The Challenge of Revenue Sharing with Bundled Pricing: An Application to Digital Music

Benjamin Shiller
The Wharton School
University of Pennsylvania

Joel Waldfogel
The Wharton School
University of Pennsylvania and NBER

Preliminary Draft – September 14, 2009

Abstract

Bundling can increase revenue and profits relative to selling products on a standalone basis, and this is an especially attractive strategy for zero-marginal-cost information products. Despite the clear benefits of bundling, it has one major problem: bundling produces revenue that is not readily attributable to particular pieces of intellectual property, creating a revenue division problem. The Shapley value provides a well-motivated solution to this problem, and we use unique survey data to create measures of bundle value and, in turn, to estimate Shapley values for each of 50 bundle elements. We then evaluate feasible revenue sharing schemes, including equal sharing, proportional sharing, and the modified Shapley value of Ginsburgh and Zang (2003, 2004). We first document that the Shapley value is highly incentive compatible (all bundle elements fare better inside the bundle than they do outside on a standalone basis). We then evaluate the feasible schemes according to both their incentive compatibility and their similarity with the Shapley value. We find, not surprisingly, that the feasible schemes are less incentive compatible than the Shapley value. Among the feasible schemes, the GZ scheme performs best while equal sharing performs worst.

It is well known that bundling products can increase revenue and profits relative to selling products on a standalone basis, provided that product valuations are not perfectly positively correlated across consumers.[1] Bundling is an especially attractive strategy for information products, whose negligible marginal costs make it easy to bundle large numbers of products (see Bakos and Brynjolffson, 1999). Digital music provides an auspicious context of considerable practical interest, and bundling of 10 or more popular songs has been shown (Shiller and Waldfogel, 2009) to raise revenue by up to a third relative to selling individual songs. The attraction of bundling is not lost on practitioners: Apple's iTunes Music Store, is reported to be contemplating bundled ("all-you-can-eat") song sales, either through one-time fees or ongoing periodic charges. Other services, such as Nokia Music Store, Napster, Rhapsody, and eMusic, are already selling bundled offerings.[2]

Despite the clear benefits of bundling, it has one major problem: bundling produces revenue that is not directly attached, or readily attributable, to particular pieces of intellectual property, leaving a seller of bundles with a revenue division problem.

Fortunately, game theorists have developed tools well suited to solving problems of this type. Since Shapley's famous 1953 paper, a large body of work in cooperative game theory has been shown useful for application to problems regarding sharing costs and distributing spoils.[3] The well-known Shapley value, in this case the expected marginal revenue of each song,

---

[1] See Stigler (1963), Adams and Yellen (1976), Schmalensee (1984), Armstrong (1999), and Bakos & Brynjolffson (1999).

[2] Nokia allows unlimited downloads from its music store for a 12-18 month period on certain phones available in many countries abroad, including the United Kingdom, Brazil, Mexico and Germany (Lionel, 2009). Napster allows users to stream music to their computer, and awards credits towards downloads for $7 monthly. At Rhapsody, users can "listen to millions of songs without paying per track. Play all the music you want for one low [$12.99] monthly price." (see http://learn.rhapsody.com/?src=rcom_navside , accessed June 5, 2008). eMusic offers tiered subscriptions, including 30 song downloads per month for $9.99 per month, 50 song downloads per month for $14.99 per month, or 75 song downloads per month for $19.99 per month. (see http://www.emusic.com/help/index.html#q4, accessed June 5, 2008).

[3] See Roth (1988) and the contributions therein.

averaging over all of its possible arrival orders in the bundle, provides a solution to the revenue-sharing problem. In general, the Shapley value calculation requires information on the revenue available to every conceivable bundle. While this is generally unknowable, we have detailed survey based data on nearly 500 college students' valuations of 50 popular songs in early 2009, obtained by direct elicitation. The data contain each individual's valuation of each song and therefore each bundle of songs, allowing us to perform pricing exercises that generate the maximal revenue available to each possible bundle. Even with our unusually good data, calculation of exact Shapley values for 50 songs is computationally challenging, so we develop a random sampling approach that produces Shapley values to arbitrary levels of accuracy.

While the Shapley value provides a theoretically coherent method for sharing bundle revenue, because of data and computational challenges, it is not in general calculable. So, in addition to calculating Shapley values, we also examine revenue sharing schemes that can, and in some cases already are, being implemented in reality. We ask how far they – and the Shapley value – deviate from incentive compatibility (whether IP owners are better off inside the bundle) as well as how closely the feasible schemes track the Shapley value.

We have two main findings. First, we find that the Shapley value produces an incentive-compatible revenue allocation in our context: each song receives more compensation inside the bundle than it would outside on a standalone basis. Second, we find that feasible schemes deviate substantially from both incentive compatibility and the Shapley value. Of the feasible schemes, equal sharing is furthest from both the Shapley value and incentive compatibility,while sharing according to the simplified Shapley value of Ginsburgh and Zang (2003, 2004) deviates least from both the Shapley value and incentive compatibility. We conclude that the revenue sharing problem is an important one for current business practice. Our paper proceeds as follows. First, we review the literature on the benefits of bundling, in particular in the context of information products. Second, we describe the Shapley value as a method for distributing bundle revenue, and we outline an approach for computing Shapley values for each of the 50 songs that can be included in bundles, which we illustrate with our data. Third, we describe our data and how we use it to develop valuations of all possible bundles of songs. Fourth, we present results, evaluating the Shapley value and its more easily implemented alternatives. A brief conclusion follows.

I. Literature Review

Bundling can raise the revenue available to a group of products. The benefits are particularly clear when products' valuations are negatively correlated across consumers (Stigler, 1963; Adams and Yellen, 1976), but bundling can raise revenue even when valuations are positively correlated (Schmalensee, 1984). Moreover, the benefits of bundling can be expected to increase as the number of products in the bundle grows large (Bakos and Brynjolffson, 1999).

In a companion paper (Shiller and Waldfogel, 2009) we use surveys of college students to characterize the distribution of consumers' valuations of 50 songs. We show that the 50 songs can generate 28.8 percent more revenue under pure bundling than they do under separate sales with uniform pricing (and 25.2 more than under separate sales with song-specific, or component, pricing). Furthermore, the benefits of bundling increase with bundle size relative to uniform (component pricing); on average bundling raises revenue by 16.4% (14.0%) for 5 song bundles, by 21.6% (18.7%) for 10 song bundles, and by 26.6% (23.2%) for 25 song bundles.

While bundling can clearly increase the size of the pie, the correct method for slicing the pie of bundle revenue is not obvious. When each song is sold separately, each has identifiable revenue, while under bundling the revenue must be divided among the songs.

Cooperative game theory provides substantial guidance on coherent methods for sharing bundle revenue. Drawing from Young (1988), one can start with some simple axioms that division schemes should obey. These include:

1) Distributions to the various songs should exhaust bundle revenue.
2) Songs that enter the bundle symmetrically should receive equal shares.
3) A song's share should depend only on its own contribution to revenue.

The Shapley value (Shapley, 1953) is the unique sharing rule obeying these properties.

The Shapley value for song *j* may be written as:

$$\phi_j^*(R_B) = \sum_{S \subseteq N-j} \frac{|S|!(|N-S|-1)!}{|N|!}[R_B(S+j) - R_B(S)].$$

Where $|S|$ is the number of songs in subset *S*, and $R_B(S)$ is the revenue from a bundle including the subset of songs *S* among the *N* songs total.

The formula has the intuitive interpretation as the average marginal contribution of each element. That is, the Shapley value is the average of song *j*'s marginal contribution when it is first – and therefore alone – in the bundle, and when it is second (following any of the other *N*-1 songs), and when it is third (following the (*N*-1)(*N*-2) combinations that can precede it), and so on[4]. When the number of elements is small, say 3, then one can easily calculate the Shapley value by hand. As the number of elements grows, the calculation required increases enormously. Before turning to our method for calculating Shapley values for each of 50 products, it is useful to describe our data.

[4] For a simple example, see the "glove game," covered in Mas-Colell, Whinston, and Green (1995).

II. Data

The basic data for this study are drawn from a survey of 433 Wharton undergraduates performed in January 2009. Those surveyed, students in an introductory economics course, were required to fill out an online form indicating their valuations for each of the top 25 songs at iTunes that week, along with 10 songs that were the top 10 songs 6 months earlier, and 15 songs drawn randomly from those ranked between 26 and 100 at the time of the survey. Students were given instructions and paper worksheets in class prior to the survey. For each song, students were told to listen to a clip to remind themselves of the song, then to write down the maximum amount they would be willing to pay to get the song from the sole authorized source. Completion of the assignment was necessary for problem set credit to motivate students to participate carefully.

In particular, students were given the following instructions:

> Imagine that, unlike in current reality, there is only one authorized source for each song. Put aside what you know about prices at existing outlets because for this survey we're pretending that they don't exist.
>
> For each song listed in the survey, indicate the maximum amount you would be willing to pay to obtain it from the sole authorized source. **For this exercise, I'm asking you to report what it is worth to you, not what price you think would be fair or what price you are accustomed to paying. That is, I'm asking you to indicate the maximum amount you would be willing to pay to obtain it from the authorized source.**
>
> For example, if you already purchased it, then at the time you bought it, you were willing to pay at least the price you paid but you might have been willing to pay more. If you would prefer not to have it even if it were free, you would indicate 0.
>
> On the following pages, you will be presented with a list of songs and artists. In the space provided for each song, enter the maximum amount you are willing to pay for the song (for example 1.75, NOT $1.75). You must enter a dollar amount for each song.

The resulting dataset included 21,650 observation from 433 respondents.

Eliciting song reservation valuations via surveys has antecedents in the marketing and operations literatures. Hanson and Martin (1990), Kalish and Nelson (1991), Venkatesh and

Mahajan (1993), and Jedidi, Jagpal, and Manchanda (2003) all use similar survey methods. Though, some remain skeptical of the validity of response data due to the sensitivity of question wording (Diamond and Hausman, 1994). Some existing earlier work suggests that buy-based direct elicitation, used in this paper, yields reasonable results in sample. Rob and Waldfogel (2006) find that implied optimal prices using buy-based elicitation tend to be close to observed prices. Kalish and Nelson (1991), testing various survey methods, find that reservation values do well in terms of fit while the other measures are superior in predicting choice on a holdout sample. Another concern with survey data is the magnitude of measurement error. Jedidi, Jagpal, and Manchanda (2003) noted that self-stated reservation prices for infrequently purchased songs are subject to measurement error. Though, in our survey , there is reason to believe the respondents were familiar with the product – we only elicited valuations for popular songs, and students were instructed to remind themselves of the song prior to providing an answer. Thus, prior research is supportive, though not unanimously, for buy-based direct elicitation of valuations for familiar products.

The data themselves provide further support for the contention that the data used in this paper are reasonable. First, survey responses do not seem to be disproportionately anchored on the well-known $0.99 price on iTunes. While there were a considerable number of $0.99 and $1.00 valuations in the survey data, the frequency of these amounts seems to be a result of rounding, not anchoring. Valuations of $0.50, and $2.00 each occurred nearly as often as valuations between $0.99 and $1.00.

A second reasonableness check is whether the reported valuations can be relied upon to predict purchase behavior. To address this, we can perform a simple check on the data: do respondents actually report possessing songs that they report valuing at or above $0.99? We begin by dropping observations obtained through file-sharing, since acquiring a song without monetary cost is consistent with valuations above or below \$0.99 cents. For the remaining song-respondent pairs, we compute the probability that that a song is owned as a function of reported valuation. Specifically, we divide, by valuation, the data into 50 equal-sized groups, and compute for each the empirical probability of song ownership. When the probability of ownership is plotted against the logarithm of the average valuation, it is apparent that the probability of ownership increases consistently with song valuation. For valuations near, but below \$0.99, about 20% reported owning the song. Just above \$0.99, the probability of ownership is 50%. As valuations increase further, so does the probability of ownership. The clear positive relationship between valuation and probability of ownership provides further evidence that the data are reasonable. For more analyses supporting the reasonableness of the data, see Shiller and Waldfogel (2009).

Our only remaining concern is the apparent rounding of responses in the data. A simple histogram of valuations (not presented here) verifies that valuations tend to be rounded to the nearest quarter, or the nearest quarter minus one cent (e.g. \$0.99). Such rounding gives rises to plateaus on the demand curve that create "sawtooth" spikes in a plot of revenue against price. We take the view that underlying preferences are smooth, so that valuations do not exhibit clumping on multiples of 25 cents. Rather, the survey observations contain measurement error arising from rounding. We could ignore this, and proceed with the raw data, but this measurement error distorts the relative effectiveness of some pricing schemes. Specifically, the estimated maximum revenue obtained under uniform pricing, used as a counterfactual below, is inflated by rounding. To see this, suppose individuals have a tendency to round their responses to the nearest \$0.25. Then, for example, the observed demand at \$1.00 includes true valuations as low as \$0.875 – it is overstated. In fact, the estimated demand at any valuation X between \$0.875 and \$1.00 will be overstated, because it will include valuations that are truly below X, but were rounded to \$1.00 in the observed data. A sufficient amount of rounding guarantees that the observed maximum revenue will exceed the true maximum revenue, even if true revenue is maximized at a price at which demand is underestimated. In contrast, rounding does not inflate bundle revenue much, because rounding at the song level tends to average out in large bundles. Thus, the benefit of bundle pricing over uniform pricing is underestimated in rounded valuation data. Because we are evaluating whether various bundle revenue distribution schemes would yield full participation in the bundle, we care a good deal about the impact of rounding. So we must find a way to "unround" the data.

One natural way to remove the impact of rounding is to fit the observed data to a smooth distribution. We found that the data seem to fit a zero inflated lognormal distribution. To fit the data to this distribution, we first modeled zero valuations as: $y_{is} = \theta_s + e_{is}$, where i denotes individual, s denotes song. We say that individual i has a positive valuation for the song if y_{is} is greater than zero. We assume e is distributed standard normal, and estimate each θ_s by running song-specific probit regressions. To estimate the correlation of e across songs, we estimate a bivariate probit regression for each song pair. We next model the log positive valuations, v_{is}, as: $v_{is} = \mu_s + \varepsilon_{is}$, where μ is the mean, and ε is normally distributed. We estimate μ_s, and the standard devation of ε, σ_s, using only positive valuations for song s. We then estimate the correlation between ε_s and ε_t, for each pair of songs s and t, by restricting the sample to observations with positive valuations for both. Note that, because μ_s and θ_s are modeled separately, there are no restrictions on their relationship.

We use these parameters to simulate valuations for 5000 individuals. Specifically, we use the estimated parameters for the positive valuations to populate the 50x1 matrix of means μ_s, and the 50x50 covariance matrix. We draw the log positive valuations for 5000 individuals and exponentiate. We then simulate the process for generating zeros by drawing each y_{is}, in a similar fashion, and zeroing individual *i*'s valuation for song *s* if y_{is} was less than or equal to zero. The simulated data yields smooth, rather than saw-toothed, revenue functions, alleviating the problems caused by rounded responses. In all subsequent analyses, we use these simulated zero-inflated, multivariate lognormally distributed data.

III. Results

Calculating Shapley Values

While Shapley values are easy to calculate for small sets of products, as the number of products grows, the calculation becomes substantially more cumbersome. With 50 songs – as in our example – there are 50 factorial orderings of songs. This gives rise to Shapley values that are each the sum of roughly 3.04 x 10^{64} terms. These sums would take a very long time to compute.

Rather than attempt to calculate the exact Shapley values, we estimate them by randomly sampling among the possible orderings. The Shapley value for song *x* in a bundle of *N* songs is the weighted sum of marginal revenues in the all of ways it can occur in each order in the bundle. There are *N-1!* orderings in which it appears first. In all of these, its value to the bundle is simply the revenue available to a bundle consisting of only *x*, or $R_B(\{x\})$. Similarly, there are

$(N\text{-}1)!$ orderings in which it appears second. Song x's value to the bundle in all of these are the 49 different possible values of $R_B(\{y,x\}) - R_B(\{x\})$ for y = 1,…,50, $y \neq x$. We can easily calculate $R_B(x)$, as well as the 49 different values of $R_B(\{y,x\})$. But suppose we are interested in each of the orderings in which x appears, say 25th. There are "49 choose 24" combinations of song orders preceding song x. That is, there are 49!/(25!24!), or 6.32 x 10^{13} combinations, leading to different values of $R_B(\{first\ 24,\ x\})$ - $R_B(\{first\ 24\})$.

Our approach is to randomly sample one hundred times from the 49 choose 24 possible orderings. This gives rise to an estimate of song x's contribution when it appears 25th in the bundle. We similarly sample to produce estimates of song x's marginal contribution when it is in each of the possible 50 orders. Each estimated Shapley value is thus built up from 5000 (50 times 100) estimates of $R_B(bundle\ including\ x) - R_B(bundle\ excluding\ x)$.

We know the revenue available to the entire bundle of 50 songs, so we have a ready check on the reasonableness of our estimated Shapley values, whether they sum to $R_B(\{whole\ bundle\})$. As we will see, we can get arbitrarily close to the true values with samples of manageable size.

Define $\Delta R^l_{B,j}$ as the average marginal revenue of song j when it appears in the l^{th} position in the bundle. The term $\Delta R^l_{B,j}$ is an average over many individual incremental revenues, and we cannot calculate it directly. However, we can estimate it, with arbitrarily high precision. Define $\Delta \tilde{R}^l_{B,j}$ as a single draw on the incremental revenue of song j when it appears in position l. That is, it is a single draw of $R_B(\{A, i\})$-$R_B(\{A\})$, where A is a set of l-1 random songs sampled from the group of songs that does not include song i. We can estimate $\Delta R^l_{B,j}$ by taking the average of draws of $\Delta \tilde{R}^l_{B,j}$. Our estimate of the average value of song j in position l, based on T draws is then: $\overset{\wedge}{\Delta R}{}^{l}_{B,j}(T) = \sum_{q=1}^{T} \Delta \tilde{R}^{l}_{B,j} \Big/ T$. The Shapley value for song *j*, based on *T* draws, the simple

average across all of its possible positions, is then estimated as: $\hat{\phi}_j(T) = \sum_{l=1}^{50} \overset{\wedge}{\Delta R}{}^{l}_{B,j}(T) \Big/ 50$.

Finally, the full bundle revenue should equal the sum of these Shapley values, across all 50 songs: $R_B(\{full\_bundle\}) = \sum_{j=1}^{50} \hat{\phi}_j(T)$.

Given the large number of terms in the exact calculation, it is not obvious how many draws (*T*) are required for accurate estimates of the Shapley values. To explore this we begin by simulating $\Delta \tilde{R}^{l}_{B,j}$ (based on single draws). We perform 100 draws for each song/order. This allows calculation of the standard deviation of a one-draw Shapley values.

Because Shapley values based on *T* draws are simply the means of *T* one-draw Shapley values, if we know the standard deviation for one-draw Shapley values, we can calculate the standard deviation of *T*-draw Shapley values using elementary statistics. If X~N(μ,σ), then $\overline{X} = \sum_{i=1}^{N} \frac{X_i}{N}$ ~N(μ,σ/√N).

Table 1 provides estimates of songs' Shapley values as well as the standard errors of these estimates. The actual total revenue of the full 50-song bundle is \$83,543, and the Shapley value estimates sum to \$83,520, providing encouragement that our procedure is reasonable. The songs' Shapley value estimates average \$1,670, and the standard errors vary between \$8.7 and \$23.6, so we conclude that our estimates of the Shapley values are rather precise.

Criteria for Evaluating Revenue Sharing Schemes

Although we will use Shapley values as a benchmark for comparing other schemes, we can also evaluate Shapley values according to incentive compatibility, or whether bundle elements receive more revenue inside the bundle than outside it.[7] Outside the bundle, a song receives its maximized standalone revenue. This can, in turn, be calculated two ways, using a uniform price that maximizes revenue across all songs or using a song-specific price. We focus on the uniform pricing values, although the uniform and song-specific revenues are quite similar.

Beyond the question of incentive compatibility *per se*, there is the question of whether deviations from incentive compatibility are systematically predictable. For example, if more popular songs systematically lose inside the bundle, then owners of promising songs might refuse to join the bundle, leading to the unraveling of bundled sales schemes.

In addition to visual inspection, we have two direct measures of incentive compatibility as well as two direct measures of proximity with the Shapley value. Our first measure of incentive compatibility is the percent of songs earning more revenue in the bundle than out. Our second incentive compatibility measure is the cumulative sum of side-payments necessary to induce full bundle participation. Specifically, cumulative side-payments equal:

$\sum_{i=1}^{N} pos\left(SR_i - R_i\right)$, where SR_i is the standalone revenue, and R_i is the payment under the revenue sharing scheme, for song i. Because bundle revenue exceeds total standalone revenue, any of these revenue sharing schemes could be incentive compatible if owners of songs that gain by joining the bundle compensate owners of songs that lose by joining the bundle. Revenue sharing schemes that on the first measure are closest to 100 percent, and on the second measure are closest to zero, are closer to being incentive compatible.

[7] We restrict attention to standalone songs outside the bundle. We would alternatively ask whether any subgroup of the 50 songs could do better as a separate bundle outside the 50-song bundle. For simplicity, we ignore the multitude of possible outside-the-bundle bundles.

We then compare remuneration under each revenue sharing scheme with Shapley values, i.e. “fair” revenue distribution payments. The third measure is the correlation between remuneration in the bundle and the Shapley value. For the fourth measure, we compute the mean absolute percent difference (MAPD) between payments under the scheme and Shapley values. Specifically, the MAPD equals $\frac{\sum_{i=1}^{N} abs\left(SV_i - R_i\right) / SV_i}{N}$. The closer the correlation is to one, and lower the MAPD, the closer a revenue sharing scheme is to the Shapley value, which has theoretical claims to appropriateness in addition, in this context, to being incentive compatibility.

We begin by evaluating Shapley values by the various measures. The lower-right panels of Figures 1 and 2 plot Shapley values, on the vertical axes, against standalone revenue on the horizontal axes. The scatter lies above the 45-degree line, indicating that the Shapley value awards songs more inside the bundle than they would receive on a standalone basis outside. All of the 50 songs are better off inside, so average required side payments are zero. Tables 2a and 2b report these criteria for Shapley values, and each of the other sharing schemes..

$\sum_{i=1}^{50} pos(SR_i - SV_i)$

Evaluating Feasible Schemes

While the Shapley value has some theoretical claim to solving the revenue distribution problem, it is difficult to implement in practice because of both data and, to a lesser extent, computational problems. This leads us to consider schemes that are actually in use, along with other feasible schemes for sharing bundle revenue.[9]

[9] Brynjolfsson and Zhang (2006) propose a manageable method for obtaining additional data (not readily available) on information products in bundles. Specifically, their method yields estimates of product-specific demand curves for products only sold as part of a bundle.

While real markets do not routinely generate valuation data, they do generally produce data on usage. For example, it is typically possible to know what share of consumers download each of the elements of the bundle, which is equivalent to knowing how many individuals, amongst those buying the bundle, value each song above zero. It is also possible that sellers would know not only the number of consumers obtaining each song but also which songs had been downloaded by which consumers. With these sorts of data in mind, we describe three kinds of revenue distribution schemes.

The first is egalitarian sharing, in which each of N bundle elements (50 in our example) receives an equal share of bundle revenue. Egalitarian sharing is straightforward. If R_B is bundle revenue, and there are N songs, then each song gets R_B/N.

The upper left plots in Figures 1 and 2 plot revenues under equal sharing against standalone revenue (available to either uniform or component pricing). While standalone revenues, on the horizontal axis, vary substantially across songs, payments under the egalitarian approach are fixed and are not universally incentive compatible. Popular songs lose, and unpopular songs gain, under bundling with equal sharing, relative to their standalone opportunities. As Table 2 shows, despite the fact that bundling produces more than a quarter more revenue overall, 30 percent of songs are worse off under bundling with equal sharing. Relative to their uniform pricing standalone revenue, the songs that fare worse are \$7,678 worse off under bundling with equal sharing. That is, side payments totaling \$7,678 could leave everyone better off. To put this in perspective, these side payments constitute 9.2 percent of total bundling revenue, or 41.1 percent of the additional revenue that bundling generates beyond the revenue available with uniform pricing. Relative to individual Shapley values, the equal division payments are perforce uncorrelated with Shapley values, and the average of these deviations is 63 percent off the Shapley value.

The second feasible scheme is proportional sharing, in which elements get revenue in proportion to how many times they are selected for consumption. Proportional sharing is used at Rhapsody and eMusic, allocating revenue proportionally to consumption, i.e. downloads or plays. If C_j is the number of times that song j is consumed (downloaded) by subscribers, N is the total number of songs, and R_B is the total subscriber revenue to be divided among holders of intellectual property, then song j receives $\frac{R_B C_j}{\sum_{k=1}^{N} C_k}$.

The upper right plots of Figures 1 and 2 depict proportional sharing. Unlike with equal sharing, songs with higher standalone revenue also get higher proportional shares, so this scheme deviates less from incentive compatibility than equal sharing. Only 24 percent of songs are worse off inside the bundle than outside, and the cumulative side payment needed for incentive compatibility is about half that needed to render equal sharing incentive compatible, about 4.5 percent of bundle revenue or roughly a fifth of the extra revenue generated by bundled sales. The correlation of proportional sharing payments with the Shapley value is 0.91, and the MAPD is 42 percent off the Shapley value

A third feasible method is the simplified Shapley values of Ginsburgh and Zang (2003, 2004). The only data required for implementation is the list of songs consumed by each bundle buyer. The GZ scheme is intuitively simple. Each bundle participant pays P_b for participation. If he consumes 5 songs, each of those songs gets a fifth of P_b. The payments to each song are then aggregated across participants to yield each song's total payment. GZ's scheme is a variant on the Shapley value when no valuation data exist. The GZ method revenue is calculated via the formula:

$$R_{B,j}^{GZ} = \sum_{i=1}^{M} \left(\frac{I(C_{ij}) P_B}{\sum_{k=1}^{N} I(C_{ik})} \right)$$

Here, i denotes individuals, j denotes songs, N and M are the number of songs and individuals, respectively, P_B is the price of the bundle, and $I(C_{ij})$ is an indicator variable denoting consumption of song j by individual i.

GZ sharing is depicted in the lower right portions of Figures 1 and 2. Like proportional sharing, GZ sharing rewards songs with higher standalone revenue more richly. Roughly a fifth (18-20 percent) of songs are worse off in the bundle under this scheme, and the cumulative side payment is about $2500, which is 3 percent of bundle revenue, or about 12 percent of the extra revenue produced by bundling. As Table 3 shows, GZ shares are more highly correlated with Shapley values than proportional shares: 0.93 vs 0.91, and the average absolute percent deviation from Shapley values is 36 percent (compared with 63 percent for equal and 42 percent for proportional).

Conclusion

Although bundling can raise the total revenue available to groups of products together, its implementation requires a method for sharing revenue among owners of the products in the bundle. Sellers of digital music are now experimenting with bundling – and with methods for sharing revenue – elevating the importance of this problem.

When songs earn less inside a bundle than they would outside, then bundling can unravel, as appears to have occurred at eMusic, a seller using proportional sharing (Harding, 2007).

When artists become relatively popular, the labels are reluctant to make their new songs available on the site. After declining to make a label's biggest new songs available to the site, the label's publicist said, "the label plans to continue using eMusic to sell smaller releases and will post major releases after a yet-to-be determined lag time."

Using unusual data that allow us to calculate the value of all possible bundles, we are able to calculate a theoretically coherent revenue sharing scheme using the Shapley value. We document that this scheme is incentive compatible in our example. We then turn to the evaluation schemes that are feasible with the sorts of usage information routinely available to sellers. We evaluate equal sharing, proportional sharing, and sharing based on the modified Shapley value of Ginsburgh and Zang according to both their incentive compatibility and similarity to the Shapley value. We find that equal sharing is the least incentive compatible and furthest from the Shapley value, while the GZ scheme improves upon both equal and proportional schemes and is best among the feasible approaches we evaluate.

Although the GZ scheme improves on proportional sharing, it does not eliminate its incentive compatibility problem entirely. We conclude that the revenue sharing problem that we identify is an important problem and an interesting topic for future research.

References

Adams, W.J. and J.L. Yellen (1976): "Commodity Bundling and the Burden of Monopoly," Quarterly Journal of Economics, 90, 475{98.

Armstrong, M. (1999): "Price Discrimination by a Many-Product Firm," Review of Economic Studies, 66(1), Special Issue: Contracts, 151-68.

Bakos, Y. and E. Brynjolfsson (1999): "Bundling Information Goods: Pricing, Profits and Efficiency," Management Science, 45(12), 1613{30.

Brynjolfsson, E. and X. Zhang (2006): "Innovation Incentives for Information Goods," Innovation Policy and the Economy, 7,99{25.

Diamond, P. A and J.A. Hausman. "Contingent Valuation: Is Some Number Better than No Number?" Journal of Economic Perspectives, vol 8, number 4, fall 1994, pp 45-64.

Fang, H. and P. Norman (2006): "To Bundle or Not to Bundle," RAND Journal of Economics, Winter 2006.

Fisher, W. (2004): "Promises to Keep: Technology, Law, and the Future of Entertainment," Stanford University Press.

Geng, X., M.B. Stinchcombe, and A.B. Whinston (2005): "Bundling Information Goods of Decreasing Value," Management Science, 51(4), 662-67.

Ginsburgh, V. and I. Zang. "Sharing the Income of a Museum Pass Program." Museum Management and Curatorship, vol. 19, no. 4 (2001): 371-383.

Ginsburgh, V. and I. Zang. "The Museum Pass Game and its Value." Games and Economic Behavior, 43, (2003):322-325.

Hanson, W. and R. Kipp Martin. "Optimal Bundle Pricing." Management Science. Vol 36, no. 2. February 1990, 155-174.

Harding, Cortney. "The Indies: eMusic's Long Tail." Billboard Magazine, September 29, 2007.

Jedidi, K., S. Jagpal and P. Manchanda (2003): "Measuring Heterogeneous Reservation Prices for Product Bundles," Marketing Science, 22(1), 107-30.

Kalish, S. and P. Nelson. "A Comparison of Ranking, Rating, and Reservation Price Measurement in Conjoint Analysis." Marketing Letters 2:4 (1991): 327-335.

Laurient, Lionel – "Nokia Delays U.S. Music Launch" Forbes.com, technology. August 31, 2009.

Mas-Colell, A., M.D. Whinston, and J.R. Green. *Microeconomic Theory*. New York: Oxford University Press, 1995.

McAfee, R.P., J. McMillan and M. Whinston (1989): “Multiproduct Monopoly, Commodity Bundling, and Correlation of Values,” Quarterly Journal of Economics, 114, 371-84.

Schmalensee, R. (1984): “Gaussian Demand and Commodity Bundling,” Journal of Business, 57(1), 211-231.

Shapley, L.. “A Value for n-Person Games.” Originally published in 1953. Reprinted in Alvin Roth, The Shapley Value: Essays in Honor of Lloyd S. Shapley, Cambridge University Press: Cambridge, UK, 1988.

Shiller, B and Waldfogel, J. Unpublished manuscript, 2009.

Stigler, G.J. (1963): "United States v. Loew's Inc.: A Note on Block Booking,” Supreme Court Review, 152-7.

Sundararajan, A. (2004) “Nonlinear Pricing of Information Goods.” Management Science 50 (12), 1660-1673.

Venkatesh, R. and V. Mahajan (1993): “A Probabilistic Approach to Pricing a Bundle of Products or Services," Journal of Marketing Research, 30(4), 494-508.

Young, H.P.. “Individual Contribution and Just Compensation,” in Alvin Roth, The Shapley Value: Essays in Honor of Lloyd S. Shapley, Cambridge University Press: Cambridge, UK, 1988.

Table 1

Revenue Attributed to Each Song Under Individual Sales, Shapley Values, and Existing Practical Bundle Revenue Sharing Methods.

					Revenue Distributed With Scheme:		
Song Title - Artist	Uniform Pricing Revenue	Component Pricing Revenue	Shapley Value	Shapley Standard Error	Equal	Proportional	GZ
7 Things - Miley Cyrus	510	511	648	11	1,671	1,182	1,059
American Boy (feat. Kanye West) - Estelle	2,090	2,104	2,722	19	1,671	1,998	2,125
Burnin' Up - Jonas Brothers	620	636	726	12	1,671	1,282	1,185
Can't Believe It (feat. Lil Wayne) - T-Pain	1,333	1,358	1,633	15	1,671	1,809	1,847
Chicken Fried - Zac Brown Band	422	443	491	11	1,671	1,058	944
Circus - Britney Spears	1,040	1,059	1,466	17	1,671	1,544	1,525
Closer – Ne-Yo	1,563	1,593	1,970	19	1,671	1,900	1,949
Corona and Lime - Shwayze	898	924	1,120	14	1,671	1,560	1,540
Dead and Gone (feat. Justin Timberlake) - T.I.	1,285	1,318	1,698	16	1,671	1,824	1,866
Decode - Paramore	538	577	674	12	1,671	1,297	1,179
Disturbia - Rihanna	2,219	2,233	2,829	17	1,671	2,020	2,141
Don't Trust Me - 3OH!3	687	733	893	13	1,671	1,403	1,305
Eye of the Tiger - Survivor	2,128	2,173	2,812	20	1,671	1,893	2,024
Fall for You - Secondhand Serenade	1,112	1,139	1,310	16	1,671	1,599	1,532
Flashing Lights - Kanye West	1,907	1,920	2,335	18	1,671	1,949	2,051
Forever - Chris Brown	2,046	2,094	2,516	17	1,671	1,986	2,052
Gives You Hell - The All-American Rejects	899	927	1,165	14	1,671	1,621	1,592
God Love Her - Toby Keith	422	456	505	10	1,671	1,171	1,055
Gotta Be Somebody - Nickelback	934	954	1,163	15	1,671	1,529	1,473
Heartless - Kanye West	2,147	2,166	2,787	20	1,671	2,014	2,124
Hot N Cold - Katy Perry	1,699	1,712	2,091	19	1,671	1,889	1,948
Human - The Killers	1,394	1,404	2,055	16	1,671	1,726	1,744
I Don't Care (Single Version) - Fall Out Boy	615	700	878	11	1,671	1,462	1,367
I Hate This Part - The Pussycat Dolls	637	682	779	13	1,671	1,368	1,272
I Kissed a Girl - Katy Perry	1,563	1,577	1,982	16	1,671	1,850	1,904

I'm So Paid - Akon, Lil Wayne & Young Jeezy	1,405	1,434	1,811	19	1,671	1,733	1,756
I'm Yours - Jason Mraz	2,320	2,389	2,825	20	1,671	1,981	2,076
If I Were a Boy - Beyonc	1,121	1,162	1,493	16	1,671	1,729	1,748
Just Dance - Lady GaGa & Colby O'Donis	2,256	2,278	2,937	18	1,671	1,988	2,103
Just a Dream - Carrie Underwood	498	561	605	12	1,671	1,343	1,216
Let It Rock - Kevin Rudolf & Lil Wayne	2,037	2,099	2,775	21	1,671	1,898	1,968
Live Your Life (feat. Rihanna) - T.I.	2,634	2,678	3,334	19	1,671	2,092	2,259
Love Lockdown - Kanye West	2,073	2,084	2,665	23	1,671	1,981	2,092
Love Story - Taylor Swift	1,201	1,225	1,575	15	1,671	1,617	1,595
Mad - Ne-Yo	1,066	1,105	1,318	15	1,671	1,716	1,680
Mercy - Duffy	671	689	780	12	1,671	1,413	1,331
Paper Planes - M.I.A.	2,409	2,483	3,456	19	1,671	2,017	2,158
Pen & Paper - The Red Jumpsuit Apparatus	502	592	787	14	1,671	1,378	1,257
Rehab - Rihanna	987	1,077	1,268	13	1,671	1,808	1,812
Right Now (Na Na Na) - Akon	1,756	1,773	2,137	16	1,671	1,917	1,971
Shattered (Turn the Car Around) - O.A.R.	1,141	1,160	1,511	18	1,671	1,664	1,633
Single Ladies (Put a Ring On It) - Beyonc	1,189	1,226	1,569	15	1,671	1,694	1,739
Sober - P!nk	523	612	748	11	1,671	1,450	1,333
Tonight - Jonas Brothers	385	436	449	9	1,671	1,132	1,005
Untouched - The Veronicas	715	753	856	11	1,671	1,385	1,299
Viva la Vida - Coldplay	3,019	3,153	4,103	24	1,671	2,055	2,202
When I Grow Up - The Pussycat Dolls	936	1,015	1,169	16	1,671	1,664	1,632
White Horse - Taylor Swift	806	816	1,068	14	1,671	1,406	1,314
Womanizer - Britney Spears	1,299	1,341	1,736	16	1,671	1,791	1,831
You Found Me - The Fray	1,203	1,203	1,298	14	1,671	1,757	1,731

Table 2a – Performance of Bundle Revenue Sharing Methods – CP Pricing Counterfactual

Revenue Distribution Method	% Above Standalone	Cumul. Side Payment overall	Cumul. Side Payment rel to bundle revenue	Cumul. Side Payment rel to extra revenue
Equal Sharing	70	8277	9.9%	44.3%
Proportional Sharing	76	4012	4.8%	21.5%
Ginsburgh-Zang	80	2638	3.2%	14.1%
Shapley Value	100	0	0.0%	0.0%

Table 2b – Performance of Bundle Revenue Sharing Methods – UP Pricing Counterfactual

Revenue Distribution Method	% Above Standalone	Cumul. Side Payment total	Cumul. Side Payment rel to bundle revenue	Cumul. Side Payment rel to extra revenue
Equal Sharing	70	7678	9.20%	41.10%
Proportional Sharing	76	3457	4.10%	18.50%
Ginsburgh-Zang	82	2114	2.50%	11.30%
Shapley Value	100	0	0.00%	0.00%

Table 3: Comparison of Feasible Schemes with Shapley Values

Revenue Distribution Method	Correlation With Shapley	MAD rel to Shapley Value	% MAPD rel to Shapley Value
Equal Sharing	0	740	63%
Proportional Sharing	0.91	531	42%
Ginsburgh-Zang	0.93	460	36%
Shapley Value	1	0	0%

Figure 1 - Performance of Bundle Revenue Sharing Methods – CP Pricing Counterfactual

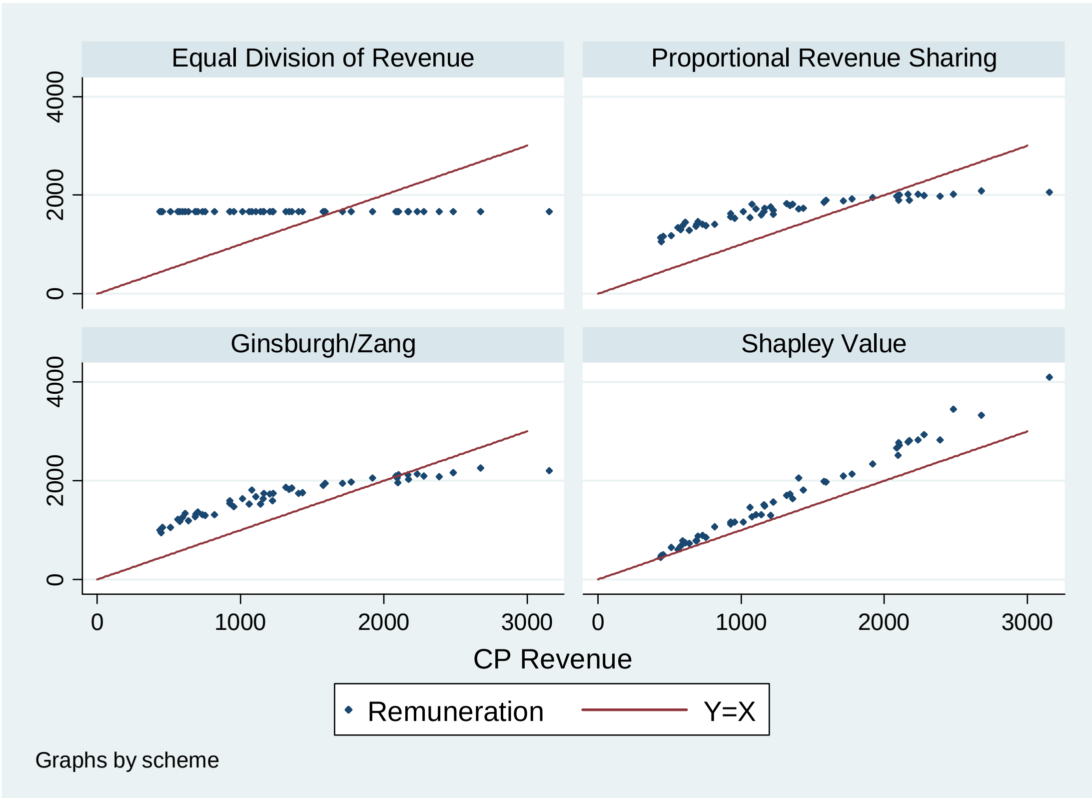

Figure 2 - Performance of Bundle Revenue Sharing Methods – UP Pricing Counterfactual