The Challenge of Revenue Sharing with Bundled Song Pricing

Benjamin Shiller
The Wharton School
University of Pennsylvania

Joel Waldfogel
The Wharton School
University of Pennsylvania and NBER

Preliminary Draft – September 11, 2008

Major online music retailers, including Apple and Nokia, are reported to be contemplating bundled ("all-you-can-eat") song sales, either through one-time fees or ongoing periodic charges. Other services, such as Rhapsody and eMusic, are already selling bundled offerings.[1] The attraction of bundling is easy to see. As a long theoretical literature shows,[2] bundling can increase revenue as long as product valuations are not perfectly positively correlated across consumers. Bundling of 10 or more popular songs has been shown (Shiller and Waldfogel, 2008) to raise revenue by 5-10 percent relative to selling individual songs. But such schemes produce revenue that is not directly attached, or readily attributable, to particular pieces of intellectual property.

Consumers of bundles typically value some products more highly than others, and have zero valuations for some products. In traditional markets without resale, consumers of bundles would acquire all products in the bundle, then dispose of (or just never use) products with zero value to them. In such markets, there is no direct measure of how many times specific products are used, and the revenue division problem requires external data and significant negotiations. In traditional markets, distributing bundle revenue 'fairly' is often prohibitively difficult. In many digital markets, by contrast, consumers of bundles choose which products to acquire when they purchase the bundle. For example, at eMusic, consumers choose each month which songs to download. Consumers have no reason to acquire and later dispose of products for which they assign

[1] 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).

[2] See Stigler, Adams and Yellen, Schmalensee, Armstrong, and Bakos & Brynjolffson.

zero value. Digital markets allow measurement of which bundle components consumers acquire, making some simple usage-based distribution schemes feasible.

The usual solution to the revenue distribution problem in digital markets, used at Rhapsody and eMusic, is to allocate revenue proportionally to consumption, i.e. downloads or plays. If C_j is the number of times that song j is consumed 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 $\frac{R_B C_j}{\sum_{k=1}^{N} C_k}$. While this approach is intuitive, it is not obviously either equitable or sufficiently attractive to all song owners to entice them to participate in the bundle. When songs are sold à la carte, consumers only buy them when they expect to value the songs above the à la carte price (typically around \$1). Hence, only relatively high valuation songs receive revenue.

Under all-you-can-eat pricing, by contrast, all songs with valuations above zero are consumed by individuals purchasing the bundle. Relative to à la carte pricing, bundling with proportional sharing can therefore distribute relatively more revenue to low-valuation songs. Even if the total pie is larger under bundling, high-valuation songs could receive absolutely less revenue. If the highest valuation songs in any bundle prefer standalone revenue to the remuneration available to them in the bundle, a race to the bottom will ensue – leaving only the least popular songs in the bundle.

Far from a fanciful concern, this already seems to be happening. Harding (2007) presents an account of eMusic's experience getting songs onto its site. The site, which specializes in independent record labels and sells songs in bundles, distributes revenue "based on the number of times a song has been downloaded." When artists become relatively popular, the labels are reluctant to make their new songs available on the site. After declining to make one of the label's biggest new records 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." This paper considers this problem as well as some possible solutions.

Assessing the significance of this problem requires information on the distribution of consumers' valuations of songs. While such information is usually unavailable, we have detailed survey based data on nearly 500 college students' valuations of 50 popular songs in early 2008, obtained by direct elicitation. Using these data, we note the pie-expanding effect of sophisticated (all you can eat) pricing schemes that we have documented elsewhere.

We then turn to the main tasks of this paper. First, using our unusually detailed data, we compute some theoretically motivated revenue allocations for each song based on Shapley values and other methods (Brynjolffson and Zang, 2006). Second, we show that the proportional revenue sharing scheme would not yield a stable bundle of the 50 songs. We then compare the proportional revenues to several other methods' implied revenues.

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, averaging over all of its possible arrival orders in the bundle, provides a candidate solution to the revenue-sharing problem. In general, the Shapley value calculation requires information on the revenue available to every conceivable

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

bundle. While this is generally unknowable, our data – each individual's valuation of each song and therefore each bundle of songs – allows us to calculate the maximal revenue available to each possible bundle. Calculation of exact Shapley values for 50 songs is computationally challenging, so we employ random sampling that produces Shapley values to arbitrary levels of accuracy.

Brynjolffson and Zang (2006) propose another bundle revenue distribution scheme, allocating revenue in proportion to each song's contribution to total surplus. Like the Shapley value method, the Brynjolffson and Zang (2006) method is data intensive, although its implementation is straightforward with our valuation data.

Finally, we review various revenue allocation schemes feasible with the data sellers would likely have, including the proportional schemes actually in use, the simplified Shapley values of Ginsburgh and Zang (2003, 2004), and a simple new method that we propose, relying on the empirical relationship between consumption in the bundle and standalone revenue. We evaluate each of these schemes against the theoretical benchmarks of Shapley and BZ.

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 to the present paper (Shiller and Waldfogel, 2008) we use surveys of college students to characterize the distribution of consumers' valuations of 50 songs. We show that the 50 songs can generate 8158 in revenue under uniform pricing, and 8471 under component pricing. The 50-song bundle, by contrast, can garner 8911 in revenue. Furthermore, on average, bundles of as few as 5 (3) songs raise revenue relative to counterfactual profit-maximizing component (uniform) pricing. Bundles of 10 songs, on average, raise revenue by 2.4 (6.1) percent, and the full bundle of all 50 songs raises revenue by 5.2 (9.2) percent, relative to standalone component (uniform) pricing (Tables 1A and 1B).

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 attributed to all songs. As noted earlier, some songs will actually fare worse under bundling than under à la carte pricing under many straightforward revenue sharing schemes.

To see this clearly, suppose a) there are two songs available, b) there are two consumers, c) firm acts optimally in each pricing scheme, and d) bundling revenues are split proportionally to number of times consumed in the bundle. Further, suppose that the consumers' values are as given in the below.

		Product	
		1	**2**
Person	**i**	**$0.50**	**$1.49**
	ii	**$0.49**	**$1.50**

The optimal à la carte uniform price is $1.49, the optimal component prices equal $0.49 and $1.49, for products 1 and 2, respectively, and the optimal bundle price is $1.99.

Under à la carte uniform pricing, product 1 does not sell at all and earns zero profit, whereas product 2 sells twice and earns \$2.98. Under à la carte component pricing, both products sell twice, the first earning \$0.98 and the second again earning \$2.98. Under bundling, both consumers buy the bundle and obtain – and use – both products. If revenue is shared proportionally, each song receives \$1.99. While simple, this example illustrates the main revenue sharing problem that bundling creates, that the higher valued product would earn more revenue outside the bundle, while the lower valued product earns more revenue inside.

This raises the question of which circumstances prompt the proportional scheme to compensate all product owners 'fairly.' Performance inside and outside the bundle can be illustrated with a diagram relating a song's revenue in the bundle (C) to its revenue outside the bundle (R). The 45 degree line represents equal revenue, so a song lies above the 45 degree line if its revenue is greater inside the bundle. If all songs have bundle revenue proportional to standalone revenue – if their consumption in the bundle is somehow proportional to their standalone revenue – then all songs will lie on a ray from the origin that is above the 45 degree line. C need not be proportional to R to guarantee that songs perform better in the bundle than outside. Rather, if the benefits of bundling are large enough, each product could earn more in the bundle, despite unequal gains. The only requirement is that all songs lie above the equal-revenue line.

How can one avoid this problem and distribute revenue 'fairly'? Cooperative game theory provides substantial guidance. 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.

While the Shapley value provides an appealing scheme for sharing revenue, even it has two shortcomings. First, it is not necessarily in the core. And, indeed, in our data, it is not in the core. Second, its calculation requires information that would not, in general, be available.

This leads to the question of what alternative schemes are easily implementable, which depends, in turn, on what is observable. The number of consumption instances of each song is likely to be observable. This makes the proportional method easily implementable. But, as we will see, the proportional method's revenues differ greatly

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

from á la carte revenues, begging the question; are there better revenue sharing rules that can be calculated using only the observable?

Brynjolfsson and Zhang (2006) propose that revenue be divided in proportion to products' value to consumers purchasing the bundle, giving producers efficient incentives for product design. Like Shapley, this approach has rather heavy information requirements, as it requires an estimate of consumers' average valuation of each song. Brynjolfsson and Zhang (2006) also suggest a method for eliciting song valuations. For each song, they propose sending coupons to a statistical sample, redeemable to each recipient if the recipient is willing to give up the song. They point out that their method is intuitive and truth telling is a 'strong and robust' equilibrium. Furthermore, even though the endowment effect may be present, it should affect elicited valuations for all songs. Brynjolfsson and Zhang's coupon method is one obvious method for eliciting valuation data. Here, of course, our song valuation data allow us to calculate consumers' average song valuations directly.[5]

Ginsburgh and Zang (2003, 2004) suggest a simple scheme for sharing revenue using little more than the data required for proportional sharing. They consider the problem of revenue sharing among museums participating in a pass program. A consumer purchases a fixed-price pass that allows him to visit any of *K* museums. They model the situation as a unanimity game in which museums visited by the same individual are assumed equally valuable to that individual. Implementation of this setup requires only the list of products consumed by each bundle participant, and the approach

[5] Fisher (2004) proposes distributing revenue to songs in proportion to the number of times each song is heard. He posits that this method would better correlate with songs' contributions to the bundle than would proportions of downloads. However, Fisher's method may be difficult in practice because some media for distributing music, such as CDs, do not have technology adequate to measure usage.

has the virtue that it leads to a very simple formula for the Shapley value.[6] But the setup departs from ours in the sense that museums are assumed symmetrically valuable within each group's set.

To our knowledge, no other revenue sharing schemes have been proposed in relevant economic literature, and only the proportional revenue sharing rule has been used in practice. The remainder of the paper is structured as follows. We present the survey data in section II. In section III, we evaluate the performance of the revenue distribution schemes with heavy data requirements, the Shapley value and the BZ scheme. In section IV, we evaluate the performance of proportional distribution, the GZ scheme and two others that we propose against the benchmarks of the theoretically motivated and data-intensive Shapley and BZ schemes. A brief conclusion follows.

II. Data

The song valuation data for this paper comes from a survey of 465 Wharton undergraduates, conducted January 16 and 17, 2008. Each student was asked to state the maximum amount they were willing to pay for each of the 50 most popular songs on ITunes that week, after listening to 30 second clips of the songs. Students were motivated to respond to the survey because it counted as a homework assignment. The exact wording of the survey is displayed below.

> 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**

[6] If passholder k visits S museums and pays P for the pass, the Shapley value for museum j is $\sum_{k=1}^{N} \frac{P}{S_k}$.

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 valuation data was merged with self-reported data on level of interest in music and the size of their music library, and limited university demographic data. The final dataset had 23,250 observations.

The data shows some signs of rounding, an issue which will be discussed later, but generally seems reasonable. The fact that 98 percent of valuations are between $0 and $10, and 86% are between $0 and $2, does not seem discordant with our expectations, given market prices. Overall, 31.2% of reported valuations equal zero. See Figure 1, which shows the distribution of valuations between $0 and $2.

Valuations in the data vary substantially across songs and individuals. Table 2 shows the mean, median, 25th and 75th percentile valuations for each song. Mean valuations range from $0.45 to $2.79. However, we find that valuations vary more across individuals than they do across songs. Regressing valuations on song fixed effects yields an R-squared of 4.4 percent, whereas regressing valuations on individual fixed effects yields an R-square of 39.5 percent. A regression on both types of fixed effects yields an R-squared of 43.9 percent. This pattern results in positive correlations between each of the 1225 sets of two songs. The mean correlation is 0.444.

Some in the field have expressed doubts about whether elicited reservation prices approximate true valuations (Diamond and Hausman, 1994). Though, they often focus on obscure public goods, such as views from the Grand Canyon. Rob and Waldfogel (2006), using data similar to the data used in this survey, found that buy-based valuations data, as used in this paper, tend to imply profit-maximizing prices similar to observed prices, while sell-based valuations data do not, likely due to the endowment effect. Furthermore, elicited reservations price data has a history in marketing and operations research. Kalish and Nelson (1991) demonstrated that reservation price data fits models better than other survey response methods, though reservation price data performed somewhat worse than other survey data, but still adequately, in out-of-sample choice predictions. Elicited reservation price data is often used in bundle pricing research (e.g. Hanson and Martin (1990), Venkatesh and Mahajan (1993), and Jedidi, Jagpal, and Manchanda (2003) because using data from other survey methods, or market data, requires either strong parametric assumptions on the distribution of valuations or a vast dataset.

We can further test the data by comparing implied purchases in the data to cumulative market sales using rankings data from Billboard's "Hot Digital" chart, which ranks songs by number of digital purchases. We measure song purchases in the survey data by the number of times each is valued at or above $0.99, the price of a song on ITunes. We do not have a direct measure of cumulative market sales, but the number of weeks a song has been on the Billboard's "Hot Digital" chart, and the highest position achieved on the chart are likely correlated with cumulative MP3 sales.[7] We therefore test the survey data by regressing implied sales in the data on the song's peak rank on the Billboard's "Hot Digital" chart and number of weeks on the chart. The regression yields:

[7] The Billboard Chart, "Hot Digital Songs," is available at http://www.billboard.com/bbcom/charts, accessed March 14, 2008.

$$\underset{(0.052)}{share(V > \$0.99) = 0.352} + \underset{(0.0017)}{0.0033} * (chart\_weeks) - \underset{(0.0015)}{0.0044} * (chart\_peak) .$$

Standard errors are in parentheses. The R-squared from this regression is 0.34. This regression shows that more weeks on the chart and higher rank, rough measures of cumulative sales, positively correlate with implied sales in the data, providing no reason to suspect the survey data's validity.

There is one problematic issue with the survey data - there appears to be a tendency to round stated valuations, often to the nearest 25 cents, as apparent in Figure 1. We show in the Appendix that the measurement error inherent in this observed rounding pattern particularly impacts the relative revenues of bundling and standalone pricing, increasing standalone revenue relative to bundling revenue. In this paper, we add a random uniform [-0.125,0.125] error to nonzero valuations to "unround" the data. This minor assumption only impacts the results in this paper through its effect on the estimated benefits of bundling.

Our analyses require we identify which songs are consumed in the bundle. We assume that an individual consumes a song in the bundle when a) they buy the bundle, and b) their valuation of the song exceeds zero. Adding random error to valuations of zero in the raw data would benefit songs with more true zero valuations, unfairly increasing revenue under the proportional revenue sharing scheme to unpopular songs, favorably biasing our results. As a robustness check, we later try redefining consumption in the bundle by raising the necessary valuation of a song above zero.

Note that, from this point on, we assume that standalone revenue equals component pricing revenue. Findings using uniform pricing revenue as standalone revenue are similar.

Lastly, we calculate in the data the revenue each song receives under uniform, component, and bundle pricing. To find optimal uniform price standalone revenues, we order valuations v_i in decreasing order, define n as the number of valuations greater than v_n, and find $\max_n (nv_n)$. To find component pricing revenues, we repeat the same procedure separately for each song. For bundling revenues, we repeat this procedure after aggregating values by individual. To find revenue under the proportional pricing scheme, we drop all individuals with total valuation for all songs less than the computed optimal bundle price, and amongst the remaining individuals define consumption of song j to be 1 if the individual's valuation for that song is strictly greater than zero, and 0 otherwise.

III. Theoretically Motivated Revenue Sharing Methods with Heavy Data Requirements: Shapley and BZ

Shapley Values

In this section, we analyze the performance of data-intensive revenue sharing methods, beginning with our estimated 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, given current technology.

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-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 $y = 1,\ldots,50$, $y \neq x$. So far, so good. 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 $R_B(\{\text{first } 24, x\}) - R_B(\{\text{first } 24\})$.

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: do 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({*l*-1 random songs sampled from -*i*, *i*})-R_B({*l*-1 random songs sampled from -*i*}).

We can estimate $\Delta R^{l}_{B,j}$ by taking the average of draws of $\Delta \tilde{R}^{l}_{B,j}$. Our estimate based on *T* draws is then: $\Delta \hat{R}^{l}_{B,j}(T) = \sum_{q=1}^{T} \Delta \tilde{R}^{l}_{B,j} \Big/ T$. The Shapley value for song *j*, based on *T* draws, is then estimated as: $\hat{\phi}_j(T) = \sum_{l=1}^{50} \Delta \hat{R}^{l}_{B,j}(T) \Big/ 50$. 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 3 provides estimates of the mean and standard deviation of single-draw Shapley values based on 100 draws. The standard deviations vary between 3.07 and 5.88. Hence, our estimates of the Shapley value based on 100 draws have standard deviations between \$0.31 and \$0.59. Given that the grand mean is \$176, these are rather precise estimates. The averages sum to 8798, which is close to the actual total revenue of the 50-song bundle (\$8804).

We now turn to the question – do Shapley values yield an incentive compatible bundle? While the correlation between Shapley values and standalone revenues is quite high (0.99), some songs do have standalone revenues exceeding their Shapley values. This is illustrated in Figure 2, which plots Shapley values against standalone revenue. We find that the probability that the Shapley value exceeds uniform price standalone revenue is independent of song value. Therefore, neither the most nor least popular songs have stronger incentives to exit the bundle. If it is costly to ascertain one's standalone revenue, then given the small and nonsystematic sometime penalty for participation in the bundle, song owners have little incentive to avoid bundle participation.

Brynjolfsson and Zhang (2006)

The Brynjolfsson and Zhang (2006) method, henceforth the BZ method, was designed to solve a different problem - to provide socially efficient incentives for product investment. Bundle remuneration under the BZ method is proportional to the total welfare the product generates amongst individuals purchasing the full bundle, i.e. marginal producer and consumer surplus. Said another way, bundle remuneration is proportional to the total area under the demand curve (V) for the song in question amongst consumers purchasing the bundle. The exact formula is:

$$R_{B,j}^{BZ} = \left(\frac{V_j}{\sum_{k=1}^{N} V_k} \right) * R_B$$

Where, V_j is the total social value added to the bundle by song j. Derived values are presented in Table 3. Implementing the BZ method requires the same data needed to estimate each song's demand curve – the full distribution of valuations for each song.

BZ allocations, like Shapley values, roughly track standalone revenues, indicating that no song owners lose substantially from participating in a bundle with revenue shared via the BZ scheme (see Figure 2). Not surprisingly, the mean absolute percent difference (MAPD) between Shapley values and BZ revenues is only about 9% - they are quite similar. However, standalone revenue exceeds revenue inside the bundle for the three most popular songs, suggesting that popular songs would opt out of the bundle under BZ. But, given the small sample of songs, it is not clear whether this is a real result or an artifact of sampling error.

IV. Practical Revenue Sharing Methods

Both the Shapley value and BZ method are data intensive, and may thus be infeasible in many situations. In this section, we analyze several simple and intuitive revenue sharing rules with weaker data requirements than Shapley or BZ. We then turn to our proposed method.

We will evaluate each method using two criteria. First, we calculate the similarity between bundle remuneration and benchmark revenues, where benchmark revenues are either Shapley values or BZ method revenues, by measuring the mean absolute percent difference (MAPD) between implied song revenues and benchmark revenues. Second, we look at whether songs benefit from joining the bundle. If many songs, or the most popular songs, lose from bundle participation, then a revenue distribution scheme would be likely to unravel.

Basic Proportional and Two-Part Proportional.

The proportional revenue sharing scheme is intuitive - and is in use - because it requires only information on the number of times each song was downloaded as part of the bundle. However, we find that it performs poorly. The MAPD between attributed revenues under the proportional scheme and Shapley values (BZ method revenues) is 30.8% (23.4), higher than the MAPDs between any other method and benchmark revenues (see Table 5). Not surprisingly, we find that many songs are not in the core.

In general, unpopular songs gain more than popular songs by joining the bundle. Despite the pie-expanding effects of bundling, the most popular 13 songs out of 50 receive less revenue in the bundle than on their own. To show this, we simply plot the revenue in the bundle vs. standalone revenue for each of the songs (Figure 3). This graph shows that the heterogeneity in revenue between songs under the proportional scheme is much less than the heterogeneity in standalone revenues between songs; the proportional scheme revenues are similar for all songs, causing popular songs to earn more outside the bundle. In fact, the least popular songs earn more than twenty percent as much in the bundle as they do on their own, while the most popular songs make about half as much in the bundle as they would on their own.

It is conceivable that unpopular songs do disproportionately well because a large number of consumers have a very low valuation for them. If so, then the proportional scheme could work if the marginal cost were raised above zero (i.e. two part pricing). Individuals would then download a song only if they purchased the bundle and valued the song above the marginal price. To find remuneration under two-part pricing, we begin by computing the optimal fixed fee for various marginal prices by grid search. Then, using these tariff schedules, we calculate, for each song, the number of consumption instances in the bundle, and attribute revenue to songs in proportion to consumption in the two-part tariff regime.

We find that, for various marginal prices, this revenue distribution scheme also performs poorly. While the MAPD between bundle remuneration and benchmark revenues initially decreases as the marginal price increases, it remains large relative to some other mechanisms, and eventually increases (see table 6). None of the two-part tariffs attempted here yield a MAPD between 2-Part tariff revenue and Shapley values (BZ method revenues) lower than 24% (17.9%). Similarly, the two-part tariff schemes do not improve incentive compatibility of the bundle. Figure 4 plots the payments in the bundle against standalone revenues. It is apparent in the figure, that when charging a marginal fee of less than or equal to 50 cents, along with the respective optimal fixed fee, revenue for each song is nearly identical to its revenue in the basic (i.e. 0 MC) proportional revenue sharing scheme. Raising the fee higher to $1.00 does not help - the revenue in the bundle still fails to increase much with standalone revenue, i.e. the line is still quite flat, and the size of the pie shrinks substantially. Popular songs still earn more revenue outside the bundle under every two-part tariff regime.

These data suggest that a bundle of these 50 songs would not be stable under the basic proportional method, or two-part proportional schemes. Moreover, one cannot create a stable bundle by merely bundling together only the songs among the 50 benefiting from bundle participation. When the most popular songs leave, the most popular of the remaining songs no longer find the bundle attractive, and so on. This would lead to a race to the bottom, resulting in a bundle of only the lowest quality songs.

Ginsburgh and Zang

The proportional revenue sharing scheme performs poorly in this context. Can other methods proposed in economic literature do better? The Ginsburgh and Zang (2003, 2004) method (henceforth GZ), like the proportional revenue sharing method, is appealing and easy to implement. The only data required for implementation is the list of songs consumed by each bundle buyer. 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. Intuitively, each song's remuneration in the bundle is calculated as follows: each bundle participant pays P_b for participation. If he consumes 5 songs, his payment is divided into 5 equal parts for those 5 songs. The payments to each song are then summed across participants to yield total payments to each song.

In our data, we find that the GZ method also performs poorly according to both criteria. The MAPD between the revenues in the GZ method and Shapley values (BZ method revenues) is 27.0% (19.6%). An absolute improvement of 3.8% MAPD (approx. 13 percent decrease) over the proportional method is notable, but small given the improvement possible through other methods (discussed later). We again find that many popular songs earn more standalone revenue than remuneration in the bundle. See Figure 3, which plots the revenue from this scheme, as well as other schemes, against standalone revenue. Under the GZ method, the 11 most popular out of 50 songs earn more revenue outside the bundle than in, again a slight improvement over the proportional method.

The fact that the GZ method improves on the proportional scheme only slightly is not surprising, because the GZ method uses a strong assumption to enable calculation of Shapley values using very limited data - only data on consumption in the bundle by individual. We find that methods requiring data on individual valuations or standalone revenues perform far better.

Empirically Based Allocation

We now turn to our proposed revenue distribution methods. Our first approach is simple. To ensure that the bundle is incentive compatible, make remuneration inside the bundle equal to or greater than standalone revenue. The most straightforward way to accomplish this is to compensate songs in proportion to their standalone revenues. The exact formula is thus:

$$R_{B,j}^{*} \frac{R_j}{\sum_{k=1}^{N} R_k} * R_B$$

Because bundle revenue exceeds the sum of standalone revenues, this scheme makes every song better off. Not surprisingly, the MAPD between the revenue under this method and Shapley values is only 6.9%. Yet, standalone revenues are not in general easily known, since they require the distribution of valuations.

Our second proposed approach attempts to use what we can easily observe, the number of times each bundle element is downloaded (C), to estimate what we would like to but cannot observe, standalone revenue (R). Of course, given our data, we can observe both R and C for 50 songs. If, based on this set of songs, the relationship between R and C appears stable, then we can use it to guide an empirically based revenue distribution scheme. We refer to this as the "proportional to outside revenue" (POR) approach.

Once we have estimated R (based on C), we propose to distribute revenue according to:

$$R_{B,j}^{*} = \frac{E(R_j \mid C_j)}{\sum_{k=1}^{N} E(R_k \mid C_k)} * R_B$$

We propose the following parameterization of the relationship between standalone revenue and consumption in the bundle[8]:

---

[8] We also tried running weighted local linear regressions, but found that they performed no better, and cannot be used in a sensible and transparent way to predict revenues for songs outside of the range of the calibration data.

$R_j = \exp\left(\alpha + \beta_1 C_j + \beta_2 C_j^2 + \varepsilon_j\right).$

Taking logs of both sides yields a straightforward linear regression, which yields an R-Square of 0.84. This functional form fits the data well, suggesting the promise of the scheme. Figure 5 plots the data and estimated function on a graph of standalone revenue against consumption.

We find that this revenue distribution method performs fairly well according to our criteria. The MAPD between its distributed revenues and Shapley values is 15.8%, much lower than the MAPD of other easily implementable methods. This method, like the BZ method, nearly yields an incentive compatible bundle (see Figure 3). Only the 3 most popular songs earn slightly more revenue outside the bundle. This problem could likely be remedied, with more data, by fitting a more flexible parametric or a non-parametric model.

Performance in-sample, though, is only of minor interest. This method should be evaluated on its out-of-sample performance. Via our proposed method, we find that estimated revenues can be estimated reasonably accurately using a subset of the 50 songs to fit the function. In fact, the estimated revenues using only 10 randomly chosen songs to estimate the function are very similar to the estimated revenues using the full sample of 50 songs. To show this, we drew 25 sets of 10, 20, 30, and 40 songs, 100 in total, and for each set fit the function of standalone revenue on bundle consumption (in the full bundle), by running the aforementioned regression. We then used the results to predict standalone revenues and corresponding payments in the bundle, for each song/iteration in and out of the estimation sample. We found that the mean of out-of-sample mean absolute percent errors (mean MAPEs) across the 25 iterations of 10 song draws was

16.0%, and the inter-quartile range was 14.8% to 17.0% - only slightly higher than the in-sample MAPE of 13.1% using the full sample (Table 7).

The magnitude of average error may be less important than the relationship between error and song popularity. Figure 6 plots the predicted revenue functions from 25 randomly drawn sets of 10 songs. The predictions are often less accurate for very popular and unpopular songs. Prediction accuracy can be improved by increasing the number of songs used in the estimation (Figure 7), or oversampling very popular and unpopular songs. To show the latter, we again randomly drew 25 sets of 10 songs each for fitting the function, but this time forcing that three songs be drawn from the top ten standalone revenue songs, three be drawn from the bottom ten, and the remaining four drawn from the middle 30 standalone revenue songs. Figure 8 graphs the predictions in the over-sampling the extremes model. The predicted out-of-sample fit is much better when over-sampling popular and unpopular songs.

Conclusion:

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

We show that the proportional revenue sharing schemes in use, by penalizing popular songs, will tend to induce a "race to the bottom" in which only low valuation songs will benefit from joining bundles. While various theoretically motivated schemes can guide revenue sharing in principle, their data requirements make them difficult to implement. We document a fairly clear relationship between an easily observable datum, the number of bundle buyers taking each song, and the songs' standalone revenue. If stable, this relationship could guide division of bundle revenue in a way that does not discourage bundle participation by popular songs. We implement this method in our data. It is premature to suggest that we have solved the problem of revenue division. But if subsequent research finds the relationship we document here elsewhere as well, then this simple method may facilitate the use of bundling in music sales.

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.

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, 2008.

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
Benefits of Bundling, as Bundle Size Increases

A

	Uniform			*Pure Bundling*			
# Products in Bundle	Revenue	Quantity of Songs	Price per song	Revenue	Revenue relative to UP	Quantity of Songs	Average Price per Song
2	332.9	235.5	1.52	332.9	0.0%	298.6	1.18
3	502.9	332.4	1.60	507.2	0.9%	504.5	1.05
4	665.6	426.6	1.65	683.3	2.7%	726.6	0.97
5	831.7	516.8	1.69	861.5	3.6%	959.0	0.92
10	1636.5	931.4	1.80	1736.9	6.1%	2157.8	0.82
25	4082.6	2183.9	1.87	4418.0	8.2%	5937.3	0.76
50	8158.9	4351.0	1.88	8911.9	9.2%	12350.0	0.72

B

	Component			*Pure Bundling*			
# Products in Bundle	Revenue	Quantity of Songs	Average Price per Song	Revenue	Revenue relative to CP	Quantity of Songs	Average Price per Song
2	339.4	226.2	1.59	332.9	-1.9%	298.6	1.18
3	515.1	332.8	1.59	507.2	-1.5%	504.5	1.05
4	684.9	437.1	1.60	683.3	-0.2%	726.6	0.97
5	857.2	554.0	1.58	861.5	0.5%	959.0	0.92
10	1695.4	1083.5	1.57	1736.9	2.4%	2157.8	0.82
25	4237.6	2738.1	1.55	4418.0	4.3%	5937.3	0.76
50	8471.3	5462.0	1.55	8911.9	5.2%	12350.0	0.72

Table 2
Survey Songs and their Valuations

Question	Mean	25th Percentil	Median	75th Perctile
Apologize (feat. OneRepublic) – Timbaland	$2.36	$0.59	$1.39	$2.67
Big Girls Don't Cry (Personal) – Fergie	$1.15	$0.00	$0.53	$1.22
Bubbly - Colbie Caillat	$1.47	$0.00	$0.68	$1.73
Clumsy – Fergie	$0.77	$0.00	$0.29	$1.01
Crank That (Soulja Boy) - Soulja Boy Tell 'Em	$2.00	$0.28	$1.01	$2.10
Crushcrushcrush – Paramore	$0.56	$0.00	$0.13	$0.71
Cyclone (feat. T-Pain) - Baby Bash	$1.28	$0.00	$0.56	$1.45
Don't Stop the Music – Rihanna	$1.39	$0.10	$0.63	$1.44
Feedback – Janet	$0.61	$0.00	$0.04	$0.57
Hate That I Love You (feat. Ne-Yo) – Rihanna	$1.29	$0.03	$0.55	$1.47
Hero/Heroine (Tom Lord-Alge Mix) - Boys Like Girls	$0.76	$0.00	$0.26	$1.00
Hey There Delilah - Plain White T's	$2.01	$0.15	$0.94	$2.02
How Far We've Come - Matchbox Twenty	$1.40	$0.04	$0.69	$1.47
Hypnotized (feat. Akon) – Plies	$1.14	$0.00	$0.48	$1.12
I Don't Wanna Be In Love (Dance Floor Anthem) – Good Charlotte	$1.06	$0.00	$0.47	$1.20
Into the Night (feat. Chad Kroeger) – Santana	$1.48	$0.00	$0.71	$1.53
Kiss Kiss (feat. T-Pain) - Chris Brown	$1.44	$0.10	$0.85	$1.70
Love Like This - Natasha Bedingfield	$1.04	$0.00	$0.43	$1.06
Love Song - Sara Bareilles	$1.01	$0.00	$0.37	$1.07
Low (feat. T-Pain) - Flo Rida	$1.60	$0.08	$0.88	$1.93
Misery Business – Paramore	$0.67	$0.00	$0.17	$0.90
No One - Alicia Keys	$1.58	$0.13	$0.83	$1.86
Our Song - Taylor Swift	$0.79	$0.00	$0.11	$0.80
Over You – Daughtry	$1.21	$0.00	$0.47	$1.12
Paralyzer - Finger Eleven	$1.10	$0.00	$0.34	$1.17
Piece of Me - Britney Spears	$0.75	$0.00	$0.03	$0.85
Ready, Set, Don't Go - Billy Ray Cyrus feat. Miley Cyrus	$0.58	$0.00	$0.00	$0.58
Rockstar – Nickelback	$1.39	$0.00	$0.50	$1.47
S.O.S. - Jonas Brothers	$0.66	$0.00	$0.15	$0.76
See You Again - Miley Cyrus	$0.67	$0.00	$0.00	$0.59
Sensual Seduction (Edited) - Snoop Dogg	$1.17	$0.00	$0.29	$1.07
Shadow of the Day - Linkin Park	$1.23	$0.00	$0.52	$1.23
Sorry – Buckcherry	$0.63	$0.00	$0.13	$0.76
Start All Over - Miley Cyrus	$0.45	$0.00	$0.00	$0.32
Stay – Sugarland	$0.62	$0.00	$0.00	$0.59
Stop and Stare – OneRepublic	$1.04	$0.00	$0.44	$1.10
Stronger - Kanye West	$2.79	$0.87	$1.74	$3.04
Sweetest Girl (Dollar Bill) [feat. Akon, Lil Wayne & Niia] - Wyclef Jean	$1.78	$0.14	$0.88	$1.98
Take You There - Sean Kingston	$1.36	$0.13	$0.78	$1.58
Tattoo - Jordin Sparks	$0.93	$0.00	$0.39	$1.00
Teardrops On My Guitar - Taylor Swift	$0.90	$0.00	$0.17	$0.93
The Great Escape - Boys Like Girls	$1.10	$0.00	$0.44	$1.25
The Way I Am - Ingrid Michaelson	$0.90	$0.00	$0.26	$0.97
The Way I Are (feat. Keri Hilson & D.O.E.) – Timbaland	$2.24	$0.42	$1.13	$2.61
Through the Fire and Flames – Dragonforce	$0.72	$0.00	$0.04	$0.90
Wake Up Call - Maroon 5	$1.54	$0.17	$0.87	$1.92
When You Were Young - The Killers	$1.61	$0.17	$0.90	$1.98

Witch Doctor - Alvin and the Chipmunks	$0.67	$0.00	$0.00	$0.43
With You - Chris Brown	$1.33	$0.00	$0.49	$1.14
Won't Go Home Without You - Maroon 5	$1.43	$0.17	$0.86	$1.57

Table 3
Remuneration under Theoretically Motivated Schemes - Estimated BZ Revenues and Shapley Values (Based on 100 Draws)

Song	*Standalone Revenues*	*BZ Revenues*	*Shapley values: Mean from 100 one-draws*	*Shapley values: Stdev*
Apologize (feat. OneRepublic) – Timbaland	362.46	336.25	380.89	4.69
Big Girls Don't Cry (Personal) – Fergie	162.09	168.44	166.45	4.26
Bubbly - Colbie Caillat	204.39	217.64	216.02	4.9
Clumsy – Fergie	123.85	112.07	127.89	4.22
Crank That (Soulja Boy) - Soulja Boy Tell 'Em	277.92	293.04	298.76	5
Crushcrushcrush – Paramore	89.74	81.66	101.81	4.03
Cyclone (feat. T-Pain) - Baby Bash	183.45	192.46	205.21	4.98
Don't Stop the Music – Rihanna	181.23	209.98	194.66	4.68
Feedback – Janet	89.11	94.72	79.96	3.3
Hate That I Love You (feat. Ne-Yo) – Rihanna	178.41	197.59	183.81	5.06
Hero/Heroine (Tom Lord-Alge Mix) - Boys Like Girls	118.8	114.04	109.09	3.68
Hey There Delilah - Plain White T's	283.31	292.49	279.75	4.86
How Far We've Come - Matchbox Twenty	186.35	201.15	192.91	4.73
Hypnotized (feat. Akon) – Plies	157.21	171.53	160.45	4.12
I Don't Wanna Be In Love (Dance Floor Anthem) - Good Charlotte	161.32	156.15	166.86	4.32
Into the Night (feat. Chad Kroeger) – Santana	188.89	216.35	203.08	4.32
Kiss Kiss (feat. T-Pain) - Chris Brown	204.46	205.27	250.04	4.96
Love Like This - Natasha Bedingfield	144.43	160.36	150.02	4.88
Love Song - Sara Bareilles	140.94	146.63	144.36	3.76
Low (feat. T-Pain) - Flo Rida	236.58	233.69	261.33	5.28
Misery Business – Paramore	106.34	99.32	120.74	3.64
No One - Alicia Keys	215.24	233.46	217.21	5.16
Our Song - Taylor Swift	108.83	123.08	101.16	3.93
Over You – Daughtry	158.36	185.01	145.03	3.87
Paralyzer - Finger Eleven	157.91	150.34	175.63	4.63
Piece of Me - Britney Spears	103.65	110.48	114.94	3.9
Ready, Set, Don't Go - Billy Ray Cyrus feat. Miley Cyrus	76.9	88.23	75.33	4.03
Rockstar – Nickelback	189.87	210.01	182.55	4.66
S.O.S. - Jonas Brothers	93.82	102.47	91.07	3.32
See You Again - Miley Cyrus	83.15	107.88	73.78	3.66
Sensual Seduction (Edited) - Snoop Dogg	165.97	169.09	160.27	4.37
Shadow of the Day - Linkin Park	160.03	176.53	191.82	4.54
Sorry – Buckcherry	92.53	92.08	96.92	3.68
Start All Over - Miley Cyrus	59.68	73.47	56.9	3.07
Stay – Sugarland	89.46	94.81	71.43	3.26
Stop and Stare – OneRepublic	147.7	156.53	152.78	3.8
Stronger - Kanye West	413.03	390.97	449.31	5.49
Sweetest Girl (Dollar Bill) [feat. Akon, Lil Wayne & Niia] - Wyclef Jean	256.43	263.37	263.52	4.76
Take You There - Sean Kingston	197.28	201.94	204.43	4.1
Tattoo - Jordin Sparks	129.7	141.90	135.24	4.11

Teardrops On My Guitar - Taylor Swift	121.58	140.60	114.9	3.88
The Great Escape - Boys Like Girls	157.6	164.93	146.72	4.17
The Way I Am - Ingrid Michaelson	130.76	131.53	133.34	4.09
The Way I Are (feat. Keri Hilson & D.O.E.) - Timbaland	340.91	326.27	372.39	5.13
Through the Fire and Flames – Dragonforce	112.99	103.72	84.36	3.27
Wake Up Call - Maroon 5	232.28	221.80	252.14	5.07
When You Were Young - The Killers	246.02	225.98	261.01	4.63
Witch Doctor - Alvin and the Chipmunks	87.86	106.65	80.1	3.56
With You - Chris Brown	160.45	208.31	186.04	4.8
Won't Go Home Without You - Maroon 5	199.85	200.78	213.24	4.79

Table 4
Alternative Revenue Sharing Schemes

Song	Uniform	Component	Proportional	GZ
Apologize (feat. OneRepublic) – Timbaland	$361.89	$362.46	$214.71	$226.96
Big Girls Don't Cry (Personal) – Fergie	$157.51	$162.09	$185.52	$191.09
Bubbly - Colbie Caillat	$193.13	$204.39	$182.69	$183.02
Clumsy – Fergie	$91.88	$123.85	$172.33	$171.13
Crank That (Soulja Boy) - Soulja Boy Tell 'Em	$277.51	$277.92	$208.12	$219.18
Crushcrushcrush – Paramore	$69.38	$89.74	$148.79	$140.51
Cyclone (feat. T-Pain) - Baby Bash	$181.88	$183.45	$190.22	$195.35
Don't Stop the Music – Rihanna	$180.01	$181.23	$208.12	$219.30
Feedback – Janet	$88.13	$89.11	$146.91	$138.51
Hate That I Love You (feat. Ne-Yo) – Rihanna	$176.26	$178.41	$196.82	$205.73
Hero/Heroine (Tom Lord-Alge Mix) - Boys Like Girls	$103.13	$118.80	$168.56	$161.49
Hey There Delilah - Plain White T's	$266.26	$283.31	$194.93	$203.37
How Far We've Come - Matchbox Twenty	$176.26	$186.35	$192.11	$194.12
Hypnotized (feat. Akon) – Plies	$144.38	$157.21	$177.98	$176.38
I Don't Wanna Be In Love (Dance Floor Anthem) - Good Charlotte	$150.01	$161.32	$178.92	$174.41
Into the Night (feat. Chad Kroeger) – Santana	$185.63	$188.89	$186.46	$192.06
Kiss Kiss (feat. T-Pain) - Chris Brown	$202.51	$204.46	$196.82	$209.64
Love Like This - Natasha Bedingfield	$144.38	$144.43	$185.52	$182.18
Love Song - Sara Bareilles	$135.01	$140.94	$177.98	$175.62
Low (feat. T-Pain) - Flo Rida	$234.38	$236.58	$198.70	$205.21
Misery Business – Paramore	$88.13	$106.34	$155.38	$147.84
No One - Alicia Keys	$213.76	$215.24	$196.82	$209.98
Our Song - Taylor Swift	$105.00	$108.83	$152.56	$144.18
Over You – Daughtry	$144.38	$158.36	$182.69	$176.63
Paralyzer - Finger Eleven	$146.26	$157.91	$160.09	$154.90
Piece of Me - Britney Spears	$101.25	$103.65	$136.55	$141.31
Ready, Set, Don't Go - Billy Ray Cyrus feat. Miley Cyrus	$61.88	$76.90	$129.01	$118.66
Rockstar – Nickelback	$187.51	$189.87	$185.52	$198.11
S.O.S. - Jonas Brothers	$82.50	$93.82	$159.15	$148.73
See You Again - Miley Cyrus	$75.00	$83.15	$134.66	$124.78
Sensual Seduction (Edited) - Snoop Dogg	$163.13	$165.97	$161.03	$164.47
Shadow of the Day - Linkin Park	$153.76	$160.03	$186.46	$186.51
Sorry – Buckcherry	$76.88	$92.53	$144.08	$128.81
Start All Over - Miley Cyrus	$52.50	$59.68	$117.71	$104.86
Stay – Sugarland	$82.50	$89.46	$136.55	$124.01
Stop and Stare – OneRepublic	$144.38	$147.70	$181.75	$177.72
Stronger - Kanye West	$412.52	$413.03	$216.59	$228.74
Sweetest Girl (Dollar Bill) [feat. Akon, Lil Wayne & Niia] – Wyclef Jean	$255.01	$256.43	$200.58	$207.30
Take You There - Sean Kingston	$196.88	$197.28	$207.17	$214.76
Tattoo - Jordin Sparks	$129.38	$129.70	$186.46	$182.55
Teardrops On My Guitar - Taylor Swift	$118.13	$121.58	$161.03	$152.40
The Great Escape - Boys Like Girls	$157.51	$157.60	$186.46	$182.77
The Way I Am - Ingrid Michaelson	$129.38	$130.76	$164.80	$155.39
The Way I Are (feat. Keri Hilson & D.O.E.) – Timbaland	$339.39	$340.91	$211.88	$220.72
Through the Fire and Flames – Dragonforce	$106.88	$112.99	$137.49	$128.69

Wake Up Call - Maroon 5	$230.63	$232.28	$203.41	$204.89
When You Were Young - The Killers	$245.64	$246.02	$189.28	$197.22
Witch Doctor - Alvin and the Chipmunks	$82.50	$87.86	$111.12	$103.92
With You - Chris Brown	$157.51	$160.45	$195.87	$204.97
Won't Go Home Without You - Maroon 5	$198.76	$199.85	$198.70	$201.92

Table 5
Mean Absolute Percent Difference Between Each Methods Revenues and Benchmark Revenues

	Less Data Intensive	-------------------------------------->			More Data Intensive
Benchmark	*Proportional*	*2-Part*†	*GZ*	*POR Estimated*	*POR Exact*
Shapley Value	30.8%	24.0%	27.0%	15.9%	6.9%
BZ	23.4%	17.9%	19.6%	12.7%	6.9%

†We report here the lowest MAPD of any of the 2-part tariff schemes tried. The corresponding 2-Part tariff's marginal fee was $0.50

Table 6
Mean Absolute Percent Difference Between Each 2-Part Tariff Methods Revenues and Benchmark Revenues

	Marginal Cost							
Benchmark	*$0.05*	*$0.10*	*$0.15*	*$0.20*	*$0.25*	*$0.50*	*$0.75*	*$1.00*
Shapley Value	30.1%	30.1%	29.4%	28.2%	27.3%	24.0%	24.7%	27.4%
BZ	22.7%	23.0%	22.3%	21.7%	20.8%	17.9%	20.0%	24.5%

Table 7
Distribution of MAPEs across Draws

Number Songs Used to Calibrate Function	Mean Out-of-Sample MAPE	25^{th} Percentile Out-of-Sample MAPE	Median Out-of-Sample MAPE	75th Percentile Out-of-Sample MAPE
10	16.0	14.8	15.1	17.0
20	15.3	14.1	14.9	15.9
30	14.7	13.1	14.9	15.8
40	14.6	11.5	13.9	16.7
50 (In Sample MAPE)	13.1	13.1	13.1	13.1

Figure 1

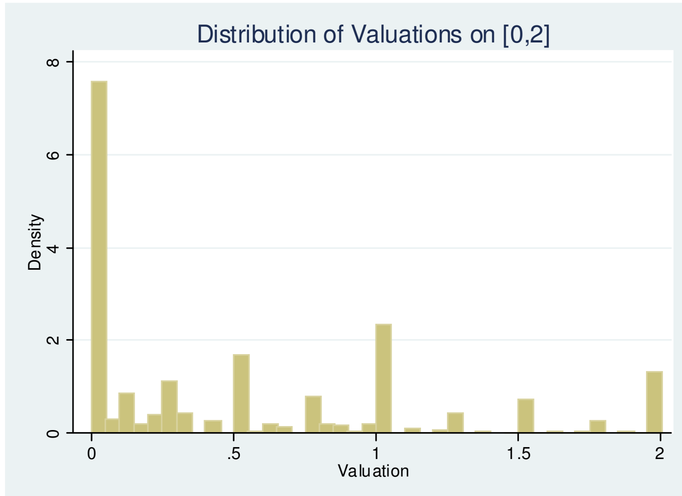


Figure 2

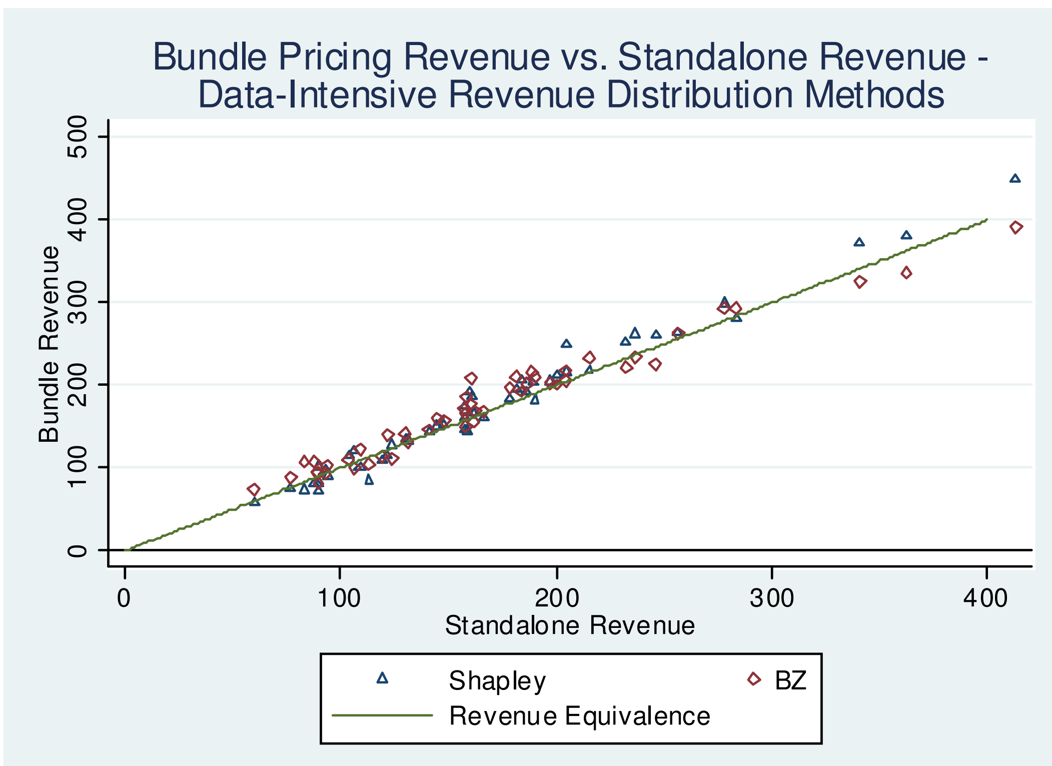


Figure 3

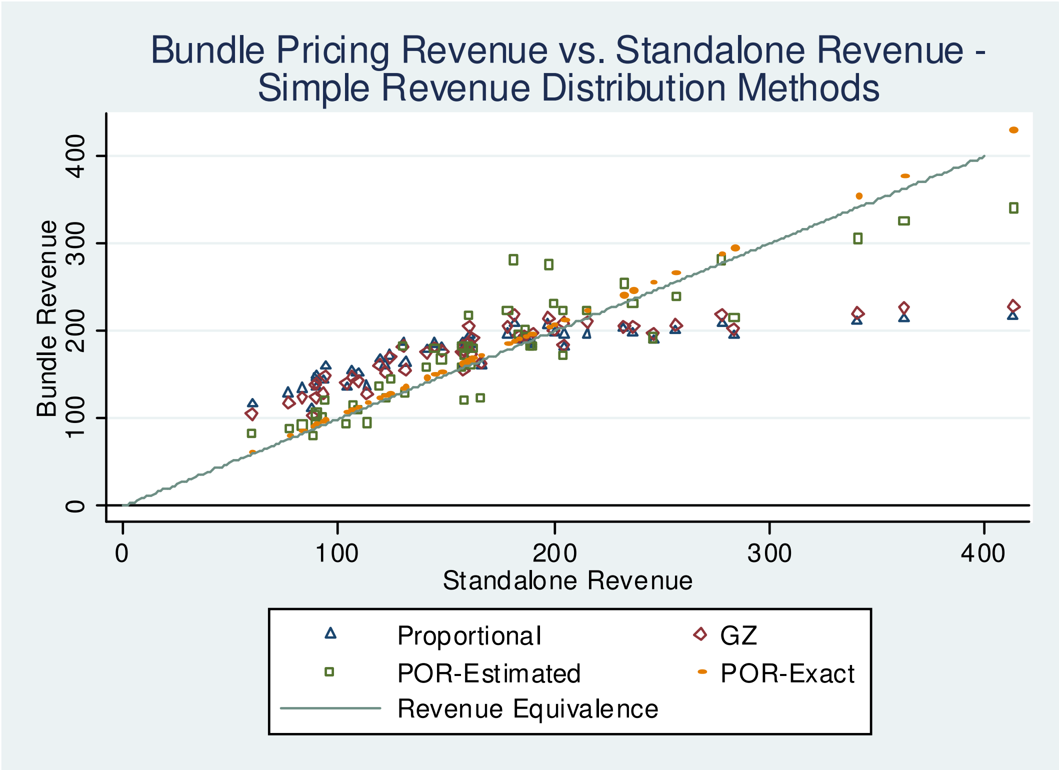


Figure 4

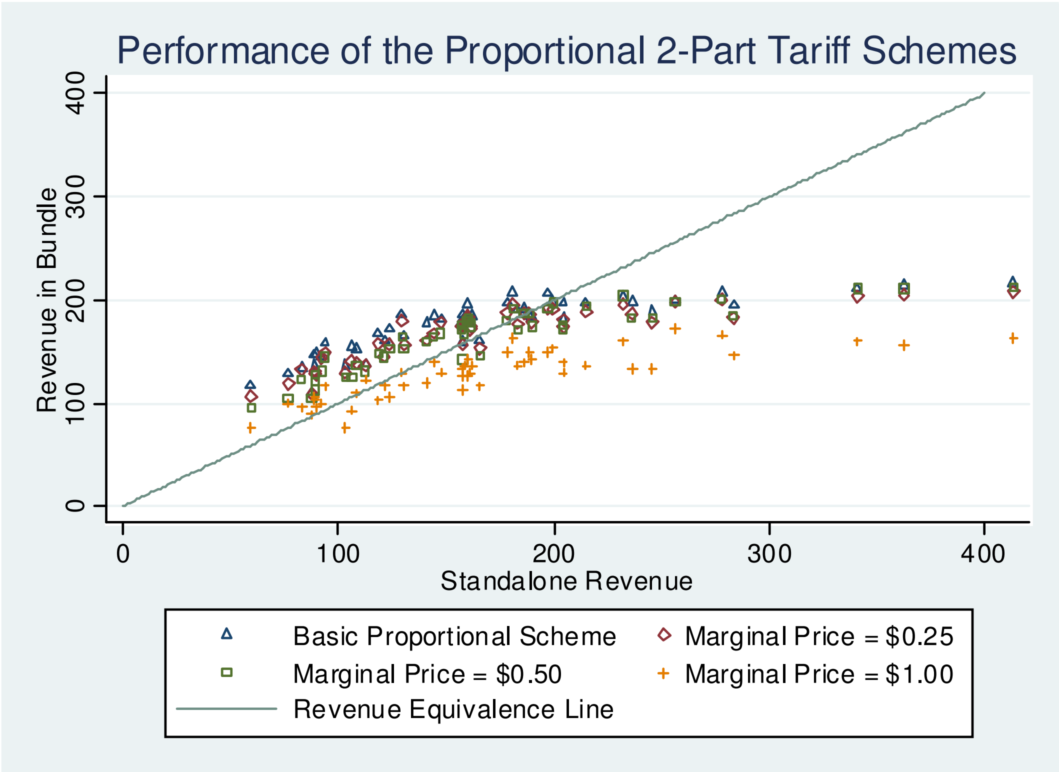


Figure 5

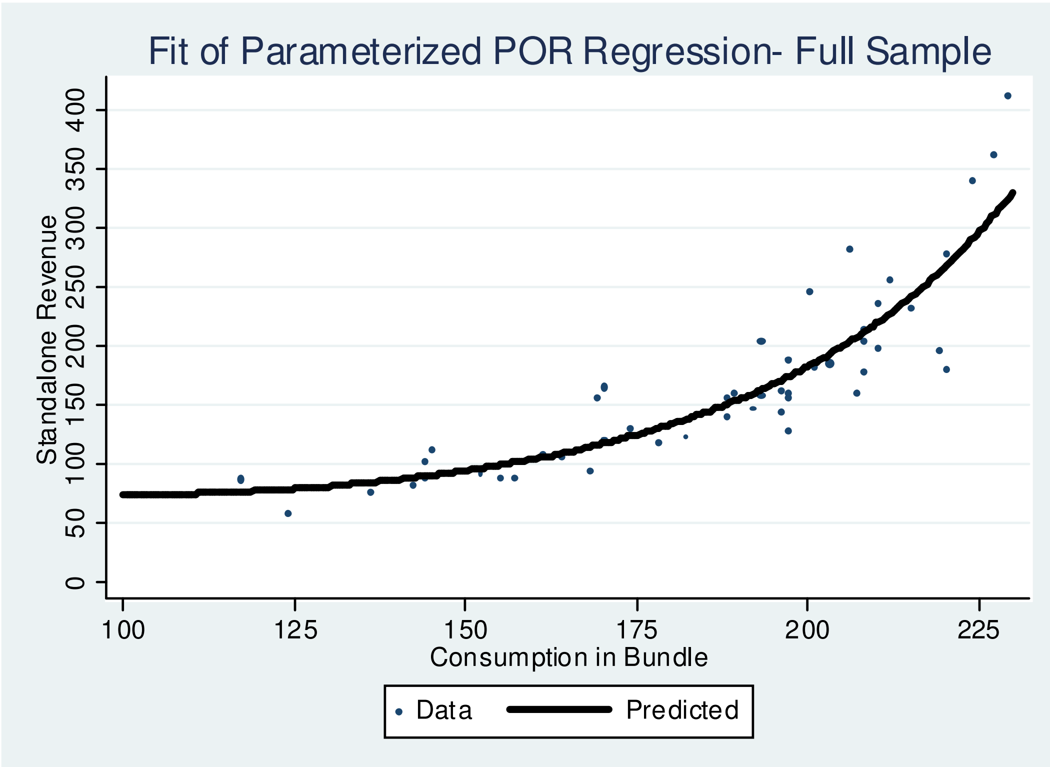


Figure 6

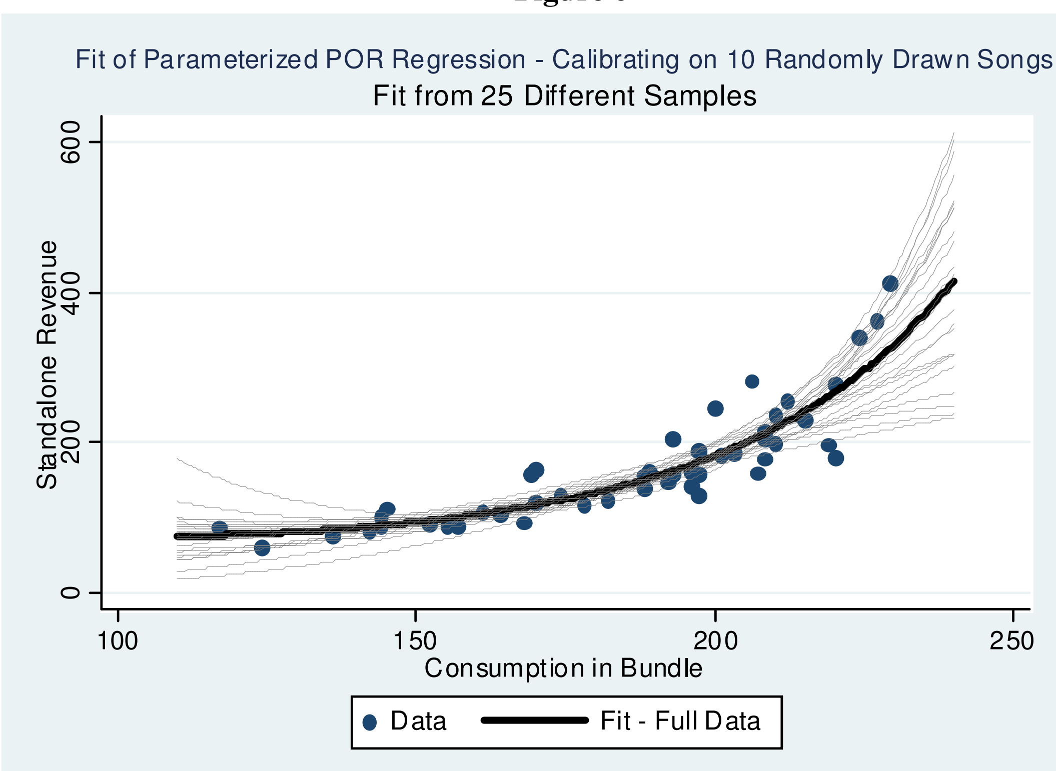
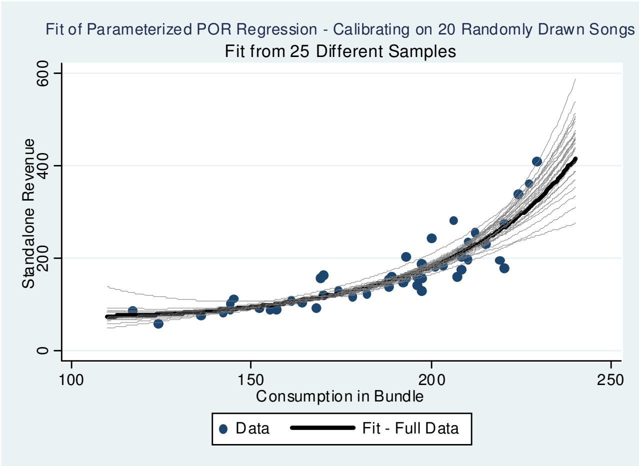
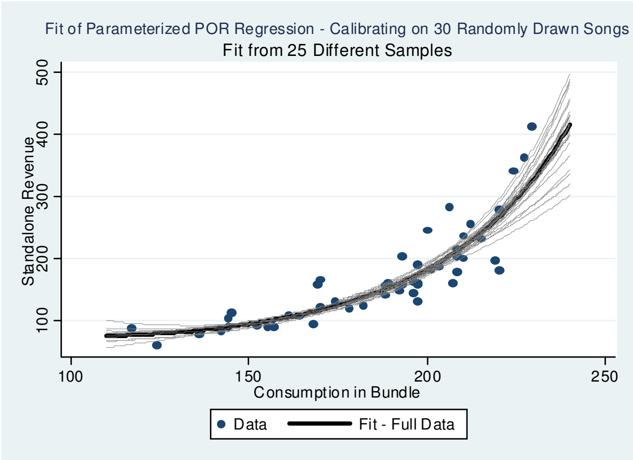


Figure 7

A


Fit of Parameterized POR Regression - Calibrating on 20 Randomly Drawn Songs
Fit from 25 Different Samples
Standalone Revenue
0
200
400
600
100
150
200
250
Consumption in Bundle
Data
Fit - Full Data

B


Fit of Parameterized POR Regression - Calibrating on 30 Randomly Drawn Songs
Fit from 25 Different Samples
Standalone Revenue
100
200
300
400
500
100
150
200
250
Consumption in Bundle
Data
Fit - Full Data

C

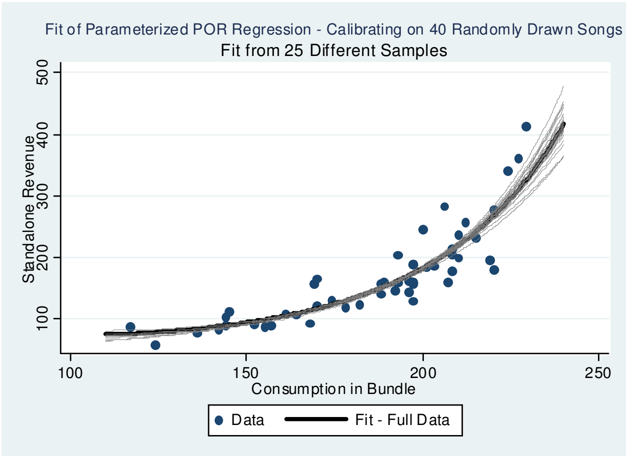
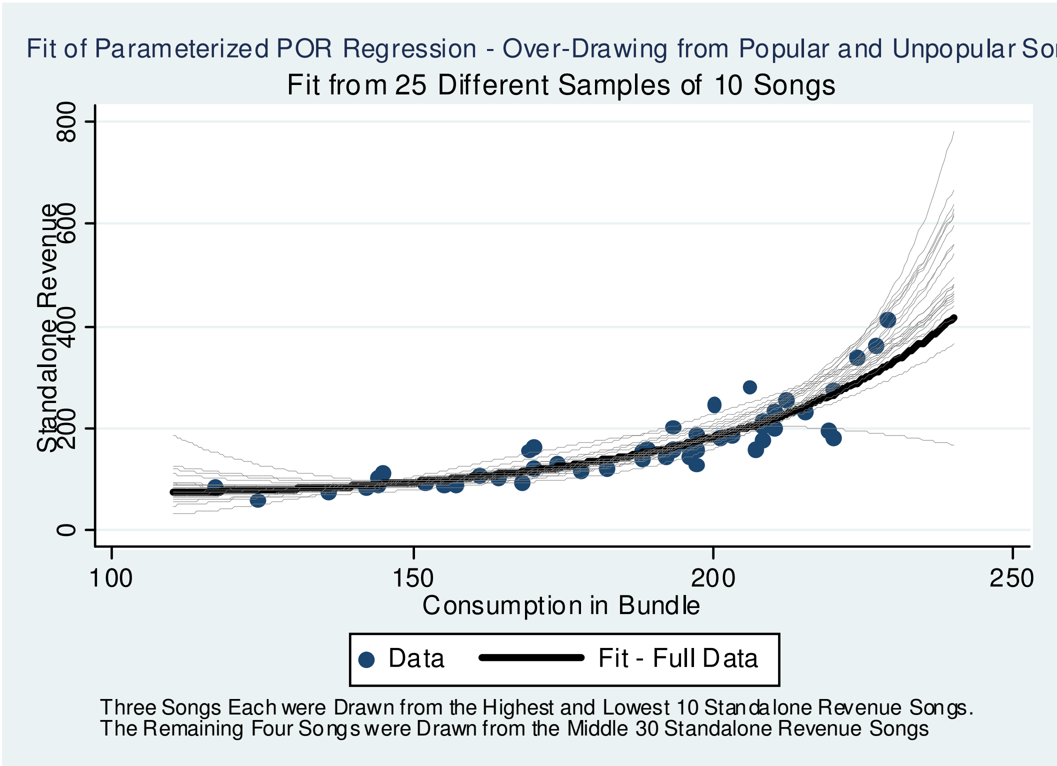

Fit of Parameterized POR Regression - Calibrating on 40 Randomly Drawn Songs
Fit from 25 Different Samples
Standalone Revenue
100
200
300
400
500
100
150
200
250
Consumption in Bundle
Data
Fit - Full Data

Figure 8


Fit of Parameterized POR Regression - Over-Drawing from Popular and Unpopular Songs
Fit from 25 Different Samples of 10 Songs
Standalone Revenue
0
200
400
600
800
100
150
200
250
Consumption in Bundle
Data
Fit - Full Data
Three Songs Each were Drawn from the Highest and Lowest 10 Standalone Revenue Songs.
The Remaining Four Songs were Drawn from the Middle 30 Standalone Revenue Songs