied and unobserved quality. Further, our analysis focuses on differential impacts of software titles across the quality dimension, so even if there exists some level of positive bias in our estimates, this does not qualitatively affect our findings unless the bias substantially differs across quality levels. This is highly unlikely since the largest effects we find are for the highest-quality games, which also generally have the longest lead time before the game is released.

4.2. Supply of Games

Following Clements and Ohashi (2005), Corts and Lederman (2009), and Dubé et al. (2010), we specify a reduced-form relationship between software supplied for a console and that console's installed base. A key departure in our model is that we distinguish among supply of varying quality levels. Consequently, we specify a system of supply equations as follows:

$$(10) \quad SW_{1jt} = \alpha_{1j} + \gamma_1 IB_{jt-104} + \eta_{1jt}$$

[15] Word length refers to the size of basic unit of data used by consoles' microprocessors.
[16] Corts and Lederman (2009) provide an extensive discussion of this assumption. In short, they argue that actual console launch dates are essentially random (due to technological uncertainty), so unobserved quality is unlikely to be correlated with exit because exiting consoles are only indirect competitors, and any correlation between unobserved quality and entry is likely mitigated by console-age controls.

…

$$SW_{Ljt} = \alpha_{kj} + \gamma_L IB_{jt-104} + \eta_{Ljt}$$

Here, α_{lj} is a fixed effect for quality level l for platform j, and IB_{jt-104} is the size of installed base for platform j at time $t - 104$. To account for the time it takes to develop a game after the production decision has been made, we use the two-year lagged level of installed base, i.e., the installed base as of 104 weeks prior to the observed software supply. We also include console-year-month dummy variables to flexibly control for console-level technological changes in the software industry over time.

These equations may suffer from potential endogeneity concerns with respect to the installed base. To address this concern, we follow the literature by using the age of the platform (and age squared) as an instrument for its installed base.

4.3. Measuring Indirect Network Effects

Given the above empirical specifications, we can test whether the two components of indirect network effects are working through differing quality channels. Following the literature (e.g., Rysman, 2004), we measure indirect network effects as the sum, across the quality dimension, of the products of supply response and marginal utilities. We can compare an estimate of the indirect network effect using our vertically differentiated model to the one we would get using a model where all games are treated equally. The alternative model using simple counts would have the following hardware and software equations, respectively:

(11) $\ln(s_{jt}) - \ln(s_{0t}) = x_{jt}\beta + \alpha p_{jt} + \lambda SW_{jt} + \sigma \ln(s_{j|g,t}) + \xi_{jt}$

(12) $SW_{jt} = \alpha_j + \gamma IB_{jt-104} + \eta_{jt}$

Within this model, we can calculate a version of the indirect network effects as the impact of a one unit change in installed base (i.e., a change of one million adopters) on the mean utility of a console. This is captured by $\gamma * \lambda$. In contrast, using our vertically differentiated model, the impact of a one unit change in installed base is captured by $\sum_{i=1}^{L}(\gamma_i * \lambda_i)$.

Why might $\gamma * \lambda \neq \sum_{i=1}^{L}(\gamma_i * \lambda_i)$? To answer this, we begin by noting the implied relationships among these parameters. First, by construction, $\gamma = \sum_{i=1}^{L} \gamma_i$. That is, the marginal

effect of installed base on the aggregate total of games is simply the sum of the marginal effects of installed base on each mutually exclusive category of games. Next, notice that $\lambda = \sum_{i=1}^{L} w_i \lambda_i$, where w_i's are 'weights' that sum to one, and whose individual magnitudes depend on the relative partial variation of each game quality group in the data, i.e., after controlling for other factors, the variance of available games for each quality group within the dataset. In other words, the marginal utility of a change in the aggregate number of games is a weighted average of the marginal utilities of the game groups, where the weights represent each group's ability to contribute toward identification in the data set. Consequently, $\gamma * \lambda = (\sum_{i=1}^{L} \gamma_i)(\sum_{i=1}^{L} w_i \lambda_i)$ and we then ask why might $\gamma * \lambda = (\sum_{i=1}^{L} \gamma_i)(\sum_{i=1}^{L} w_i \lambda_i) \neq \sum_{i=1}^{L} (\gamma_i * \lambda_i)$.

Suppose that the weights are all equal (i.e., they all equal $1/L$). Then, comparing the two models' indirect network effect measures simply boils down to determining the correlation between γ_i and λ_i. If they are positively correlated, using aggregate measures will underestimate the true effect, and overestimate the true effect if they are negatively correlated. For instance, if high-quality games both generate a higher marginal utility for consoles and have a lower supply response to changes in the installed base, then we have negative correlation; hence the use of aggregate measures will overstate the true indirect network effect. We illustrate this idea in Figures 1 and 2 below. In Figure 1, there is a negative correlation, as described above, and in Figure 2, there is a positive correlation between marginal utility and supply response.

[Figures 1 and 2 about here]

Now, consider an alternative scenario where γ_i and λ_i are uncorrelated, but there is correlation between w_i and λ_i. In this case, we would have $(\sum_{i=1}^{L} \gamma_i)(1/L)(\sum_{i=1}^{L} \lambda_i) = \sum_{i=1}^{L} \gamma_i * \lambda_i$, but $(\sum_{i=1}^{L} w_i \lambda_i) \neq (1/L)\sum_{i=1}^{L} \lambda_i$. Hence, in contrast to above, if w_i and λ_i are positively correlated, using aggregate measures will overestimate the true effect, and underestimate the true effect if they are negatively correlated. For instance, if marginal utility is increasing in game quality, but partial variation in available games is declining in game quality, there is a negative correlation between λ_i and w_i, implying the use of aggregate measures will underestimate the true effect, ceteris paribus.

In sum, the relationship between measured indirect network effects using aggregate games

versus quality-differentiated games will depend on whether correlation along the quality dimension exists between the marginal utility of software (λ_i) and either 1) the response of software supply to an increase in installed base (γ_i), or 2) conditional variation in software availability (w_i).

5. Results

According to our hypotheses, we expect to observe the largest coefficient estimates in our hardware equation for the highest quality games. However, it is not clear that these games are also the most responsive to a console's installed base. In fact, our prediction in Section 2 was that the supply response from a unit change in installed base will be hump-shaped across the quality dimension.

5.1. Hardware Results

Table 2a contains our IV nested-logit results for console demands, using two different standard functional forms for utility with respect to software availability – linear and logarithmic. These results are reported in columns (1) and (3). Both sets of results tell us a similar story. Specifically, they show that the highest quality games tend to have the highest marginal utilities for consumers. In both specifications, we see particularly high coefficient estimates for titles ranked above 80. This finding is consistent with prior work (e.g., Binken and Stremersch, 2009), which indicates a disproportionate impact on utility from high-quality games.

[Table 2a about here]

In columns (2) and (4) of Table 2 we estimate our model using total counts of game titles, in both linear and logarithmic form. In both cases, we get the standard result of a positive effect on utility from an increase in total titles available. We will utilize these results for further comparison purposes when we discuss the implied indirect network effects in Section 5.3.

To test the robustness of these findings, we consider two alternative analyses. First, we split our set of game titles into those that are exclusive and those that are non-exclusive, to see if the pattern differs between these different types of games. The results are presented in the first four columns of Table 2b. Here we see a generally similar pattern for both types compared to our

baseline results that combine the two; namely, the highest-quality games have the largest impact on hardware demand, and that effect generally diminishes for lower-quality games. We also see that changes in aggregate counts have a similar impact on console demand across the two types of games.

Second, we allow for the possibility that games might become obsolete after enough time elapses since their release. In columns (5) and (6) of Table 2b, we present our results for the baseline model where we no longer count a title after it has reached an age of three years. We again observe similar results for the quality-differentiated regression in column (5). Further, our result in column (6) using the aggregate counts indicates that the indirect network effects for the aggregate titles essentially go to zero.[17] Thus, a significant mis-measurement could arise if one uses aggregate counts and even a reasonable cutoff. We observe similar qualitative results when imposing a cutoff of one year and two years although unsurprisingly the estimates get less similar as the cutoff gets shorter.

[Table 2b about here]

5.2. Software Results

Tables 3a and 3b contain our regressions for the supply of game titles as a function of the size of the installed base. The first row in each table contains IV results using a two-year lag in the installed base, and the second row contains IV results using a one-year lag in the installed base. Lastly, the third row contains results using a two-year lag but without instrumenting. In all specifications, we see notable variation in software supply along the quality dimension, with a peak around a quality level of 75. The effect tapers down from this point, and we note that, although the effect seems to jump up for titles ranked under 50, this group covers a much wider quality range than the others; hence, the effect clearly continues to diminish if we were to

[17] When cutoff thresholds are imposed, there are periods where the number of titles decreased or increased very little. Consequently, this pushes the estimate on the aggregate titles to be a smaller number.

consider similar quality ranges within this group[18].

[Tables 3a and 3b about here]

The last column of Table 3b shows our results when we aggregate all quality levels into a total count measure. Here, we see installed base has a positive effect on the overall supply of software.

We note here that, as with prior studies of the relationship between software supply and installed base, it is difficult to completely rule out spurious correlation due to mutual trending between installed base and software supply. Hence, our estimates could suffer from a positive bias. However, it is less clear that this would be the case when imposing a two-year lag, and our results without IVs indicate a bias that is negative when no IVs are used. Despite the challenge of finding ideal IVs, it is reasonable to believe, as our theory predicts and prior studies have argued, that there does exist a positive causal effect of installed base on software supply. Further, to the extent that our analysis can identify differing relative magnitudes of this effect (implying that bias, if there is any, is stable across software quality), our analysis comparing measured indirect network effects when one accounts for heterogeneity across quality levels versus when one does not is still valid in qualitative and percentage terms, but the level differences may be overstated. It is to this comparison that we now turn.

5.3. Indirect Network Effects

In this subsection, we compare, quantitatively, the indirect network effect we would find using a standard, aggregate measure of software variety to the effect we find using quality-differentiated measures. To do so, we consider an increase in installed base of one million, which is a one unit increase in our model (since we measure this variable in millions). Using aggregate measures, this would increase total titles by 19.9062 (using our 2-year lagged model). This would then increase mean utility—the dependent variable in our hardware regressions—by 19.9062 * 0.0129 = 0.2528. If, instead, we allow for differential effects along the quality dimension, an increase in installed base of one million will change mean utility by the sum of the

[18] We also run these estimates using instruments in the form of console age interacted with console dummies. We again observe the same general pattern.

pair-wise products of its effect on software supply and marginal utility for each quality level. This amounts to 0.939 * 0.0536 + … + 2.4687 * 0.0098 = 0.3584.

Hence, we see that the use of aggregate measures underestimates the true indirect network effect on mean utility by approximately 30 percent. Why do we see this underestimation for this market when using aggregate measures? As discussed in Section 4.3, there are two factors driving this result. The first is the correlation along the quality dimension between the marginal utility of software (λ_i) and the response of software supply to an increase in installed base (γ_i). We graph these two measures in Figure 3 below. While it may appear that these two measures are positively correlated in the figure, they are in fact mildly negatively correlated (-0.24). This negative correlation would tend to cause measured indirect network effects using aggregate software to be overstated (the opposite of what we find). Therefore, it must be the case that the second factor—correlation along the quality dimension between the marginal utility of software (λ_i) and conditional variation in software availability (w_i)—is driving our result. While we don't have direct measures for w_i, we can assess whether $\lambda = \sum_{i=1}^{L} w_i \lambda_i = (1/L) \sum_{i=1}^{L} \lambda_i$. From our results in Table 2, $\lambda = 0.0129$ and the average of the quality-specific marginal utilities is 0.0201. This implies that the impact of higher quality games has relatively less weight due to less overall variation in software coming from variation in high-quality games.

[Figure 3 about here]

In sum, our analysis indicates the negative correlation between marginal utility of software and partial variation in software availability is the dominant factor behind the underestimation of indirect network effects when using aggregate software counts. The direction and magnitudes we find using our data may vary for other studies. However, to the extent that there is variation in marginal utility of software, and either variation in software supply response to installed base or conditional variation in software availability, in a given market and associated data, the issues we address here are relevant.

5.4. Impact on Market Shares

It is also instructive to interpret the magnitude of this underestimation in indirect network effects in terms of market shares. For a given observed market share of a console at a given point

in time (s_{jt}), we can ask what is the implied percentage increase in this share when installed base increases by one million two years prior (at t-104). We can then compare this measure for the case where we allow for heterogeneous effects along the quality dimension and the case where we do not. When we allow for heterogeneous effects, the change to mean utility when installed base at t-104 increases by one million is 0.3584, as calculated above.

The calculation is done by noting that $s'_{jt} = \exp(\ln(s'_{0t}) + \delta_{jt} + 0.3584)$ with the increased installed base and $s_{jt} = \exp(\ln(s_{0t}) + \delta_{jt})$ without it, where δ_{jt} is mean utility with no increase in installed base. Since, on average, the share of the outside option is approximately 99 percent, and an increase in market share of one of the consoles would diminish the share of the outside option by far less than one percentage point, $s'_{0t}/s_{0t} \cong 1$. Hence, if we assume $s'_{0t} = s_{0t}$, then $\ln(s_{jt}) - \ln(s_{0t})$ would increase by 0.3584. It then follows that the market share s_{jt} increases by approximately ($e^{0.3584} - 1$), or 43 percent.

We can compare this to the predicted increase when we do not allow for heterogeneous effects. In this case, s_{jt} increases by ($e^{0.2528} - 1$), or 29 percent. Thus, by not allowing for heterogeneous effects, we would underestimate the indirect network effect as it pertains to market shares by approximately one third.

6. Conclusion

Economists have long noted the importance of network effects in many industries. Existing empirical studies have abstracted from the quality dimension of the complementary goods in estimating the magnitude of indirect network effects. In this paper, we introduced vertical differentiation into the standard monopolistic competition model and examined the implications for supply response and marginal utilities. In particular, the use of aggregate software counts can mis-measure indirect network effects when there is correlation along the quality dimension between the marginal utility of software, and either 1) the response of software supply to an increase in installed base, or 2) conditional variation in software availability in the data being utilized. In our analysis, we found that using aggregate measures underestimates the indirect network effects in the 7th-generation video game console market, implying the latter, negative correlation dominates. This approach is applicable to many markets where indirect network

effects may be present, and our analysis suggests that accounting for heterogeneity in quality can result in economically significant, different measured effects and implications for market shares.

References

Ackerberg, D. and Gautam, G. (2006), "Quantifying Equilibrium Network Externalities in the ACH Banking Industry," *RAND Journal of Economics* 37: 738–761.

Basu, A., Mazumdar, T., and Raj, S.P. (2003), "Indirect Network Externality Effects on Product Attributes," *Marketing Science* 22: 209–221.

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

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

Binken, J. and Stremersch, S. (2009), "The Effect of Superstar Software on Hardware Sales in System Markets," *Journal of Marketing* 73: 88–104.

Bresnahan, T., Stern, S., and Trajtenberg, M. (1997), "Market Segmentation and the Sources of Rents from Innovation: Personal Computers in the Late 1980s," *RAND Journal of Economics* 28: S17–S44.

Caillaud, B. and Jullien, B. (2003), "Chicken and Egg: Competition Among Intermediation Service Providers," *RAND Journal of Economics* 34: 309–328.

Chou, C.-F. and Shy, O. (1990), "Network Effects without Network Externalities," *International Journal of Industrial Organization* 8: 259–270.

Church, J. and Gandal, N. (1992), "Network Effects, Software Provision, and Standardization," *Journal of Industrial Economics* 40: 85–103.

——— and ——— (1993), "Complementary Network Externalities and Technological Adoption," *International Journal of Industrial Organization* 11: 239–260.

Clements, M. and 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. and Lederman, M. (2009), "Software Exclusivity and the Scope of Indirect Network Effects in the US Home Video Game Market," *International Journal of Industrial Organization* 27: 121–136.

Derdenger, T. (2012), "Technological Tying and the Intensity of Competition: Empirical Analysis of the Video Game Industry," working paper, Carnegie Mellon University.

Dixit, A. and Stiglitz, J. (1977), “Monopolistic Competition and Optimum Product Diversity,” *American Economic Review* 67: 297–308.

Dubé, J.P., Hitsch, G., and Chintagunta, P. (2010), “Tipping and Concentration in Markets with Indirect Network Effects,” *Marketing Science* 29: 216–249.

Gandal, N., Kende, M., and Rob, R. (2000), “The Dynamics of Technology Adoption in Hardware/Software Systems: The Case of Compact Disc Players,” *RAND Journal of Economics* 31: 43–61.

Gaudiosi, J. (2007), “How the Wii Is Creaming the Competition,” CNNMoney.com, http://money.cnn.com/magazines/business2/business2_archive/2007/05/01/8405654/index.htm.

Gowrisankaran, G. and Stavins, J. (2004), “Network Externalities and Technology Adoption Lessons from Electronic Payments,” *RAND Journal of Economics* 35: 260–276.

Gretz, R. (2010), “Hardware Quality vs. Network Size in the Home Video Game Industry,” *Journal of Economic Behavior & Organization* 76: 168–183.

Grossman, G. and Helpman, E. (2002), “Integration versus Outsourcing in Industry Equilibrium,” *Quarterly Journal of Economics* 117: 85–120.

Hagiu, A. (2006), “Pricing and Commitment by Two-Sided Platforms,” *RAND Journal of Economics* 37: 720–737.

Hogendorn, C. and Yuen, K.Y. (2009), “Platform Competition with ‘Must-Have’ Components,” *Journal of Industrial Economics* 57: 294–318.

Katz, M. and Shapiro, C. (1985), “Network Externalities, Competition, and Compatibility,” *American Economic Review* 75: 424–440.

Lee, R.S. (2013), “Vertical Integration and Exclusivity in Platform and Two-Sided Markets,” forthcoming in *American Economic Review*.

Nair, H., Chintagunta, P., and Dubé, J.-P. (2004), “Empirical Analysis of Indirect Network Effects in the Market for Personal Digital Assistants,” *Quantitative 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 and Management Strategy* 9: 513–548.

Prieger, J. and Hu, W.-M. (2012), “Applications Barriers to Entry and Exclusive Vertical Contracts in Platform Markets,” *Economic Inquiry* 50: 435–452.

Rysman, M. (2004), “Competition between Networks: A Study of the Market for Yellow Pages,” *Review of Economic Studies* 71: 483–512.

Stokey, N. (1981), “Rational Expectations and Durable Goods Pricing,” *Bell Journal of Economics* 12: 112–128.

Waldfogel, J. (2011), “Bye, Bye, Miss American Pie? The Supply of New Recorded Music since Napster,” NBER Working Paper No. 16882.

Zhu, F. and Iansiti, M. (2012), “Entry into Platform-based Markets,” *Strategic Management Journal* 33: 88–106.

Tables

Table 1

Summary Statistics

Variable	Mean	Std. Dev.	Min	Max	Obs.
Console share	0.002532	0.003781	0.000209	0.04563	905
Outside share	0.992475	0.010475	0.903163	0.999659	905
Console inside share	0.371271	0.196686	0.073846	1	905
Price	314.7966	102.0075	149.99	599.99	905
Titles 90-99	12.94365	10.36876	0	38	905
Titles 85-89	23.97901	20.09757	0	67	905
Titles 80-84	41.94365	29.30338	2	103	905
Titles 75-79	44.62983	28.79726	3	117	905
Titles 70-74	46.43425	29.09856	0	104	905
Titles 65-69	41.14475	28.02259	0	96	905
Titles 60-64	34.41657	24.06499	0	80	905
Titles 55-59	28.21326	19.04367	0	64	905
Titles 50-54	24.45525	16.65602	0	54	905
Titles <50	45.70276	35.67635	0	114	905
Console age (weeks)	152.3105	88.8662	1	336	905
Holiday	0.125967	0.331996	0	1	905
Installed base	16.78098	12.17484	0.15547	43.72277	905
Installed base 52wk lag (mil.)	13.66823	10.28382	0.15547	43.72277	749
Installed base 104wk lag (mil.)	10.44496	8.033499	0.15547	32.74407	593
Households with TV (mil.)	113.8279	1.640153	109.9945	115.9	905
# of Competitors	3.363536	0.690803	3	5	905
Sum competitors' processing speed	9894.24	4259.831	1518	15368	905
Sum competitors' memory	761.3116	247.5779	143	1167	905
Sum competitors' word length	200.7337	41.37922	160	288	905
# competitors from same generation	1.885083	0.463309	0	2	905
# competitors from same manufacturer	0.492818	0.500225	0	1	905

Table 2a[19]

Hardware IV nested logit results

Variable	(1)	(2)	(3)	(4)
	Titles in Level Form		Titles in Log Form	
Price	-0.0045+ (0.0025)	-0.0126** (0.0031)	-0.0023 (0.0027)	-0.0086** (0.0025)
Log Inside Share	0.2461+ (0.1322)	0.2496* (0.1248)	0.4944** (0.1079)	0.4463** (0.1157)
Titles 90-99	0.0536** (0.0130)		0.4108** (0.1096)	
Titles 85-89	0.0803** (0.0181)		0.7950** (0.2152)	
Titles 80-84	0.0299** (0.0112)		0.5722** (0.2213)	
Titles 75-79	0.0236** (0.0083)		-0.0171 (0.2028)	
Titles 70-74	0.0165+ (0.0096)		-0.3429+ (0.1774)	
Titles 65-69	-0.0293** (0.0082)		-0.2985* (0.1374)	
Titles 60-64	0.0232** (0.0089)		0.2624+ (0.1223)	
Titles 55-59	0.0081 (0.0101)		0.2783** (0.1024)	
Titles 50-54	-0.0147 (0.0158)		-0.0797 (0.1127)	
Titles <50	0.0098 (0.0075)		-0.0343 (0.1054)	
Titles		0.0129** (0.0019)		0.7379* (0.2985)
R-squared	0.891	0.858	0.884	0.871
Observations	905	905	905	905

[19] All regressions include console-level fifth-order polynomial in console age (in weeks), console fixed effects, console-level dummy variables for Holiday, and a constant term. Standard errors are robust to arbitrary heteroskedasticity. Inference results also robust to clustering at the console-month level, except for marginal loss of statistical significance for titles ranked 60-64 and 70-74 in the logged regression (3rd column of results). ** is significant at 1%, * is significant at 5%, and + is significant at 10%.

Table 2b[20]

Hardware IV nested logit results

Variable	(1)	(2)	(3)	(4)	(5)	(6)
	Exclusive Titles Only		**Non-exclusive Titles Only**		**Titles Dropped after 3 years**	
Price	-0.0039* (0.0018)	-0.0121** (0.0032)	-0.0069** (0.0022)	-0.0125** (0.0048)	-0.0107** (0.0034)	-0.0124** (0.0031)
Log Inside Share	0.4947** (0.1122)	0.3253* (0.1287)	0.1448 (0.1294)	0.2450 (0.1798)	0.1221 (0.1426)	0.3420** (0.1209)
Titles 90-99	0.1609** (0.0358)		0.0835** (0.0179)		0.0645** (0.0147)	
Titles 85-89	0.1850** (0.0329)		0.0510** (0.0188)		0.0401** (0.0149)	
Titles 80-84	-0.0284 (0.0206)		0.0226* (0.0115)		-0.0157 (0.0126)	
Titles 75-79	-0.0522** (0.0135)		0.0851** (0.0105)		-0.0156* (0.0074)	
Titles 70-74	0.0040 (0.0179)		0.0557** (0.0128)		-0.0017 (0.0069)	
Titles 65-69	-0.0186 (0.0154)		-0.0260* (0.0109)		-0.0399** (0.0089)	
Titles 60-64	0.0500** (0.0184)		0.0121 (0.0112)		0.0084 (0.0090)	
Titles 55-59	-0.0045 (0.0155)		0.0344* (0.0134)		-0.0044 (0.0099)	
Titles 50-54	0.0134 (0.0327)		-0.0208 (0.0146)		-0.0118 (0.0127)	
Titles <50	0.1015** (0.0161)		-0.0116 (0.0085)		0.0220** (0.0077)	
Titles		0.0178** (0.0050)		0.0196** (0.0042)		-0.0043 (0.0026)
R-squared	0.896	0.857	0.887	0.860	0.859	0.854
Observations	905	905	905	905	905	905

[20] All regressions include console-level fifth-order polynomial in console age (in weeks), console fixed effects, console-level dummy variables for Holiday, and a constant term. Standard errors are robust to arbitrary heteroskedasticity. ** is significant at 1%, * is significant at 5%, and + is significant at 10%.

Table 3a[21]

Software regression results

	Titles 90-99	Titles 85-89	Titles 80-84	Titles 75-79	Titles 70-74	Titles 65-69
(IV) Installed $Base_{t-104}$	0.9867** (0.1373)	1.5494** (0.1898)	2.4790** (0.2917)	2.7688** (0.3181)	2.6395** (0.2736)	2.1591** (0.2364)
(IV) Installed $Base_{t-52}$	0.9390** (0.1177)	1.5389** (0.1609)	2.4944** (0.2470)	2.6793** (0.2641)	2.7564** (0.2338)	2.3249** (0.2066)
(no IV) Installed $Base_{t-104}$	0.4400** (0.0990)	0.7211** (0.1485)	0.9969** (0.1986)	1.3581** (0.2538)	1.3360** (0.2280)	1.1131** (0.1845)

Table 3b

Software regression results

	Titles 60-64	Titles 55-59	Titles 50-54	Titles < 50	Titles
(IV) Installed $Base_{t-104}$	2.0348** (0.2257)	1.7280** (0.1815)	1.0923** (0.1425)	2.4687** (0.2856)	19.9062** (1.7747)
(IV) Installed $Base_{t-52}$	2.1369** (0.2025)	1.7469** (0.1581)	1.2646** (0.1316)	2.7090** (0.2538)	20.5904** (1.4665)
(no IV) Installed $Base_{t-014}$	1.1521** (0.1899)	1.0267** (0.1869)	0.6064** (0.1227)	1.2964** (0.2577)	10.0469** (1.6396)

[21] All regressions include console-year-month fixed effects and a constant term. Standard errors are robust to arbitrary heteroskedasticity. Standard error estimates change only trivially when applying 3SLS to quality groupings (e.g., Titles 90-99, Titles 85-89, and Titles 80-84). Inference results also robust to clustering at the console-month level. ** is significant at 1%, * is significant at 5%, and + is significant at 10%.

Figure 1

Correlation causing overestimation of indirect network effects using an aggregate model (with equal weighting of quality groups)

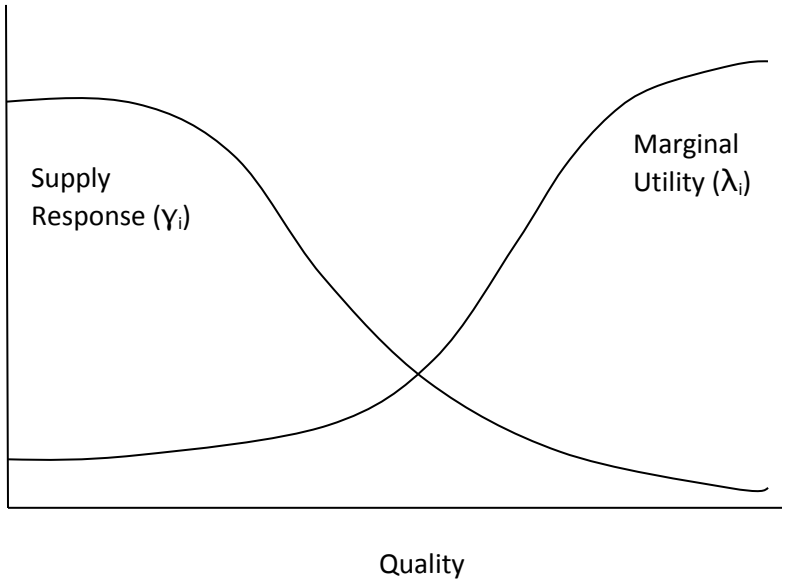


Figure 2

Correlation causing underestimation of indirect network effects using an aggregate model (with equal weighting of quality groups)

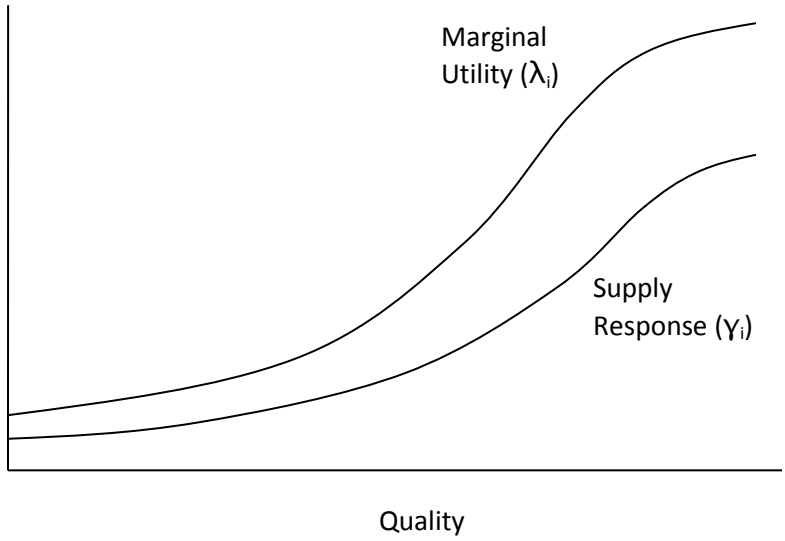
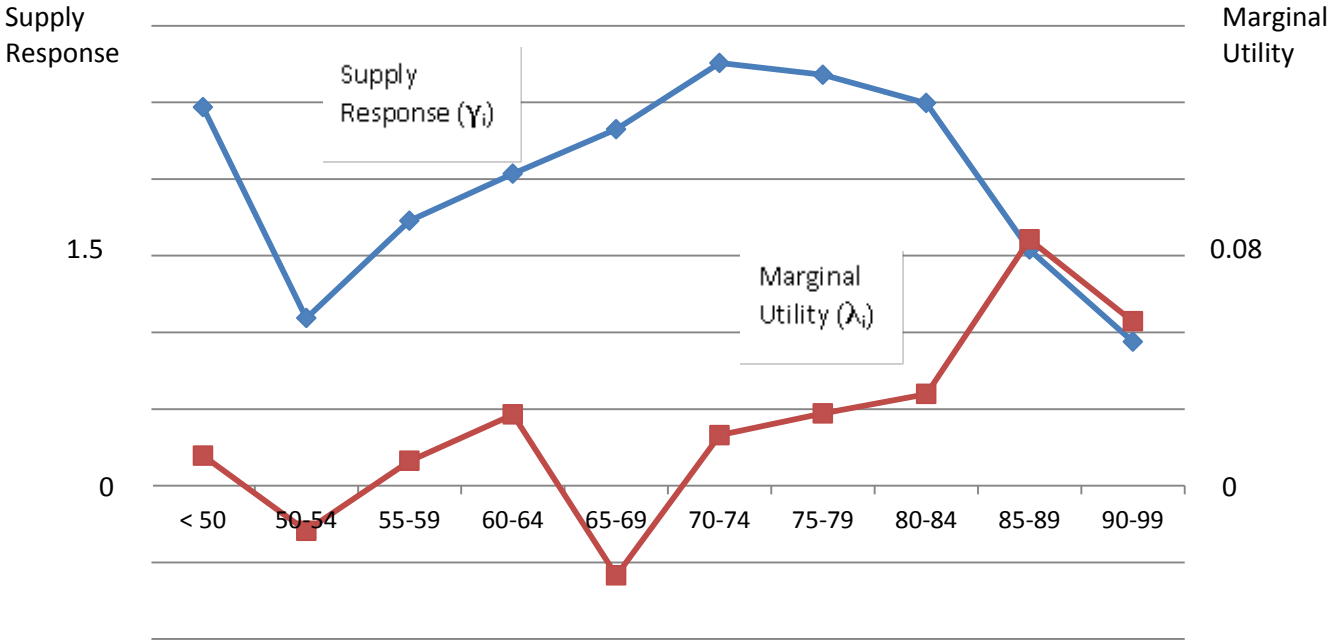


Figure 3

Observed pattern of marginal utility and supply response