Valuing Non-Contractual Firms Using Common Customer Metrics

Daniel M. McCarthy

Peter S. Fader[1]

February 24, 2017

[1]Daniel McCarthy is a Ph.D. Candidate in the Department of Statistics at the Wharton School of the University of Pennsylvania (email: danielmc@wharton.upenn.edu, web: http://www-stat.wharton.upenn.edu/~danielmc/). Peter S. Fader is the Frances and Pei-Yuan Chia Professor of Marketing at The Wharton School of the University of Pennsylvania (address: 771 Jon M. Huntsman Hall, 3730 Walnut Street, Philadelphia, PA 19104-6340; phone: (215) 898-1132; email: faderp@wharton.upenn.edu). We acknowledge financial support from Adobe Corporation through the Digital Marketing Research Award, the American Statistical Association through the Doctoral Research Award, the Baker Retailing Center, the ISMS through the Doctoral Dissertation Award, and the Marketing Science Institute through the Alden Clayton Award. We are grateful for the research assistantship of Elliot Oblander and Namita Nandakumar. We thank the anonymous retailer who contributed their data and insights.

Electronic copy available at: https://ssrn.com/abstract=2923466

Abstract

Valuing Non-Contractual Firms Using Common Customer Metrics

There is growing interest in "customer-based corporate valuation," explicitly tying the value of a firm's customer base to its financial valuation. While much progress has been made in building a well-validated customer-based valuation model for contractual (or subscription-based) firms, there has been little progress for non-contractual firms. Non-contractual businesses have more complex transactional patterns because they are characterized by latent attrition instead of observable churn behavior, and often have highly irregular purchase timing and spend amounts, making it harder to reconstruct granular purchase behaviors from aggregated metrics (e.g., active users and the repeat rate). Nevertheless, a number of firms disclose a variety of customer metrics to their shareholders. The authors use *indirect inference*, a well-established econometric estimation procedure, to estimate the aforementioned customer behaviors using aggregated metrics. The authors show how the predictive validity of the models varies as a function of the metrics used. They apply this methodology to data from a large business unit of an e-commerce retailer, valuing the business unit as a whole, decomposing this valuation into existing and yet-to-be-acquired customers, and analyzing customer profitability.

Keywords: customer lifetime value; customer equity; valuation; marketing metrics; indirect inference

Electronic copy available at: https://ssrn.com/abstract=2923466

Executives, marketing managers, and financial professionals are increasingly aware that current and future customer relationships are a valuable – if not the most valuable – asset of a firm. As a result, customer lifetime value (CLV, the net present value of a customer's future after-tax marginal profits) is being discussed, studied, and leveraged more than ever before (Braun et al. (2015), Datta et al. (2015), Maycotte (2015)). Accurate estimation of CLV gives external stakeholders (shareholders, creditors, suppliers, competitors, regulators, and other company outsiders) the ability to estimate customer equity (CE), which is equal to the remaining lifetime value of all existing customers plus the net present value of the CLV of all yet-to-be-acquired customers (Bauer and Hammerschmidt (2005), Blattberg and Deighton (1996), Rust, Lemon, and Zeithaml (2004)). For many firms, CE represents the majority of shareholder value (SHV) of the firm, enabling a direct, explicit link between customer behavior and the overall financial valuation of the firm.

"Customer-based corporate valuation" (CBCV) is the process of valuing the firm by forecasting current and future customer behavior using customer data in conjunction with traditional financial data. However, the correct valuation framework and the customer data needed to estimate the parameters of the model underlying that valuation framework depend on the nature of the relationship that the firm has with its customers. Performing CBCV for "contractual" (i.e., subscription-based) firms entails forecasting three main quantities – (1) future customer acquisitions, (2) how long acquired customers will remain with the firm before they churn, and (3) the monetary value associated with customers, on average, while those customers are active. Prior work has shown that two customer metrics are needed to estimate CE for contractual firms – total customers acquired and total customers churned each period[1] (Gupta, Lehmann, and Stuart (2004), Libai, Muller, and Peres (2009), McCarthy, Fader, and Hardie (2017), Schulze, Skiera, and Wiesel (2012), Wiesel, Skiera, and Villanueva (2008)). However, it is unclear what the appropriate valuation framework is for *non-contractual* firms (e.g., retail, travel/hospitality, media/entertainment, gaming, and more), what metrics external stakeholders need as inputs, and what estimation procedure external stakeholders can use to infer the parameters of the model given a particular collection of metrics. It is also unclear what collection of metrics is the most informative for external stakeholders if the firm is only willing to publicly disclose a limited number of them. The contribution of this article is to address these issues, proposing a novel methodology to estimate an integrated individual-level model of customer acquisition, repeat purchase, and spend behavior using a small collection of commonly disclosed non-contractual customer metrics.

There are two main aspects of this problem which have limited progress to date. First, non-contractual firms are more difficult to value than contractual ones because customer churn is unobserved and purchase behavior less predictable. Second, the data available to estimate the model is limited. We address each of these aspects in turn.

Recall the definition of expected customer lifetime value (E(CLV)):

$$\mathrm{E(CLV)} = \mathrm{E}\left[\int_0^\infty \mathrm{p}(t)\ \mathrm{m}(t)\ \mathrm{S}(t)\ \mathrm{d}(t)\ \mathrm{d}t\right], \tag{1}$$

where $\mathrm{p}(t)$ is the purchase rate at time t, $\mathrm{m}(t)$ is the monetary value associated with a transaction at time t, $\mathrm{S}(t)$ is the survivor function, and $\mathrm{d}(t)$ is the discount factor which accounts for the time value of money and the non-diversifiable riskiness of the customer and/or firm.

While the empirical survivor function $\mathrm{S}(t)$ in Equation 1 is observable for contractual firms, it is *unobservable* for non-contractual firms. When customers of a cell phone provider would like to end their relationships with the firm, they must let the firm know. In contrast, if customers of an e-commerce retailer decide to end their relationships with the firm, they simply discontinue purchasing from the firm. This complicates the underlying model (and thus data) required to predict future customer activity because non-contractual firms cannot report customers lost. The most popular and widely adopted/referenced workaround to this complication (Bain and Company (2000), Blattberg, Getz, and Thomas (2001), Gupta, Lehmann, and Stuart (2004), Libai, Muller, and Peres (2009), Seybold (2000)) is to create some notion of an observable "retention rate" which, in truth, does not exist in a non-contractual setting, and proceed with the same framework that is used for contractual settings. This proxy may be defined, for instance, as the repeat rate, or proportion of customers who made a purchase last year who made another purchase this year (Farris et al. (2010)). While the repeat rate may be a useful proxy for customer longevity, using it in a traditional

CLV formula in a non-contractual setting will understate future purchase activity (and thus future profits) dramatically because customers who have not purchased in one year may still be alive. For example, consumer products seller QVC notes that 6% of total sales in 2015 came from customers who had not purchased in over a year (QVC (2015)). In addition, retention propensities may differ across customers, and ignoring this variation will further undervalue them (Fader and Hardie (2010)).

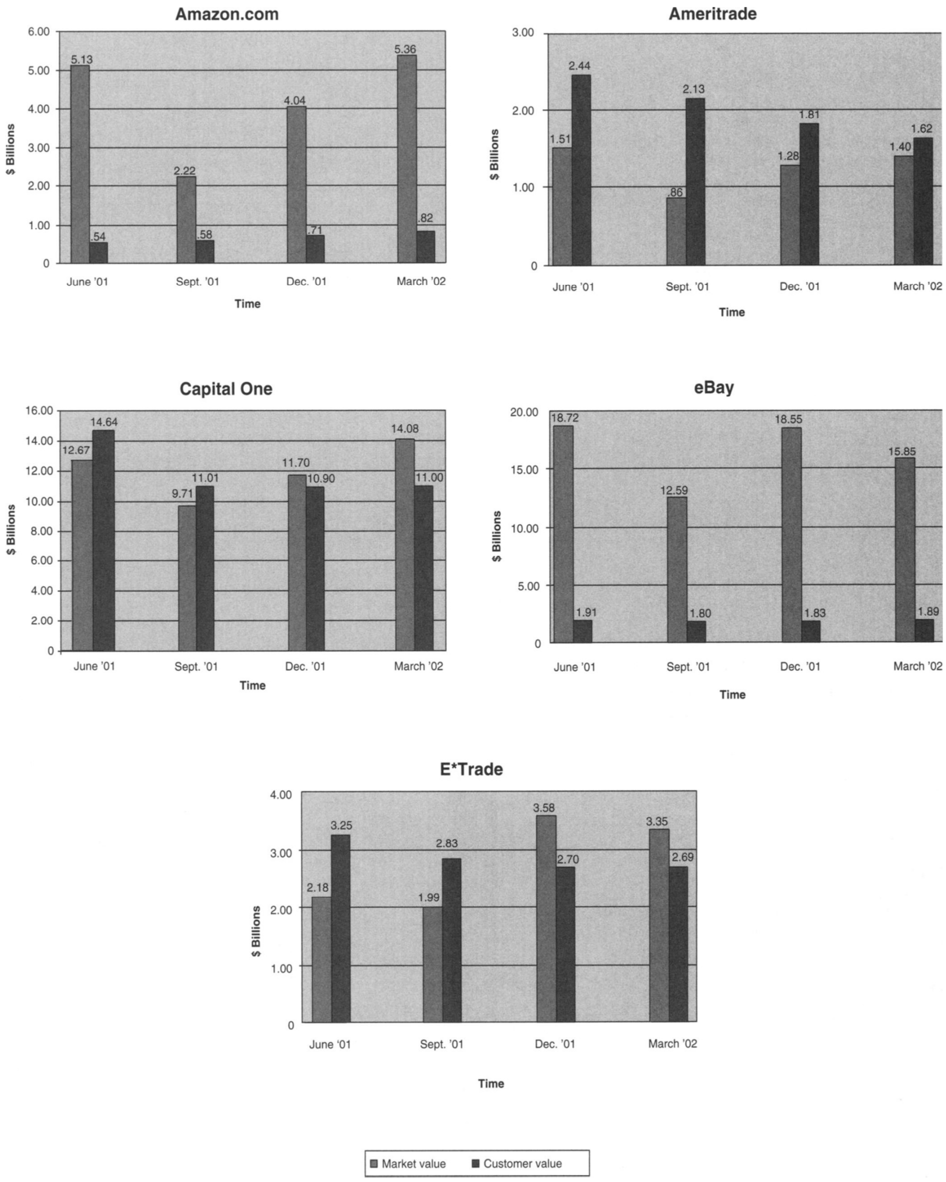
Even if a non-contractual customer were known to be alive, there are additional complexities associated with her repeat purchase and spend behavior that must be taken into account. Purchase rates p(t) and spend amounts associated with each of those purchases m(t) can be highly variable over time and across customers for non-contractual firms. A principled model for the future purchase and spend behavior for non-contractual customers must be able to estimate heterogeneity in these behaviors. In contrast, customers of contractual firms traditionally pay a fixed subscription fee each period, which is equivalent to exactly one purchase each period for all customers (i.e., p(t)=1) with a spend amount that is relatively constant over time and relatively homogeneous across customers.

A second statistical challenge is that external stakeholders of a firm do not get to observe individual-level customer data. Performing CBCV for non-contractual firms would be a solved problem if external stakeholders had access to individual-level data, because there are well-established, well-validated models specifically suited for non-contractual customer behavior (Fader, Hardie, and Shang (2010), Schmittlein, Morrison, and Colombo (1987)). In the absence of such data, customer behavior must be estimated using common customer metrics disclosed quarterly (e.g., active customers, or the repeat rate), or a collection of repeated cross-sectional summaries (Kinshuk, Fader, and Hardie (2016)). While estimating individual-level behaviors using aggregate information is a well-studied problem (Albuquerque and Bronnenberg (2009), Chen and Yang (2007), Feit et al. (2013), Musalem, Bradlow, and Raju (2008), Musalem, Bradlow, and Raju (2009)), our problem setting differs from previous works in three ways. First, the data in our setting is a possibly incomplete, wide-ranging collection of cross-sectional data summaries, and not cross-sectional market share data alone, or market share data augmented by customer purchase frequencies marginalized over time. Second, our goal is not only to estimate a model given a fixed set of disclosures, but also to recommend what set of metrics offers the best predictive performance for the smallest number of disclosures. Third, our data often encompasses millions of prospects and customers over years or even decades.

In the pioneering customer-based valuation work of Gupta, Lehmann, and Stuart (2004) (hereafter, GLS), both contractual and non-contractual companies were valued with the same model. For non-contractual firms, a retention rate proxy was used as the retention rate in their CLV formula. Of the five companies valued in their empirical analysis, the only two non-contractual businesses, eBay and Amazon, were also the two most misvalued. These companies were undervalued by an average of 88% and 83%, respectively[2], even though these valuations were performed after the stock market had fallen sharply in the aftermath of the "tech bubble," and prominent Wall Street analysts were publicly questioning Amazon's solvency at the time (Arango (2001), Streitfeld (2001)). Figure 1 reproduces the valuation estimates from GLS for reference. A model specifically suited to non-contractual businesses is needed.

The procedure we use to estimate the proposed valuation model relies on indirect inference (Gallant and Tauchen (1996), Gourieroux, Monfort, and Renault (1993), Smith (1993), hereafter denoted as "II"). II is a versatile simulation-based estimation method which is particularly useful when fitting complex models using limited data (e.g., censored, missing, and/or aggregate data, or omitted covariates (Jiang and Turnbull (2004))), characteristics that describe our problem setting well. Generalized method of moments (GMM, Hansen (1982)) can be thought of as a special example of II (Jiang and Turnbull (2004)). II subsumes simulated method of moments (SMM, McFadden (1989)), which is a popular parameter estimation technique within the marketing literature (Chintagunta (1992), Dutta, Narasimhan, and Rajiv (1999), Erdem (1996), Gönül and Srinivasan (1993), Mehta, Rajiv, and Srinivasan (2003)). SMM is not generally applicable in our setting because the set of metrics that we consider may not be moments in a conventional sense (e.g., firms will not disclose metrics involving skewness or even variance). As with SMM, however, II is useful when it is difficult or impossible to compute the likelihood function of a proposed model,

Figure 1: Market Value and Customer Value Over Time: Gupta, Lehmann, and Stuart (2004)



but easy to simulate from that model. We use II to find the set of true model parameters that minimize the expected distance between moments of convenient but misspecified "auxiliary models," which are derived from the observed disclosures, and what we expect those auxiliary moments should be.

In the next section, we present the model governing customers' acquisition, repeat purchasing, and spend, and how this model is used to drive an overall valuation for the firm. After providing the common customer metrics which will be used to estimate the customer model, we show how II is used to perform model estimation. We then analyze the predictive performance of all possible collections of these metrics through a large-scale simulation analysis. We apply this methodology to a 5.5-year transaction log data set from an e-commerce retailer business unit. We provide an overall valuation for the business unit, then decompose this valuation into existing versus yet-to-be-acquired customers, and analyze the unit economics of newly acquired customer cohorts. We conclude with a discussion of the results.

MODEL DEVELOPMENT

In this section, we specify the individual-level model for the customer which we will use to forecast future customer activity and show how this model is embedded within an overall valuation framework for the firm.

Valuation Framework

We adopt the discounted cash flow (DCF) model as our firm valuation framework (Damodaran (2012), Greenwald et al. (2004), Holthausen and Zmijewski (2013), Koller, Goedhart, and Wessels (2010), Schulze et al. (2012)). Assuming a weekly "clock" for disaggregate customer purchase activity, shareholder value (SHV) is equal to the value of the firm's operating assets (OA) plus the non-operating assets (NOA), minus the net debt (ND):

$$\text{SHV}(w) = \text{OA}(w) + \text{NOA}(w) - \text{ND}(w), \tag{2}$$

where $w = 1$ represents the first week of the company's commercial operations. SHV is observed in effectively continuous time for publicly traded firms, while OA, NOA, and ND are observed at the end of each quarter by external stakeholders. The value of a firm's operating assets (OA) is equal to the sum of all future free cash flows (FCF's) the firm will generate, discounted at the weighted average cost of capital (WACC, assuming weekly compounding):

$$\text{OA}(w) = \sum_{w'=0}^{\infty} \frac{\text{FCF}(w + w')}{(1 + \text{WACC})^{w'}}. \tag{3}$$

If we assume no cash flow-related adjustments (see Damodaran (2012)), weekly FCF for customer-based businesses is equal to weekly revenues (R) multiplied by the contribution margin ratio ($1-$VC), minus total customer acquisition expenses (cost per acquired customer multiplied by the number of acquired customers, or CAC$\times$A) and fixed operating costs (FC), after taxes (where TR is the firm's tax rate):

$$\mathrm{FCF}(w) \;=\; \{\mathrm{R}(w) \times [1 - \mathrm{VC}(w)] - \mathrm{FC}(w) - \mathrm{CAC}(w) \times \mathrm{A}(w)\} \times [1 - \mathrm{TR}(w)]. \quad (4)$$

Financial professionals typically model and forecast revenues and expenses using time-series models. This may be sensible when firm financial disclosures do not include customer data. If customer data is available, however, forecasting accuracy can be improved by decomposing customer-driven financial line items into their constituent parts, estimating models and then forecasting these parts into the future, then aggregating these parts together again to form projections of future customer-driven financial line items.

Firms incur costs to acquire customers. At non-contractual firms, customers are acquired when they make their first purchase. After they are acquired, they make repeat purchases in future weeks until they end their relationship with the firm. Recognizing this, customer data informs Equation 4 in two ways. First, it informs revenues (R), which may be decomposed into total purchases (initial and repeat) and average spend per purchase. Second, it informs total customer acquisition expenses, which are a function of the total number of customers acquired (A). The remaining variables in Equation 4 – VC, FC, CAC, TR, WACC, NOA, and ND – are modeled and projected in essentially the same way that a financial professional would normally do (for example, by obtaining an estimate from a third party, or by using a simple time series extrapolation of historical data).

Our goal, then, is to specify processes for the acquisition of new customers, how many repeat purchases these customers make after they have been acquired before they churn, and how much they will spend on each of those purchases. We estimate the parameters of these models so that the weekly behaviors implied by the model are consistent with the quarterly disclosures provided by the firm. We combine the projections from these processes to forecast R and A in Equation 4 into the future, then plug the resulting FCF(w) projections into Equation 3 to value the firm's operating assets.

If these processes predict future customer activity well, the valuation forecasts which flow from the model will be more accurate as a result. This is a prediction problem, where the modeler is assumed to be an external stakeholder who only has access to limited data. As such, our model does not correct for endogeneity and does not include pricing/marketing mix data. While it is possible to incorporate endogenous variables (Schweidel and Knox (2013)), this would require additional data that external stakeholders are unlikely to have access to, and would be costly for firms to disclose publicly. Furthermore, an endogeneity-corrected valuation model may have lower holdout predictive validity than the same model which does not correct for endogeneity, particularly because endogenous variables are not observed in the holdout period (Ebbes, Papies, and Van Heerde (2011)), even by the firm itself. Examples of prior literature in which endogeneity-corrected models are proposed that underperform their non-corrected counterparts on holdout data include Besanko, Gupta, and Jain (1998) and Neslin (1990). Endogeneity correction is less helpful in our prediction-focused, limited-data setting.

Model Specification

Our proposed model for the timing of customer adoption consists of two parts: (1) the formation of pools of prospective customers over time (e.g., through household formation/population growth), and (2) the duration of time which elapses from the time a prospect is "born" to when the prospect adopts the service. We drive the creation of prospect pools over time off of the population size. At the beginning of the firm's commercial operations, there is an initial prospect pool $\text{M}(0)$ which is equal to the population size at the time, $\text{POP}(0)$ (i.e., the total number of US households at the time of incorporation). This prospect pool will eventually adopt in future weeks $w = 1, 2,$ and so on. The size of the prospect pool in a given week w is equal to population growth during the week:

$$\text{M}(w) = \text{POP}(w) - \text{POP}(w-1), \quad w = 1, 2, \ldots \tag{5}$$

After a prospect pool is formed, we model the duration of time until customer adoption through a flexible yet parsimonious hazard model. Let $\mathrm{F}_A(w' - w|w)$ denote the probability that a prospect from week w will adopt by the end of week w'. Letting $\mathrm{A}(w')$ be the number of adopters in week w', $\boldsymbol{\psi}_A$ be the collection of parameters underlying the acquisition model, and $\mathbf{x}_A(w')$ be a p-length vector containing values of our covariates in week w', then

$$\mathrm{A}(w') = \sum_{w=0}^{w'-1} \mathrm{M}(w) \times \{\mathrm{F}_A[w' - w|w, \mathbf{x}_A(w'), \boldsymbol{\psi}_A] - \mathrm{F}_A[w' - w - 1|w, \mathbf{x}_A(w' - 1), \boldsymbol{\psi}_A]\}. \quad (6)$$

The hazard model we use to model a prospect's time until adoption is a Weibull(λ_A) baseline with covariate effects incorporated through proportional hazards, with cross-sectional heterogeneity in the baseline propensity, λ_A, characterized by a gamma(r_A, α_A) distribution. In previous literature, this model specification has very successfully modeled duration data in general (Fader and Hardie (2009), Moe and Fader (2002), Morrison and Schmittlein (1980), Schweidel, Fader, and Bradlow (2008)), and customer acquisitions in a CBCV setting in particular (McCarthy, Fader, and Hardie (2017)). Given a homogeneous acquisition shape parameter (c_A), homogeneous but possibly time-varying acquisition covariates up to week w' ($\mathbf{X}_A(w') = [\mathbf{x}_A(1),\ \mathbf{x}_A(2),\ \ldots,\ \mathbf{x}_A(w')]$), and coefficients for acquisition covariates ($\boldsymbol{\beta}_A$), the unconditional probability that a customer from prospect pool w will be acquired by the end of week w' is

$$\begin{aligned}
\mathrm{F}_A(w' - w|w, \mathbf{X}_A(w'), \boldsymbol{\psi}_A) &= \int_0^\infty \mathrm{F}_A[w' - w|w, \lambda_A, c_A, \mathbf{X}_A(w'), \boldsymbol{\beta}_A]\mathrm{f}(\lambda_A|r_A, \alpha_A)d\lambda_A \\
&= 1 - \left[\frac{\alpha_A}{\alpha_A + B_A(w, w' - w)}\right]^{r_A}, \quad \text{where} \qquad (7) \\
\mathrm{B}_A(w, w' - w) &= \sum_{i=w+1}^{w'} [(i - w)^{c_A} - (i - w - 1)^{c_A}] e^{\boldsymbol{\beta}_A^T \mathbf{x}_A(i)} \qquad (8)
\end{aligned}$$

This formulation allows the propensity to acquire services to be affected by factors such as a recession and seasonality.

Our baseline model for the number of repeat purchases the customer makes is a generalization of the Beta-Geometric/Beta-Binomial (BG/BB) model for non-contractual customer base analysis (Fader, Hardie, and Shang (2010)). Immediately after the customer places his/her initial order, he/she is in an alive state. While alive, he/she makes a purchase in week w with probability $q(w)$. This probability may be higher or lower due to unobserved heterogeneity, or due to external factors such as the state of the macroeconomy or seasonality. We allow for both effects through a logit-normal formulation, so that

$$q(w) = \frac{\exp[b_p + \boldsymbol{\beta}_p^T \mathbf{x}_p(w)]}{1 + \exp[b_p + \boldsymbol{\beta}_p^T \mathbf{x}_p(w)]}, \tag{9}$$

where the baseline purchase propensity b_p is distributed across the population according to a normal(μ_p, σ_p^2) distribution, $\mathbf{x}_p(w)$ are covariates associated with week w, and $\boldsymbol{\beta}_p$ are the coefficients associated with those covariates.

Each period, the customer may churn with probability θ. We assume that a customer who has churned has a 0% probability of making a future purchase. Because customer churn is not observed, this is a so-called "leaky always-a-share" model (Fader and Hardie (2014)). We let θ vary across the population according to a gamma(γ, δ) distribution. Going forward, we will refer to this model as the Beta-Geometric/Mixed-Logit (BG/ML) model. Had we assumed that $q(w)$ varied across customers but not time and that $q(w)$ is distributed across the population according to a beta distribution, this model reduces to the BG/BB model.

Finally, we model the underlying expected amount spent per purchase and how it varies across customers (Fader, Hardie, and Lee (2005b) and Schmittlein and Peterson (1994)). The expected amount spent on a given purchase in week w, $\mathrm{E}[s(w)]$, is driven by a baseline propensity to spend, b_s, and a time trend term to allow expected spend to vary as a function of which quarter q the customer is in. A time trend in average spend per purchase is often necessary – for example, it is evident in the empirical example which follows, and in the two public company valuations in McCarthy, Fader, and Hardie (2017). The baseline propensity to spend varies across customers according to a lognormal distribution:

$$\mathrm{E}[s(w)] = b_s + q\mu_q, \quad \text{where} \quad \log(b_s) \sim \mathcal{N}(\mu_0, \tau^2) \tag{10}$$

The spend formulation ensures that customers' expected spends are strictly positive with a heavy right tail, consistent with non-contractual transaction log data.

These models are flexible in their ability to capture variation across customers and over time while remaining parametrically parsimonious. Parsimony is important in our setting because of the limited data which is available to train the model upon. In our empirical example, we train the three models summarized above using four quarterly customer metrics.

Valuation Procedure

Our goal is to use the fitted models from the previous section to forecast R and A in Equation 4 into the future, then use these projections to come up with an overall valuation estimate. In this section, we outline how to estimate the overall value of the firm's operating assets, after the parameters of the model have been estimated. We will discuss estimation (and the data used to do the estimation) over the next two sections.

First, we estimate weekly customer acquisitions far into the future, to a point in time when the present value of all future cash flows is negligible. Let the calibration period and forecasting horizon be W and W^* weeks long, respectively (in practice, we use a 50-year forecasting horizon, so $W^* = 2600$). To come up with weekly acquisition projections, we need estimates of POP(w) over the forecasting horizon. In the analyses which follow, we assume that projections of POP(w) over the forecasting horizon are provided to us. If projections of POP(w) are not available, we could use a simple model which estimates future growth as an extrapolation of historical growth. We use Equation 5 to turn POP(w) projections into prospect pools M(w) over all periods $w = 1, 2, \ldots, W + W^*$. We then use the fitted acquisition parameters $(r_A, \alpha_A, c_A, \boldsymbol{\beta}_A)$, and the historical and projected future prospect pools [M(1),. . ., M$(W + W^*)$] to obtain historical and future expected customer acquisitions [A(1),. . ., A$(W + W^*)$] using Equation 6.

Second, we estimate weekly total (initial plus repeat) purchases, given the historical and future customer acquisitions [A(1),. . ., A$(W + W^*)$] and the parameters of the repeat purchase process $(\mu_p, \sigma_p^2, \boldsymbol{\beta}_p)$. Because there is no closed-form expression for the unconditional expected number of repeat purchases by week, we use Monte Carlo simulation instead. Letting A* be the total number of customers that will eventually be acquired by the end of the forecasting horizon (i.e., A$^* = \sum_{w=1}^{W+W^*}$ A(w)), we repeatedly simulate a binary "total purchases matrix," TP, which records the purchases each eventually-acquired customer will make in each week. The first A(1) rows correspond to the customers acquired in week one, the following A(2) rows correspond to customers acquired in week 2, and so on. Letting TP$^{(k)}$ be the kth simulated realization of this matrix,

$$\text{TP}^{(k)} = \begin{bmatrix} \text{y}_{1,1}^{(k)} & \text{y}_{1,2}^{(k)} & \cdots & \text{y}_{1,W+W^*}^{(k)} \\ \text{y}_{2,1}^{(k)} & \text{y}_{2,2}^{(k)} & \cdots & \text{y}_{2,W+W^*}^{(k)} \\ \vdots & \vdots & \ddots & \vdots \\ \text{y}_{A^*,1}^{(k)} & \text{y}_{A^*,2}^{(k)} & \cdots & \text{y}_{A^*,W+W^*} \end{bmatrix}, \tag{11}$$

where $\text{y}_{j,w}^{(k)}$ is a binary variable equal to one if the kth simulation of customer j made a purchase in week w, 0 otherwise. Assuming that customer j is acquired in week w, we simulate his/her individual-level parameters from their respective heterogeneity distributions, $\theta_j^{(k)} \sim \text{beta}(\gamma, \delta)$ and $b_{p,j}^{(k)} \sim \mathcal{N}(\mu_p, \sigma_p^2)$. The binary sequence associated with customer j in each week w' (i.e., the jth row of TP$^{(k)}$) is determined by the following 4 scenarios:

$$\text{y}_{j,w'} = \begin{cases} 1 \text{ with probability } 1 & w' = w \\ 0 \text{ with probability } 1 & w' < w \\ 1 \text{ with probability } (1-\theta_j^{(k)})^{w'-w} \dfrac{\exp[b_p^{(k)} + \boldsymbol{\beta}_p^T \mathbf{x}_p(w')]}{1+\exp[b_p^{(k)} + \boldsymbol{\beta}_p^T \mathbf{x}_p(w')]} & w < w' \\ 0 \text{ otherwise} \end{cases}$$

The expected purchases made by all customers over all time periods, $\widehat{\text{TP}}$, is formed by taking the element-wise average of TP$^{(k)}$ over K replications (in practice, we set $K = 100$, but the appropriate value for K should be a function of A*). The column sums of $\widehat{\text{TP}}$ are equal to the expected total number of purchases for the firm as a whole over the entire calibration period and forecasting horizon. We denote the $W+W^*$-length vector of expected historical and future purchases by [P(1),. . ., P$(W + W^*)$].

Third, we estimate weekly revenues, given the vector of expected weekly total purchases [P(1),. . ., P$(W + W^*)$] and the parameters of the spend process (μ_0, τ^2, μ_q). Let E$[s(w)]$ be the unconditional expected spend per purchase in week w. Assuming week w is in quarter

q,

$$\mathrm{E}[s(w|\mu_0, \tau^2, \mu_q)] = \exp(\mu_0 + \tau^2/2) + q\mu_q.$$

Expected firm revenues in week w, $\mathrm{R}(w)$ (Equation 4), is equal to the product of expected total purchases in week w ($\mathrm{P}(w)$) with expected spend per purchase in week w ($\mathrm{E}[s(w)]$). We use future weekly revenues $[\mathrm{R}(W+1), \ldots, \mathrm{R}(W+W^*)]$ to estimate the valuation of the firm in week W.

As noted in the valuation framework, once we have estimated future revenues R and gross customer acquisitions A, the remaining inputs needed to obtain FCF, OA, and SHV (Equations 2, 3, and 4) are obtained using the same procedure that a financial professional would use. See McCarthy, Fader, and Hardie (2017) for two public company valuations. The operating unit we value in our empirical example is privately held, so all of these non-customer-driven line item projections were provided to us by the company's management team.

The valuation procedure assumes that we have already estimated the parameters of the model. In the next two sections, we discuss the data which is available to perform the estimation, and how we can estimate the model parameters with this information.

CANDIDATE CUSTOMER METRICS

As we described previously, public non-contractual firms do not publicly disclose individual-level customer data. However, some companies do disclose aggregated customer metrics. Furthermore, these "first-party" disclosures released by the firms themselves are supplemented by third-party disclosures, mined by business intelligence and market intelligence firms such as 1010data, SecondMeasure, SurveyMonkey Intelligence, AppAnnie, Slice Intelligence, and Prosper Analytics. What are the metrics that these firms do, and/or should, disclose? We discuss customer data summaries most relevant to each of the non-contractual customer model processes in turn. For all customer metrics, we follow firms' usual convention of providing this number on a trailing twelve month basis (i.e., the firm reports the number of customers acquired over the past year).

For the acquisition process, gross customers acquired is the most natural and most important customer data summary (McCarthy, Fader, and Hardie (2017)). Dozens of contractual firms, and even some forward-thinking non-contractual firms (QVC (2015)), regularly disclose this measure. From a statistical relevance standpoint, gross customers acquired reliably identifies the parameters of our proposed acquisition model (see Appendix A for a full factorial parameter recovery analysis). While we could easily consider other acquisition-related measures, these factors create little practical incentive to do so.

For the spend process, we obtain mean spend "for free," because it can be derived from total revenues (which must be disclosed by public firms) in conjunction with the purchase process. As noted earlier, however, spend per transaction is more variable than a routine subscription payment and thus it is more important for us to capture the shape of the distribution, rather than simply projecting the means. Median spend is a very natural companion measure to mean spend. Median spend is a communicable, intuitively appealing measure of basket size for external stakeholders. Disclosures of mean and median spend over time identify the parameters of the spend process well. As is the case with the acquisition process, while other disclosures could in theory be considered, it is not apparent what spend measures would be better, practically and statistically.

What is less clear is what should be recommended for the repeat purchase process. We focus on the following six common repeat purchase measures:

1. Active users (AU): The number of customers who make 1+ purchases in the past year.
2. Heavy users (HAU): The number of customers who make 2+ purchases in the past year.
3. Repeat rate (RR): The proportion of customers who made a purchase last year, who purchase again this year.
4. Repeat buyer proportion-customers (RBPC): The proportion of customers who made a purchase this year who also purchased before this year began.
5. Repeat buyer proportion-orders (RBPO): The proportion of orders made this year by customers who also purchased before this year began.
6. Average frequency (F): The average number of purchases made by all active customers over the past year.

A non-exhaustive list of first-party disclosures of these metrics (or very closely-related metrics) are provided in Table 1. Third-party disclosures are provided in parentheses. While the most commonly disclosed metric is active users, the other metrics are also frequently disclosed.

Table 1: Common Disclosures (Third-Party Disclosures in Parentheses)

Metric	Firms
Active Users	Amazon, Camping World, Caesars Acquisition Co., Evine, Facebook, Gamzio Mobile, Glu Mobile Inc., GoPro, HSN, International Game Technology, LinkedIn, MEDL Mobile Holdings, MeetMe, Inc., Quepasa Corp., QVC, Scientific Games Corp., Snap Interactive, Social Reality, Sohu Com Inc., Twitter, Uber, Wayfair, Zynga (Niantic)
Heavy Active Users	QVC (Aveda, Bare Essentials, ULTA)
Repeat Rate	QVC, Uber (Amazon)
Repeat Buyer Proportion	Amazon, Etsy, QVC, Wayfair (HSN, Instacart, Jet.com)
Average Frequency	Evine

We provide a simple numerical example in Table 2 to further illustrate how granular transaction log data is summarized along different dimensions by the customer metrics. For expositional purposes, we only show disclosures at the end of each *year* in this numerical example (i.e., at the end of years one, two, and three). In the analysis that follows, however, we assume that disclosures can be made each *quarter*. For example, AU as of the end of the fourth quarter of commercial operations represents the number of customers who made at least one purchase in quarters one through four, AU as of the end of the fifth quarter of commercial operations represents the number of customers who made at least one purchase in quarters two through five, and so on. While our estimation procedure can be used for firms that only disclose once per year by treating the first three quarterly disclosures of each year as missing data, annual disclosure periodicity implies a very small number of data points to perform modeling with and thus may not be empirically identified.

An important question is how do we use these customer metrics to estimate the latent variable model specified in the previous section. We answer this question in the next section.

Table 2: Numerical Example of Acquisition and Total Purchase Metrics

Number of Purchases by Customer/Year

	Y1	Y2	Y3
Customer 1	1	1	0
Customer 2		3	3
Customer 3			1
Gross Acquisitions	1	1	1
AU	1	2	2
HAU	0	1	1
RR	NA	100%	50%
RBPC	0%	50%	50%
RBPO	0%	25%	75%
F	1	2	2

ESTIMATION WITH INDIRECT INFERENCE

The disclosures from the previous section are not conventional model-dependent moments. For example, the purchase disclosures are overlapping, cross-sectional, often highly non-linear summaries involving both the acquisition process and the repeat purchase process at once (i.e., these measures do not distinguish between initial and repeat purchases). In this section, we provide a brief overview of II and how it enables us to estimate the model parameters using these data summaries, when evaluation of the likelihood function is analytically intractable. We start by providing a general framework which can be applied to any panel dataset. We then provide an illustration of how we use II to estimate the parameters of the acquisition process, and how the other two processes are estimated. We provide a detailed description of the auxiliary model specifications for the repeat purchase and spend processes in Appendix B.

General Parameter Estimation Framework

Let Y_{jt} be a random variable representing the behavior of customer j at time t, for $j = 1, 2, \ldots, J$ and $t = 1, 2, \ldots, T$. Assume that the distribution of Y_{jt} is a function of p model parameters $\boldsymbol{\psi}$ and time-varying covariates $\mathbf{x}_{jt}$, so that $y_{jt} \sim f(Y_{jt}|\boldsymbol{\psi}, \mathbf{x}_{jt})$. We estimate $\boldsymbol{\psi}$

through the following three-step procedure:

1. Compute a q-dimensional "auxiliary statistic" $\hat{\mathrm{s}}$ which is a deterministic function of the empirical cumulative distribution function (ECDF) $\hat{F}_{JT}$ (i.e., $\hat{\mathrm{s}} = g(\hat{F}_{JT})$), where $q \geq p$. $\hat{\mathrm{s}}$ should represent key characteristics of the data. These auxiliary statistics are often moments of the data or deterministic functions of the true model parameters under a convenient but misspecified "auxiliary model." This allows us to match off of non-moments of the data under the true model, unlike the GMM estimator. In the next section we provide the auxiliary model we use for the acquisition process, and in Appendix B we provide the auxiliary models for the repeat purchase and spend processes.

2. Let the "auxiliary parameter" s be the limiting value of the auxiliary statistic $\hat{\mathrm{s}}$ as $(J, T) \to \infty$ (i.e., $\mathrm{s} = g(F_{\boldsymbol{\psi}})$). Define the "binding relationship," which relates the unknown model parameters $\boldsymbol{\psi}$ to the auxiliary parameter s. Then $\mathrm{s} = \mathrm{s}(\boldsymbol{\psi})$.

3. Use the binding relationship to find the model parameters $\boldsymbol{\psi}$ which minimize the "distance" between the observed auxiliary statistic $\hat{\mathrm{s}}$ and the auxiliary parameter under $\boldsymbol{\psi}$, $\mathrm{s}(\boldsymbol{\psi})$, under a suitable estimate of the distance or weight matrix W to be defined below[3]:

$$\hat{\boldsymbol{\psi}} = \mathrm{argmin}_{\boldsymbol{\psi}}[\mathrm{s}(\boldsymbol{\psi}) - \hat{\mathrm{s}}]^T \hat{W} [\mathrm{s}(\boldsymbol{\psi}) - \hat{\mathrm{s}}]. \tag{12}$$

II is often used when no closed-form expression is available relating model parameters $\boldsymbol{\psi}$ to auxiliary parameters $\mathrm{s}(\boldsymbol{\psi})$. When there is no such closed-form expression, we simulate the binding relationship. For a given $\boldsymbol{\psi}$, we simulate K new datasets. Letting $\hat{F}_{JT}^{(k)}$ be the ECDF associated with simulation k, compute K auxiliary statistics under the K simulated datasets, $[\hat{\mathrm{s}}^{(1)}(\boldsymbol{\psi}), \hat{\mathrm{s}}^{(2)}(\boldsymbol{\psi}), \ldots, \hat{\mathrm{s}}^{(K)}(\boldsymbol{\psi})]$. Then for K large, $\mathrm{s}(\boldsymbol{\psi}) \approx \sum_{k=1}^{K} \hat{\mathrm{s}}^{(k)}(\boldsymbol{\psi})/K$. In the empirical analyses which follow, K is set so that stochastic variability in the average is acceptably low (in practice, it is set so that the average empirical coefficient of variation is less than .5%).

This procedure is consistent under mild regularity conditions as long as the weight matrix W is positive definite (for example, the identity matrix). However, the optimal weight matrix is the asymptotic inverse variance-covariance matrix of the estimator solution:

$$\widehat{W} = \lim_{K \to \infty} \frac{1}{K-1} \sum_{k=1}^{K} \{[\widehat{s}(\boldsymbol{\psi}_A)^{(k)} - \mathbf{s}(\boldsymbol{\psi}_A)][\widehat{s}(\boldsymbol{\psi}_A)^{(k)} - \mathbf{s}(\boldsymbol{\psi}_A)]^T\}^{-1}. \tag{13}$$

The optimal weight matrix is preferable to the identity matrix for three reasons: (1) well-identified moments have higher weights, (2) parameter estimation is invariant to changes in the scale of the auxiliary statistics, and (3) correlations between auxiliary statistics are accounted for (i.e., there is no "double counting" of highly correlated auxiliary statistics).

In contrast, standard nonlinear least squares (NLS) estimation gives equal weight to all observed data points, which diminishes its finite sample performance. This issue is particularly acute in our setting, where different disclosures operate on very different scales (e.g., we may observe RR of 50% and AU of 5MM), and firm disclosures may be strongly correlated with one another (e.g., AU and HAU). As with NLS, however, this procedure does not require calculation of the likelihood function. One must solve the objective function provided in Equation 12.

As an illustration, we show how we use this procedure to estimate the parameters of the acquisition process in the next section.

Acquisition Process Parameter Estimation

Assume that we observe an arbitrary collection of n gross acquisition disclosures, **ADD**,

$$\textbf{ADD} \equiv [\text{ADD}_{q_1}, \text{ADD}_{q_2}, \cdots, \text{ADD}_{q_n}],$$

whose ith element, ADD_{q_i}, is equal to the trailing year sum of gross customers acquired. We posit a non-parametric auxiliary model in which all of the prospects which exist by the end of the calibration period, $\text{POP}(W)$, are independent and identically distributed, and can be acquired starting immediately after commercial operations begin. The acquisition time of each of the prospects under the auxiliary model is determined by a multinomial$(\pi_1^{(a)}, \pi_2^{(a)}, \cdots, \pi_Q^{(a)}, \pi_\emptyset^{(a)})$ draw. A non-parametric auxiliary model maximizes the amount of information we obtain from the observed data and over-identifies the true model. $\pi_q^{(a)}$ is the probability that the customer is acquired in quarter q, and $\pi_\emptyset^{(a)}$ is the probability that the customer is not acquired during the calibration period. **ADD**/POP(W) are sums of empirical moments of the auxiliary model under the observed data. For example, for $q_i \geq 4$,

$$\text{ADD}_{q_i}/\text{POP}(W) = \widehat{E}\left(\pi_{q_i-3}^{(a)} + \pi_{q_i-2}^{(a)} + \pi_{q_i-1}^{(a)} + \pi_{q_i}^{(a)}\right).$$

ADD/POP(W) is the auxiliary statistic for the acquisition process, $\widehat{\mathrm{s}}^{(a)}$.

For each set of acquisition parameters $\boldsymbol{\psi}_A$ that we consider, we obtain the corresponding auxiliary parameter $\mathbf{s}(\boldsymbol{\psi}_A)$. We do this by repeatedly simulating a "total acquisitions matrix", TA($\boldsymbol{\psi}_A$), which records the number of prospects within each prospect pool who are acquired in each week. The first row corresponds to the M(0) prospects from the week-0 prospect pool, the second row corresponds to the M(1) prospects from the week-1 prospect pool, and so on. Letting TA$^{(k)}$ be the kth simulated realization of this matrix,

$$\text{TA}(\boldsymbol{\psi}_A)^{(k)} = \begin{bmatrix} \mathrm{a}_{0,1}^{(k)} & \mathrm{a}_{0,2}^{(k)} & \cdots & \mathrm{a}_{0,W}^{(k)} \\ \mathrm{a}_{1,1}^{(k)} & \mathrm{a}_{1,2}^{(k)} & \cdots & \mathrm{a}_{1,W}^{(k)} \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{a}_{W-1,1}^{(k)} & \mathrm{a}_{W-1,2}^{(k)} & \cdots & \mathrm{a}_{W-1,W}^{(k)} \end{bmatrix}, \tag{14}$$

where $\mathrm{a}_{w,w'}^{(k)}$ is equal to the kth realization of the number of prospects from M(w) prospect pool who are acquired in week w'. Consider the $(w+1)$st row of TA$(\boldsymbol{\psi}_A)^{(k)}$, corresponding to prospect pool M(w). Given $\boldsymbol{\psi}_A$, a simulated realization of this row is equal to

$$(\mathrm{a}_{w,1}^{(k)}, \ldots, \mathrm{a}_{w,w}^{(k)}) = 0 \text{ and } (\mathrm{a}_{w,w+1}^{(k)}, \ldots, \mathrm{a}_{w,W}^{(k)}) \sim \text{multinomial}[\mathbf{M}(w); \phi_{w,w+1}, \cdots, \phi_{w,W}, \phi_{w,\emptyset}],$$

where $\phi_{w,w'}$ is defined via difference of CDF's using Equation 7:

$$\phi_{w,w'} \equiv \mathrm{F}_A[w' - w | w, \mathbf{X}_A(w'), \boldsymbol{\psi}_A] - \mathrm{F}_A[w' - w - 1 | w, \mathbf{X}_A(w' - 1), \boldsymbol{\psi}_A].$$

The kth realization of the number of acquisitions across all prospect pools is equal to the column sums of $\mathrm{TA}^{(k)}$, so the corresponding auxiliary statistic $\widehat{s}(\boldsymbol{\psi}_A)^{(k)}$ is equal to

$$\widehat{s}(\boldsymbol{\psi}_A)^{(k)} = \left(\sum_{w'=52q_1-51}^{52q_1} \sum_{w=0}^{W-1} \mathrm{a}_{w,w'}^{(k)}, \cdots, \sum_{w'=52q_n-51}^{52q_n} \sum_{w=0}^{W-1} \mathrm{a}_{w,w'}^{(k)} \right) \Big/ \mathrm{POP}(W)$$

The auxiliary parameter corresponding to a particular set of acquisition parameters $\boldsymbol{\psi}_A$ is equal to $\mathrm{s}(\boldsymbol{\psi}_A) = \sum_{k=1}^{K} \widehat{s}(\boldsymbol{\psi}_A)^{(k)}/K$.

To estimate the parameters of the acquisition process using indirect inference assuming this auxiliary model, we perform the following steps.

First, we initialize the distance metric to an n-dimensional identity matrix, $\widehat{W} = I_n$.

Second, we update our estimate of the acquisition parameters, $\widehat{\boldsymbol{\psi}}_A$. We find the set of parameters that minimizes the expected Wald distance of $\widehat{\mathrm{s}}^{(a)}$ from $\mathrm{s}(\boldsymbol{\psi}_A)$ under the distance metric $\widehat{W}$ using Equation 12.

Third, we update the distance metric $\widehat{W}$ given $\widehat{\boldsymbol{\psi}}_A$, using Equation 13.

Finally, we repeat steps two and three as many times as desired. Following Wooldridge (2002) (p. 193), we perform two iterations, and did not observe any material improvement when we used more than two iterations.

The process is for all intents and purposes the same when we introduce the repeat purchase and spend processes into the estimation procedure. We posit auxiliary models corresponding to each process, then find the set of true model parameters whose corresponding auxiliary parameters are as close as possible to the observed auxiliary statistics which are derived from the observed disclosures using a two-step estimator for the optimal weight matrix. Auxiliary model specifications for the repeat purchase and spend processes are provided in Appendix B. We exploit the fact that the true acquisition process is independent of the repeat purchase and spend processes, and that the true repeat purchase process is independent of the spend process, to improve the convergence speed of the algorithm[4]. This effectively allows us to optimize over no more than five parameters at a time.

Firms are unlikely to disclose all possible customer metrics within their quarterly and annual filings because of the perceived costs of disclosure (Lev (1992)). Therefore, we must understand how the predictive validity of the model varies as a function of the size and composition of the collection of metrics used to train the model, and as a function of contextual factors. We generate many different data sets or "worlds" reflecting different patterns of customer acquisition, purchasing, and latent attrition by varying the values of the acquisition and repeat purchase processes. We vary six key inputs to these processes in a full factorial design to create $N_W = 64$ such data sets. We generate each dataset for an initial prospect pool of 10MM customers which grows at an annualized growth rate of 1.2% per year. We observe the behavior of this prospect pool over a six-year ($Q = 24$ quarter) period, which is used to train the model. To test the robustness of the metric collections to missing data, we left censor the data, deleting the first year of activity. This leaves us with 20 quarterly disclosures for each metric which is included in the metric collection, corresponding to the 5 years of uncensored quarterly data for the repeat purchase metrics and quarterly customers acquired. We estimate the parameters of the acquisition and repeat purchase processes with this data.

We consider all $N_C = 63$ possible combinations of the 6 repeat purchase metrics – all collections from size one (i.e., each of the metrics individually) to size six (i.e., all of the available metrics at once). We assess metric pair performance within a particular data set based upon how well that metric pair predicts aggregate incremental purchases within a holdout period, which is a managerially relevant quantity linked to customer-based corporate valuation. We performed a similar analysis to assess parameter recovery and the results are qualitatively similar. The error measure that we choose is holdout mean absolute percentage error (MAPE) over the next five years ($Q^* = 20$ quarters). The results are robust to other error measures. Formally, letting $\mathrm{P}_q^{(k)}$ denote the total observed number of purchases made within quarter q of data set k, and $\widehat{\mathrm{P}}_q^{(k,c)}$ the expected number of repeat purchases within quarter q of data set k using metric collection c, the MAPE is equal to

$$\text{MAPE}_{kc} = \frac{1}{Q^*} \sum_{q=Q+1}^{Q+Q^*} \left| \frac{\text{P}_q^{(k)} - \widehat{\text{P}}_q^{(k,c)}}{\text{P}_q^{(k)}} \right|, \quad c = 1, 2, \ldots, N_C, \;\; k = 1, 2, \ldots, N_W. \tag{15}$$

We obtain the expected MAPE corresponding to each metric collection c over each data set k. These estimates form the basis for our assessment of the performance of these metrics. For details regarding the simulation design, see Appendix C.

As a first overall assessment, we provide in Figure 2 a scatterplot of the average MAPE for each collection and each particular size, averaged across the 64 data sets. The average MAPE associated with the best size-one collection is 15.6%. The average MAPE of the best size-two collection is considerably lower, at .7%, and there is no further improvement when we move to larger collections (the best collections of size three to six all have average MAPE values of approximately .6%). This suggests that we need no more than two quarterly repeat purchase metrics to achieve adequate predictive performance. We also observe that there are size-two collections which have significantly smaller average MAPE values than other size-four and even size-five collections. This implies that the right pair of metrics has better predictive performance than other collections which are far larger in size.

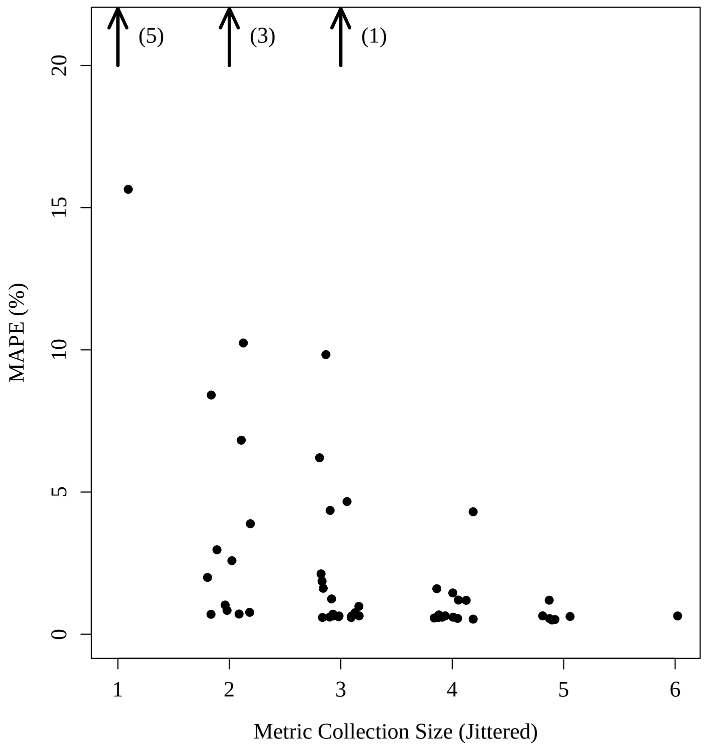
We explore the various size-two collections next. To better visualize which pairs have the best prediction accuracy, we plot in Figure 3 the MAPE of all 15 size-two collections, averaged across all 64 data sets. The results are shown in Figure 3[5], ordered from highest to lowest average MAPE, along with 95% standard error intervals.

Five size-two pairs have excellent overall predictive accuracy (average MAPE values less than or equal to 1%), and another four pairs have very good overall predictive accuracy (MAPE values between 1% and 3%). All other pairs have average MAPE figures exceeding 5%, often significantly so.

The five best-performing pairs consist of F coupled with each of the other metrics, which suggests that F is the most informative metric in a marginal sense. The five F-pairs are also very robust – the maximum MAPE across all data sets is less than 4% for all F-pairs.

While the conceptual distinction between RBPO and RBPC may seem slight, there are

Figure 2: Average MAPE (%) of Each Collection by Collection Size (Count of Collections Exceeding 20% Average MAPE)

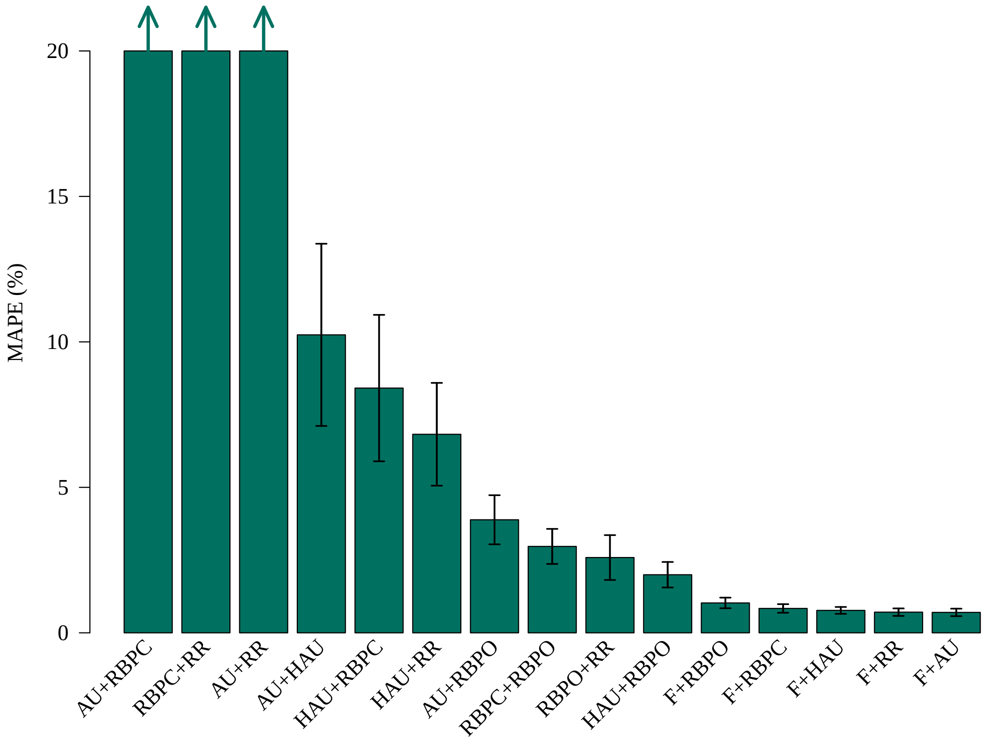


significant differences in performance between these two measures. RBPO is an included in the fifth through ninth best pairs by average MAPE. In contrast, RBPC is in three of the five worst pairs by average MAPE, including two pairs with average MAPE values in excess of 45%. RBPO is a significantly more informative metric than RBPC in a marginal sense.

The quality of metrics within a particular collection is significantly more important for predictive accuracy than the quantity of metrics. For example, all but one of the size-two F-pairs have statistically significantly lower average MAPE values than the size-five collection which does not include F (AU+HAU+RBPC+RBPO+RR has an average MAPE of 1.2%, and the t-statistics of associated paired t-tests range from -2.72 to -3.74). Furthermore, some collections which are large in size have low predictive accuracy in an absolute sense (e.g., the average MAPE of AU+RBPC+RR is 46.3%).

This analysis suggests that F is the most informative of the proposed metrics to investors, in terms of its ability to improve overall purchasing (and thus revenue) forecasting accuracy.

Figure 3: MAPE (%) by Size-Two Metric Collection, Averaged Across Data Sets



F is popularly used and studied by marketing scholars, often in conjunction with penetration/reach (Cheong, De Gregorio, and Kim (2010), Danaher (2007), Sharp and Sharp (1997)). Despite this, F is one of the lesser-disclosed common marketing metrics by public firms (see Table 1). We recommend that firms disclose F in conjunction with AU. While F is less popularly disclosed, AU is already widely adopted by firms and discussed by financial professionals. F and AU is closely related to so-called "means and zeroes" parameter estimation (Hardie and Fader (2000)). This combination of metrics has the highest predictive accuracy in this simulation study, is small in number, and is an intuitive pair of key performance indicators for firms. AU summarizes purchase breadth (how well the firm is able to get a large number of buyers), while F summarizes purchase depth (how well the firm is able to get buyers to buy frequently). Furthermore, these metrics allow external stakeholders to obtain total purchases over time when we observe that total purchases are the product of F with AU. This makes F+AU particularly compatible with investors' traditional financial models, which may decompose revenues into total purchases and average revenue per purchase, then further decompose total purchases into active customers and average purchases per active customer.

EMPIRICAL ANALYSIS

We now apply the method developed over the previous sections to a data set of purchases from a large geographic business unit of an e-commerce retailer in the apparel industry. The data set consists of all disaggregate purchase data for the business unit since the beginning of commercial operations at the end of June 2010. The data set ends at the end of December 2015, so the observation period is 5.5 years (22 quarters) in length. Based upon the results from the large-scale simulation, we implement the estimation procedure using our recommended collection of customer metrics – gross customer acquisitions, active users, average purchase frequency, median spend, and total revenues. The results are virtually identical when we estimate the model upon all repeat purchase metrics. We split the 22-quarter data set into two periods – a $Q = 18$ quarter calibration period which is used to estimate the model, and a $Q^* = 4$ quarter holdout period which is used to assess the predictive validity of the model. The retailer disclosed the breakdown of their fixed and variable costs (FC and VC), their future expected subscriber acquisition costs (CAC), their weighted average cost of capital (WACC), and the size of the applicable market (POP). We summarize these projections when we value the business unit.

Model Assessment

Validation of acquisition and repeat purchase model fit focuses upon a series of six commonly used in-sample and out-of-sample goodness-of-fit diagnostic plots (Braun, Schweidel, and Stein (2015), Fader, Hardie, and Lee (2005a), Fader, Hardie, and Shang (2010), Feit, Wang, Bradlow, and Fader (2013)). These plots summarize salient aspects of the data and test our ability to model and project them. Expected quantities of interest are formed by simulating K datasets using the estimated model hyperparameters, computing the desired quantity of interest each dataset, then averaging across datasets (we set $K = 1000$)).

Parameter Estimates and Evaluation of Fit

We begin with the customer acquisition process. The fitted model rejected heterogeneity in the propensity to spend across customers ($\widehat{r}_A = 9.1\text{E}^9$). There was notable quarterly seasonality, with an uptick in calendar Q4 customer acquisitions. The estimated model parameters of the resulting Weibull with covariates model (and their associated standard errors) are $\widehat{\lambda}_A = 6.1\text{E}^{-5}$ (1.7E^{-6}), $\widehat{c}_A = 1.15$ $(.008)$, and $\widehat{\boldsymbol{\beta}}_A = .18$ $(.006)$. Figure 4 shows expected and actual quarterly customer acquisitions on a quarter-by-quarter basis (left) and on a cumulative basis (right). The vertical dotted line denotes the beginning of the holdout period. While there is noise from quarter to quarter, we model and predict the flow of acquisitions well over time.

Turning to the latent attrition and repeat purchase process, seasonality was evident in the second calendar quarter for repeat purchases, so we include a seasonal covariate to account for it. The estimated BG/ML model parameters are $\widehat{\mu}_p = -3.92$ $(.042)$, $\widehat{\sigma}_p^2 = 1.67$ $(.036)$, $\widehat{\boldsymbol{\beta}}_p = .06$ $(.001)$, $\widehat{\gamma} = .53$ $(.014)$, and $\widehat{\delta} = 3.65$ $(.873)$. The actual data shown in the validation plots are constructed using the granular data. On the left within Figure 5, we show the expected and actual number of people making 0, 1, ..., $10+$ repeat purchases during the calibration period. We correctly infer that the majority of customers (71%) are "one and done," making only their initial purchase and no subsequent repeat purchases. This suggests that the company's overall valuation is currently reliant upon a relatively small number of highly engaged customers. We will study this in more detail in the next section.

On the right within Figure 5, we plot the expected and actual number of purchases a customer will make in the holdout period, conditional upon the number of purchases the customer makes within the calibration period. We observe that the BG/ML model generates accurate predictions of expected behavior in the holdout period conditional upon what has occurred within the calibration period.

Figure 4: Validating Incremental (left) and Cumulative (right) Gross Acquisitions

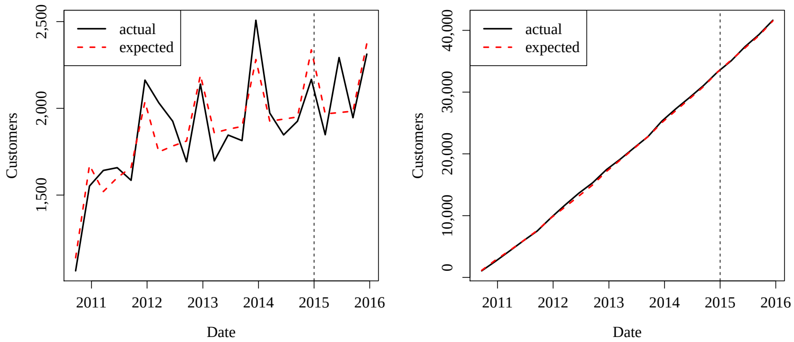


Figure 5: Predicted vs. Actual Frequency of Repeat Transactions (left), Conditional Expected Purchases (right)

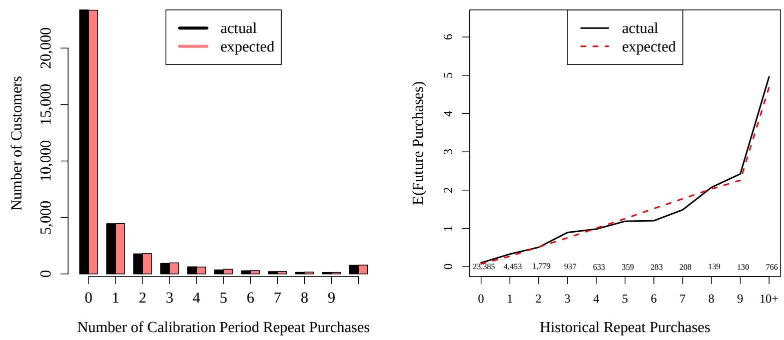
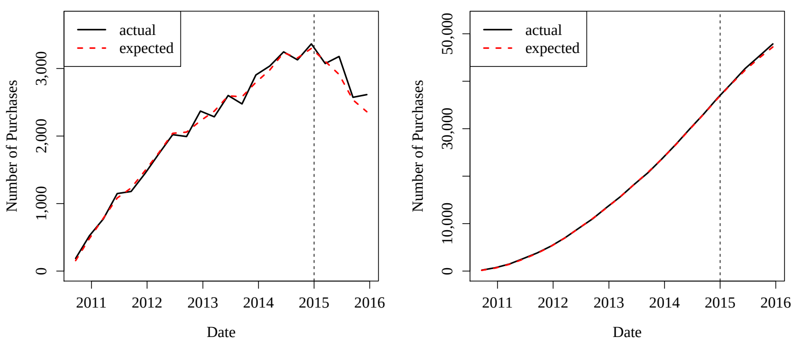


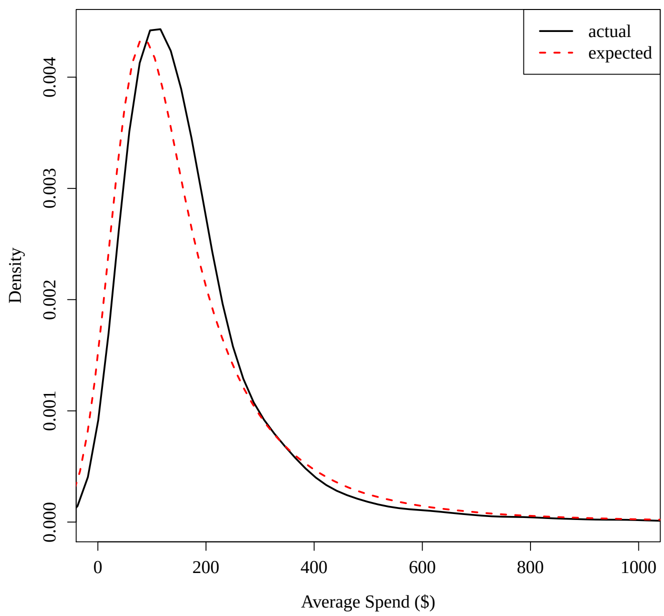
Figure 6: Validating Incremental (left) and Cumulative (right) Repeat Purchases



In Figure 6, we examine the flow of repeat purchases over time. The plot on the left corresponds to the expected and actual number of repeat purchases quarter-by-quarter, during the calibration period and the holdout period. The plot on the right shows the expected and actual number of repeat purchases cumulatively over time. The model generates good predictions of aggregate repeat purchase behavior longitudinally over time. Note that repeat purchases generally increase during the calibration period because the introduction of newly acquired customers during the calibration period (Figure 4) more than offsets the decline in purchases due to latent attrition. Repeat purchases decline during the holdout period because we only make projections in the holdout period for customers acquired during the calibration period.

The final step is to validate the spend model. The estimated model parameters are $\widehat{\mu}_0 = 4.78\ (.018)$, $\widehat{\mu}_q = 1.63\ (.065)$, and $\widehat{\tau}^2 = .77\ (.009)$. To visualize the fit of the model, in Figure 7 we compute the expected marginal distribution of average spend across customers, and compare it to the empirical density of the observed average spends across customers. The resulting fit is plotted in Figure 7. While the estimated modal spend amount is slightly less than the observed modal spend, we capture the overall shape of the distribution, including its long right tail. Figure 7 reinforces the need to model heterogeneity in spend across customers.

Figure 7: Marginal Distribution of Average Spend Across Customers



Model Comparison

Although the analysis so far shows that the in-sample and out-of-sample fit of our proposed model is very good, it does not provide us with insight into how our model's performance compares to alternative models. In this section, we compare our performance to two benchmark models.

We compare the proposed model to the methodology used in GLS. We also compare the proposed model to an extension of GLS (hereafter, GLS+) which allows for (1) seasonal fluctuation in the customer acquisition process, (2) time trend and seasonality covariates in the number of purchases made per active customer. In addition to validating in-sample and out-of-sample fit, we provide overall valuation estimates for the three models in the next section.

We summarize the relative performance of these three models in Table 3, which reports the mean absolute percentage error (MAPE) for (1) total quarterly customers acquired and (2) total quarterly purchases made by existing customers over the calibration and holdout periods.

	Customer Acquisitions			Total Purchases		
Model	Overall	OOS	IS	Overall	OOS	IS
Proposed	6.32	6.86	6.20	3.94	5.46	3.60
GLS+	26.20	53.41	19.80	13.56	15.25	13.16
GLS	29.12	58.84	22.13	41.59	25.08	45.48

Table 3: E-commerce Retailer: MAPE of Quarterly Customer Acquisitions and Total Purchases by Time Frame

The in-sample and out-of-sample fits of the proposed model are better than the benchmarks. We find that GLS underestimates future customer acquisitions. While allowing for seasonality in GLS's acquisition model reduces its error, the level of error remains high. This is because the Bass-like technological substitution model (TSM) for customer acquisition infers that customer acquisition has already hit its peak. Because the TSM must be symmetric about the period of peak acquisition (as with the Bass model), GLS predicts that future acquisitions will fall quickly in the holdout period.

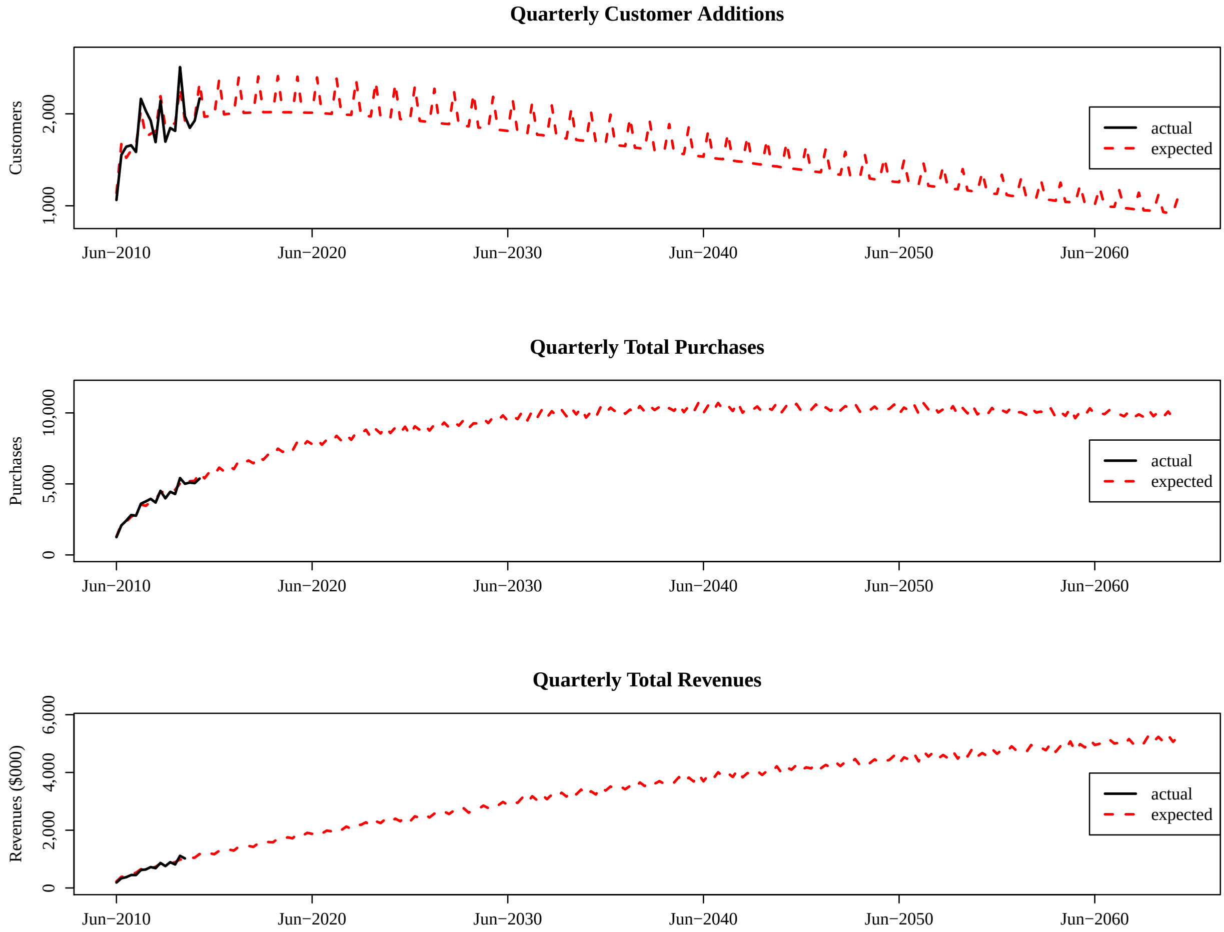
GLS underestimates future total purchases more than it underestimates future customer acquisitions, with MAPE values in excess of 25% in-sample and out-of-sample. GLS assumes a homogeneous retention rate which is equal to the then-current repeat rate. Because the vast majority of customers make only one purchase (Figure 5) the annual repeat rate is 25.9%, which quickly "kills off" active customers as we move forward into the holdout period. In reality, there is a considerable amount of heterogeneity in customers' propensity to remain with the firm, as evidenced by the repeat purchase parameter estimates. While most customers will churn quickly, some customers will remain with the firm for a long time. This underestimation is mitigated but not eliminated with the GLS+ model, which allows the number of purchases per active customer to trend upwards over time. For a more detailed analysis of these alternative specifications, see Appendix D.

Valuation

Following the process we outlined in the valuation procedure section, we move from a one-year forecast to 50-year ($2,600$-week) forecast, which is far enough into the future that the time value of money makes the incremental impact of all future profit or loss upon overall valuation negligible. In Figure 8, we plot actual and expected quarterly customer acquisitions (top panel), total purchases (middle panel), and revenues (bottom panel). The acquisitions forecasts in Figure 8 are formed by "compressing" the observed acquisition data in Figure 4 to make room for a 50-year holdout period. These valuation results are not sensitive to the estimated total number of prospects, which was provided to us by the firm.

The weekly customer acquisition estimates over the 50-year forecasting horizon after the 4.5-year calibration period (A(235),A(236),$\ldots$,A(2834)), and the corresponding revenue estimates over the same period (R(235),R(236),$\ldots$,R(2834)) are the primary customer model-driven inputs to future FCF using Equation 4. Projections of other inputs needed to estimate SHV in Equation 2 were provided to us by the company's management team. We use these figures to obtain weekly FCF projections over the forecasting horizon, the net present value of the resulting FCF stream, and finally, shareholder value (SHV). We estimate that SHV is equal to \$22.8MM. The corresponding estimated valuations using GLS and GLS+ are \$2.6MM and \$2.8MM, respectively. We provide a summary of the inputs

Figure 8: E-commerce Retailer: Summary of Projections



driving the overall valuation in Table 4.

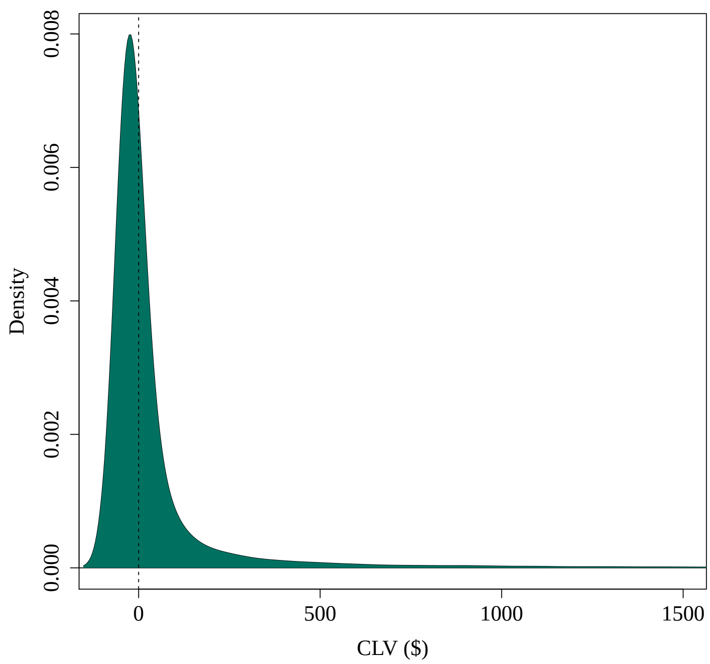
Table 4: Summary of Valuation

Total Initial Prospects (POP(0))	500,000
Annual Growth Rate of Prospects	0%
Variable contribution margin (VC)	76.4%
Weekly fixed costs (FC)	$4,260
Cost per acquired customer (CAC)	$76
Tax rate (TR)	35%
Weighted average cost of capital (WACC)	6%
Non-operating assets - Net debt (NOA - ND)	$0
Shareholder value (SHV)	$22.8MM

Model Implications

In addition to estimating of the overall valuation of the business unit, we may also use our model to make inferences into the value of newly acquired customers. The average customer lifetime value (CLV) of a newly acquired cohort of customers will be $76.2 per customer (a return of approximately 100% over the customer's acquisition cost), but that there will be substantial variability in profitability across these customers. The expected distribution of CLV across newly acquired customers is shown in Figure 9.

Figure 9: Expected Distribution of CLV for Newly Acquired Customers



The overall valuation of the business unit is driven by a small number of customers. For example, we predict that 68% of newly-acquired customers will not be profitable and just 2.9% of newly-acquired customers will generate 80% of their collective value.

We may also infer how much of the overall valuation of the firm will come from existing customers versus not-yet-acquired customers. We estimate that the net present value of all future profits expected from the existing customer base (current customer equity or CCE) is $3.5MM, or approximately 15% of the $22.8MM overall valuation of the firm. This relatively low percentage is to be expected, given the relatively young age of the business unit. CCE is predicted to represent an increasing proportion of the overall valuation of the business unit in the future, as the existing customer base is increasingly comprised of customers with high retention and purchase propensities.

The business unit's valuation will be a "concentrated growth story." The chief valuation driver is whether the business unit can continue acquiring "die-hard" customers, whose CLV's exceed $250 on average, because these highly engaged customers will generate virtually all of the business unit's profits and recoup the losses that the firm will incur on the 68% of acquired customers who will be unprofitable.

These results also suggest that the business unit may increase its valuation if it is able to find a way to reduce the number of unprofitable customers and/or acquire fewer like them, because they currently represent a rather large proportion of newly-acquired customers. When we spoke with an executive at the retailer about this, he noted that the relatively large number of unprofitable customers are most likely due to international shipping issues (e.g., delays and errors) causing new customers to sour on the company after their initial purchases and not make repeat purchases. Fixing these shipping issues could reduce the large number of "one and done" customers in future acquisition cohorts, which could increase the overall valuation of the business if the incremental value created exceeds the cost of the fix.

DISCUSSION

This article proposes a novel methodology with which common customer metrics are used to estimate a latent variable model for non-contractual customer acquisition, repeat purchase, and spend. We not only show that some collections of customer metrics can reliably identify the customer model, but also which collections are the most predictive. The customer model is used to drive an overall valuation model for non-contractual firms, provide helpful color about where the value is coming from, and how much individual customers are worth.

The methodology has uses which extend beyond corporate valuation. These techniques may be useful for expert testimony in litigation cases where firms would like to provide enough information to confirm or deny specific points raised within the case, but no more than that. Furthermore, it may often be the case that external stakeholders are primarily interested in sales forecasts (e.g., for forecasts of overall economic activity, or for valuation via a multiple of future sales (Liu, Nissim, and Thomas, 2002)), which the proposed methodology provides.

The methodology may also be useful in industries that would not usually be amenable to such an analysis. For example, it may be the case for consumer packaged goods firms that customer acquisition figures are not available, but that a larger panel of repeat purchase metrics is. The right collection of repeat purchase metrics may identify the acquisition process as well. We leave these topics to future work.

One of the limitations of this work is that we have assumed that the disclosure decision is not strategic. It could be that there is a forward-looking component to firms' decision to disclose (Mintz et al. (2016)). The lack of non-contractual companies disclosing metrics makes it impossible to verify whether or not this is the case empirically, and our primary goal in this article is to study the informativeness of metrics is a valuation setting, but the methodology presented here would not preclude a study of this sort if the data were available. Another limitation of this work is that our valuation predictions would be highly uncertain for firms with very little historical data. Our approach is not immune to the challenges one may face when forecasting new product sales or the diffusion of a new innovation when data is available over a short window of time. However, while the acquisition process may be highly uncertain in these cases, firms can still learn about the valuation of current customers using the proposed repeat purchase and spend processes.

This work highlights an important under-appreciated use for common customer metrics. These metrics need not be ends in their own right (i.e., as standalone firm-wide key performance indicators) – they can be leveraged to better understand the true underlying propensity of customers to acquire services, make purchases, and spend, and how these propensities vary across customers. The proposed methodology turns backward-looking customer metrics into important forward-looking measures, which should decrease investor uncertainty regarding future cash flows and thus increase the valuation of the firm (Bayer, Tuli, and Skiera (2017)). Moreover, while it is currently in vogue to work with ever-larger datasets, this work highlights just how much we can make that data smaller while still being able to make accurate future revenue projections. A small quantity of the right data can be far more useful than a large quantity of the wrong data.

As companies, business intelligence firms, and investors realize these other uses for customer metrics, we believe the demand for their disclosure will continue to grow. This would not be the first time – one of the most commonly disclosed and tracked retail metrics, same store sales (SSS), became popular because a Wall Street analyst showed just how useful it was at uncovering the true underlying financial condition of a fast-growing retailer in the 1970's (Blumenthal (2008)). We believe the right set of customer metrics could allow investors to track the quality of existing customers much the same way that SSS allows investors to track the quality of existing stores. With physical stores ceding share to internet-based retail, the need for such a metric is more important than ever.

REFERENCES

Abe, Makoto. 2009. Counting your customers one by one: A hierarchical bayes extension to the Pareto/NBD model. *Marketing Science* **28**(3) 541–553.

Albuquerque, Paulo, Bart J Bronnenberg. 2009. Estimating demand heterogeneity using aggregated data: An application to the frozen pizza category. *Marketing Science* **28**(2) 356–372.

Arango, Timothy. 2001. Exec says publisher declined to extend Amazon's credit terms. URL `https://news.thestreet.com/story/1340697/1/exec-says-publisher-declined-to-extend-amazons-credit-terms.html`.

Bain and Company. 2000. The value of online customer loyalty. URL `http://www.bain.com/Images/Value_online_customer_loyalty_you_capture.pdf`.

Bauer, Hans H, Maik Hammerschmidt. 2005. Customer-based corporate valuation: Integrating the concepts of customer equity and shareholder value. *Management Decision* **43**(3) 331–348.

Bayer, Emanuel, Kapil R Tuli, Bernd Skiera. 2017. Do disclosures of customer metrics lower investors' and analysts' uncertainty, but hurt firm performance? *Journal of Marketing Research (forthcoming)* .

Besanko, David, Sachin Gupta, Dipak Jain. 1998. Logit demand estimation under competitive pricing behavior: An equilibrium framework. *Management Science* **44**(11-1) 1533–1547.

Blattberg, Robert C, John Deighton. 1996. Manage marketing by the customer equity test. *Harvard Business Review* **74**(4) 136.

Blattberg, Robert C, Gary Getz, Jacquelyn S Thomas. 2001. *Customer equity: Building and managing relationships as valuable assets*. Boston, MA: Harvard Business Press.

Blumenthal, Karen. 2008. *Grande Expectations: A Year in the Life of Starbucks' Stock*. New York, NY: Crown Publishing.

Braun, Michael, David A Schweidel, Eli Stein. 2015. Transaction attributes and customer valuation. *Journal of Marketing Research* **52**(6) 848–864.

Chen, Yuxin, Sha Yang. 2007. Estimating disaggregate models using aggregate data through augmentation of individual choice. *Journal of Marketing Research* **44**(4) 613–621.

Cheong, Yunjae, Federico De Gregorio, Kihan Kim. 2010. The power of reach and frequency in the age of digital advertising. *Journal of Advertising Research* **50**(4) 403–415.

Chintagunta, Pradeep K. 1992. Estimating a multinomial probit model of brand choice using the method of simulated moments. *Marketing Science* **11**(4) 386–407.

Damodaran, Aswath. 2012. *Investment valuation: Tools and Techniques for Determining the Value of Any Asset*. Hoboken, N.J.: John Wiley & Sons.

Danaher, Peter J. 2007. Modeling page views across multiple websites with an application to internet reach and frequency prediction. *Marketing Science* **26**(3) 422–437.

Datta, Hannes, Bram Foubert, Harald J Van Heerde. 2015. The challenge of retaining customers acquired with free trials. *Journal of Marketing Research* **52**(2) 217–234.

Duffee, Gregory R, Richard H Stanton. 2008. Evidence on simulation inference for near unit-root processes with implications for term structure estimation. *Journal of Financial Econometrics* **6**(1) 108–142.

Dutta, Shantanu, Om Narasimhan, Surendra Rajiv. 1999. Success in high-technology markets: Is marketing capability critical? *Marketing Science* **18**(4) 547–568.

Ebbes, Peter, Dominik Papies, Harald J Van Heerde. 2011. The sense and non-sense of holdout sample validation in the presence of endogeneity. *Marketing Science* **30**(6) 1115–1122.

Erdem, Tülin. 1996. A dynamic analysis of market structure based on panel data. *Marketing science* **15**(4) 359–378.

Fader, Peter S, Bruce GS Hardie. 2009. Probability models for customer-base analysis. *Journal of Interactive Marketing* **23**(1) 61–69.

Fader, Peter S, Bruce GS Hardie. 2010. Customer-base valuation in a contractual setting: The perils of ignoring heterogeneity. *Marketing Science* **29**(1) 85–93.

Fader, Peter S, Bruce GS Hardie. 2014. The Pareto/NBD is not a lost-for-good model. Working paper.

Fader, Peter S, Bruce GS Hardie, Ka Lok Lee. 2005a. "Counting your customers" the easy way: An alternative to the Pareto/NBD model. *Marketing Science* **24**(2) 275–284.

Fader, Peter S, Bruce GS Hardie, Ka Lok Lee. 2005b. RFM and CLV: Using iso-value curves for customer base analysis. *Journal of Marketing Research* **42**(4) 415–430.

Fader, Peter S, Bruce GS Hardie, Jen Shang. 2010. Customer-base analysis in a discrete-time noncontractual setting. *Marketing Science* **29**(6) 1086–1108.

Farris, Paul W, Neil T Bendle, Phillip E Pfeifer, David J Reibstein. 2010. *Marketing metrics: The definitive guide to measuring marketing performance*. Pearson Education.

Feit, Eleanor McDonnell, Pengyuan Wang, Eric T Bradlow, Peter S Fader. 2013. Fusing aggregate and disaggregate data with an application to multiplatform media consumption. *Journal of Marketing Research* **50**(3) 348–364.

Gallant, A Ronald, George Tauchen. 1996. Which moments to match? *Econometric Theory* **12** 657–681.

Gönül, Füsun, Kannan Srinivasan. 1993. Modeling multiple sources of heterogeneity in multinomial logit models: Methodological and managerial issues. *Marketing Science* **12**(3) 213–229.

Gourieroux, Christian, Alain Monfort, Eric Renault. 1993. Indirect inference. *Journal of Applied Econometrics* **8**(S1) S85–S118.

Greenwald, Bruce, Judd Kahn, Paul D Sonkin, Michael Van Biema. 2004. *Value Investing: From Graham to Buffett and Beyond*. Hoboken, N.J.: John Wiley & Sons.

Gupta, Sunil, Donald R Lehmann, Jennifer Ames Stuart. 2004. Valuing customers. *Journal of Marketing Research* **41**(1) 7–18.

Hansen, Lars Peter. 1982. Large sample properties of generalized method of moments estimators. *Econometrica: Journal of the Econometric Society* 1029–1054.

Hardie, Bruce, P Fader. 2000. Applied probability models in marketing research. Working paper.

Holthausen, Robert W, Mark E Zmijewski. 2013. *Corporate Valuation: Theory, Evidence and Practice*. Cambridge, MA: Cambridge Business Publishers.

Jiang, Wenxin, Bruce Turnbull. 2004. The indirect method: inference based on intermediate statistics—a synthesis and examples. *Statistical Science* **19**(2) 239–263.

Kinshuk, Jerath, Peter S Fader, Bruce GS Hardie. 2016. Customer-base analysis using repeated cross-sectional summary (rcss) data. *European Journal of Operational Research* **249**(1) 340–350.

Koller, Tim, Marc Goedhart, David Wessels. 2010. *Valuation: Measuring and Managing the Value of Companies*. Hoboken, N.J.: John Wiley & Sons.

Lev, Baruch. 1992. Information disclosure strategy. *California Management Review* **34**(4) 9–32.

Libai, Barak, Eitan Muller, Renana Peres. 2009. The diffusion of services. *Journal of Marketing Research* **46**(2) 163–175.

Liu, Jing, Doron Nissim, Jacob Thomas. 2002. Equity valuation using multiples. *Journal of Accounting Research* **40**(1) 135–172.

Maycotte, H.O. 2015. Customer lifetime value – the only metric that matters. URL `http://www.forbes.com/sites/homaycotte/2015/08/25/customer-lifetime-value-the-only-metric-that-matters/#1350aa975886`.

McCarthy, Daniel M, Peter S Fader, Bruce GS Hardie. 2017. Valuing subscription-based businesses using publicly disclosed customer data. *Journal of Marketing* **81**(1) 17–35.

McFadden, Daniel. 1989. A method of simulated moments for estimation of discrete response models without numerical integration. *Econometrica: Journal of the Econometric Society* **57**(5) 995–1026.

Mehta, Nitin, Surendra Rajiv, Kannan Srinivasan. 2003. Price uncertainty and consumer search: A structural model of consideration set formation. *Marketing science* **22**(1) 58–84.

Michaelides, Alexander, Serena Ng. 2000. Estimating the rational expectations model of speculative storage: A monte carlo comparison of three simulation estimators. *Journal of Econometrics* **96**(2) 231–266.

Mintz, Ofer, Timothy J Gilbride, Imran S Currim, Peter Lenk. 2016. Metric effectiveness and use in marketing-mix decisions: Correcting for endogenous selection effects and ex-ante expectations. Working paper.

Moe, Wendy W, Peter S Fader. 2002. Using advance purchase orders to forecast new product sales. *Marketing Science* **21**(3) 347–364.

Morrison, Donald G, David C Schmittlein. 1980. Jobs, strikes, and wars: Probability models for duration. *Organizational Behavior and Human Performance* **25**(2) 224–251.

Musalem, Andres, Eric T Bradlow, Jagmohan S Raju. 2008. Who's got the coupon? Estimating consumer preferences and coupon usage from aggregate information. *Journal of Marketing Research* **45**(6) 715–730.

Musalem, Andres, Eric T Bradlow, Jagmohan S Raju. 2009. Bayesian estimation of random-coefficients choice models using aggregate data. *Journal of Applied Econometrics* **24**(3) 490–516.

Neslin, Scott A. 1990. A market response model for coupon promotions. *Marketing Science* **9**(2) 125–145.

QVC. 2015. 10-K. URL `http://www.sec.gov/Archives/edgar/data/1254699/000125469916000050/qvc_10kx12312015.htm`.

Rust, Roland T, Katherine N Lemon, Valarie A Zeithaml. 2004. Return on marketing: Using customer equity to focus marketing strategy. *Journal of Marketing* **68**(1) 109–127.

Sabavala, Darius J, Donald G Morrison. 1977. A model of TV show loyalty. *Journal of Advertising Research* **17**(6) 35–43.

Schmittlein, David C, Donald G Morrison, Richard Colombo. 1987. Counting your customers: Who are they and what will they do next? *Management science* **33**(1) 1–24.

Schmittlein, David C, Robert A Peterson. 1994. Customer base analysis: An industrial purchase process application. *Marketing Science* **13**(1) 41–67.

Schulze, Christian, Bernd Skiera, Thorsten Wiesel. 2012. Linking customer and financial metrics to shareholder value: The leverage effect in customer-based valuation. *Journal of Marketing* **76**(2) 17–32.

Schweidel, David A, Peter S Fader, Eric T Bradlow. 2008. Understanding service retention within and across cohorts using limited information. *Journal of Marketing* **72**(1) 82–94.

Schweidel, David A, George Knox. 2013. Incorporating direct marketing activity into latent attrition models. *Marketing Science* **32**(3) 471–487.

Seybold, Patricia. 2000. Don't count on Amazon. *Business 2.0* URL `http://boards.fool.com/business-20-dont-count-out-amazon-13436734.aspx?sort=postdate`.

Sharp, Byron, Anne Sharp. 1997. Loyalty programs and their impact on repeat-purchase loyalty patterns. *International journal of Research in Marketing* **14**(5) 473–486.

Smith, Anthony A. 1993. Estimating nonlinear time-series models using simulated vector autoregressions. *Journal of Applied Econometrics* **8**(S1) S63–S84.

Streitfeld, David. 2001. Analysts, vendors increasingly wary about Amazon. URL `https://www.washingtonpost.com/archive/business/2001/02/21/analysts-vendors-increasingly-wary-about-amazon/82aee135-065e-4ce9-a293-edbbc83a6552`.

Wiesel, Thorsten, Bernd Skiera, Julian Villanueva. 2008. Customer equity: An integral part of financial reporting. *Journal of Marketing* **72**(2) 1–14.

Wooldridge, J. 2002. *Introductory Econometrics Analysis of Cross Section and Panel Data*. Cambridge, MA: The MIT Press.

Zhang, Yao, Eric T Bradlow, Dylan S Small. 2014. Predicting customer value using clumpiness: From RFM to RFMC. *Marketing Science* **34**(2) 195–208.

Notes

[1]Instead of disclosing the number of customers lost during the period, the firm may equivalently disclose the per-period churn rate, or the total number of customers that are active as of the end of each period.

[2]The other three businesses valued – Ameritrade, Capital One, and E*Trade – are all contractual as first defined by Schmittlein, Morrison, and Colombo (1987) because their churn is observable. Customers are required to maintain positive account balances to remain customers.

[3]This is the so-called "distance"-based II estimator of Gourieroux, Monfort, and Renault (1993) and Smith (1993). There is also a "score"-based II estimator (Gallant and Tauchen (1996)). Studies have shown that the finite-sample performance of the score-based estimator is not as strong (Duffee and Stanton (2008), Michaelides and Ng (2000))

[4]We estimate the parameters of the acquisition process using the procedure above, then estimate the repeat purchase process given the acquisition parameter estimates, then estimate the spend process given the acquisition and repeat purchase estimates, then estimate the processes jointly.

[5]The average MAPE values for AU+RR, RBPC+RR, and AU+RBPC are 45.6%, 47.4%, and 201.0%, respectively. The y-axis of the plot is truncated so that these pairs do not impede our ability to summarize the performance of the other 12 pairs.

A ACQUISITION PARAMETER RECOVERY ANALYSIS

We test our ability to recover the parameters of the acquisition model across a wide variety of data generating processes. We consider $3^3 = 27$ different scenarios: $r_A \in (0.5, 1.0, 1.5)$, median time to acquisition of 10, 15, or 20 years; and $c_A \in (1.0, 1.25, 1.50)$. We simulate five years of weekly data, aggregate the weekly data to report acquisition data every quarter, and estimate the parameters by MLE. We assume an initial population of 1MM households and 1.2% annual growth in the population. For each scenario considered, we report the mean absolute percent error (MAPE) for each parameter, as well as the overall MAPE averaged across the parameters. Results are averaged over 100 replications and are shown in Table 5.

Table 5: True Parameters, Estimated Parameters, and MAPE

True Parameters			Estimated Parameters			MAPE (%)			
r_A	α_A	c_A	r_A	α_A	c_A	r_A	α_A	c_A	Overall
0.5	173.3	1.00	.50	173.36	1.00	.40	.27	.09	.25
1.0	520.0	1.00	1.00	519.86	1.00	.83	.60	.09	.51
1.5	885.3	1.00	1.50	884.78	1.00	1.30	1.04	.09	.81
.5	260.0	1.00	.50	260.02	1.00	.56	.38	.10	.35
1.0	780.0	1.00	1.00	779.68	1.00	1.28	1.01	.10	.80
1.5	1327.9	1.00	1.50	1327.99	1.00	2.13	1.82	.11	1.35
.5	346.7	1.00	.50	346.68	1.00	.73	.50	.11	.45
1.0	1040.0	1.00	1.00	1039.77	1.00	1.86	1.54	.12	1.17
1.5	1770.5	1.00	1.50	1771.16	1.00	2.97	2.65	.12	1.91
.5	827.7	1.25	.50	827.79	1.25	.42	.29	.08	.27
1.0	2483.2	1.25	1.00	2481.84	1.25	.95	.63	.09	.56
1.5	4227.4	1.25	1.50	4224.11	1.25	1.49	1.14	.09	.91
.5	1374.0	1.25	.50	1374.06	1.25	.66	.41	.10	.39
1.0	4122.1	1.25	1.00	4121.58	1.25	1.70	1.30	.11	1.04
1.5	7017.5	1.25	1.50	7018.66	1.25	2.80	2.35	.12	1.75
.5	1968.7	1.25	.50	1968.48	1.25	.94	.60	.11	.55
1.0	5906.0	1.25	1.00	5908.64	1.25	2.71	2.23	.13	1.69
1.5	10054.4	1.25	1.50	10067.73	1.25	4.76	4.18	.14	3.02
.5	3952.6	1.50	.50	3953.21	1.50	.47	.35	.09	.30
1.0	11857.8	1.50	1.00	11853.64	1.50	1.14	.72	.10	.65
1.5	20186.9	1.50	1.50	20177.63	1.50	1.85	1.37	.10	1.11
.5	7261.4	1.50	.50	7260.54	1.50	.88	.50	.11	.50
1.0	21784.2	1.50	1.00	21796.16	1.50	2.45	1.78	.13	1.45
1.5	37085.8	1.50	1.50	37150.16	1.50	4.34	3.57	.14	2.68
.5	11179.7	1.50	.50	11181.92	1.50	1.47	.90	.14	.84
1.0	33539.0	1.50	1.01	33688.46	1.50	4.56	3.72	.16	2.81
1.5	57097.3	1.50	1.52	57699.00	1.50	8.11	7.18	.17	5.15

B AUXILIARY MODEL SPECIFICATIONS

In this section, we specify the auxiliary models for all processes except the acquisition process, which was specified in the section "Acquisition Process Parameter Estimation."

There are two non-parametric auxiliary models associated with the repeat purchase process. One auxiliary model is associated with AU, HAU, and F while the other auxiliary model is associated with RR, RBPC, and RBPO. As with the acquisition process, these non-parametric auxiliary models were chosen to minimize loss of information.

We posit an auxiliary model associated with AU, HAU, and F in which all customers who will eventually be acquired by the end of the calibration period, denoted by A^* (i.e., $\mathrm{A}^* = \sum_{i=1}^{Q} \mathrm{ADD}_q$), are exchangeable with one another and may make purchases in the first quarter of the data. The auxiliary model does not distinguish between initial purchases and repeat purchases. The number of purchases made in quarter q by all customers who will eventually be acquired by the end of the calibration period is determined by a multinomial$[\pi_0^{(p)}(q), \ldots, \pi_{12}^{(p)}(q)]$ draw. $\pi_x^{(p)}(q)$ is the probability that the customer makes x purchases within quarter q, and the multinomial distribution associated with each quarter is independent of other quarters. Assume that we observe an arbitrary collection of n AU disclosures, **AU**, where

$$\mathbf{AU} = [\mathrm{AU}_{q_1}, \mathrm{AU}_{q_2}, \cdots, \mathrm{AU}_{q_n}],$$

whose ith element, AU_{q_i}, is equal to the number of customers who made at least one purchase in the four quarters preceding quarter q_i. Then $\mathrm{AU}_{q_i}/\mathrm{A}^*$ is equal to the complement of the probability of no purchases taking place over the prior four quarters under the auxiliary model:

$$\mathrm{AU}_{q_i}/\mathrm{A}^* = 1 - \widehat{E}\left[\prod_{t=0}^{3} \pi_0^p(q_i - t)\right].$$

Assume that we observe an arbitrary collection of n HAU disclosures, **HAU**, where

$$\mathbf{HAU} = [\mathrm{HAU}_{q_1}, \mathrm{HAU}_{q_2}, \cdots, \mathrm{HAU}_{q_n}],$$

whose ith element, HAU_{q_i}, is equal to the number of customers who made at least two purchases in the four quarters preceding quarter q_i. Then $\mathrm{HAU}_{q_i}/\mathrm{A}^*$ is equal to the complement of the probability of either zero or one purchases taking place over the prior four quarters under the auxiliary model:

$$\mathrm{HAU}_{q_i}/\mathrm{A}^* = 1 - \widehat{E}\left[\prod_{t=0}^{3} \pi_0^p(q_i - t)\right] - \widehat{E}\left[\sum_{s=0}^{3}\prod_{t=0}^{3} \pi_{\mathbb{1}(s=t)}^p(q_i - t)\right],$$

where $\mathbb{1}(s = t)$ is a binary indicator variable equal to one if $s = t$ and zero otherwise.

Finally, assume that we observe an arbitrary collection of n F disclosures, $\mathbf{F}$, where

$$\mathbf{F} = [\mathrm{F}_{q_1}, \mathrm{F}_{q_2}, \cdots, \mathrm{F}_{q_n}],$$

whose ith element, F_{q_i}, is equal to the average number of purchases made over the past 12 months by a customer, given the customer was active over this time period. Using Bayes' theorem, it follows that

$$\mathrm{F}_{q_i} = \widehat{E}\left[\frac{\sum_{t=0}^{3}\sum_{s=1}^{13} s\pi_s^p(q_i - t)}{1 - \prod_{t=0}^{3} \pi_0^p(q_i - t)}\right]$$

We posit a different auxiliary model associated with RR, RBPC, and RBPO. All A^* customers who will eventually be acquired by the end of the calibration period are exchangeable with one another and may make purchases in the first quarter of the data, as with the auxiliary model associated with AU, HAU, and F. However, the number of purchases made each quarter for each customer depends upon whether the customer made at least one purchase in any prior quarter or not. The number of purchases made in quarter q by all customers who have not yet made their first purchase is determined by a multinomial$[\pi_{A,0}^{(p)}(q), \pi_{A,1}^{(p)}(q), \cdots, \pi_{A,13}^{(p)}(q)]$ draw, while the number of purchases made in quarter q by all customers who have made at least one purchase in any prior quarter is determined by a multinomial$[\pi_{B,0}^{(p)}(q), \pi_{B,1}^{(p)}(q), \cdots, \pi_{B,13}^{(p)}(q)]$ draw. Conditional upon the prior purchase history of the customer, the multinomial distributions associated with each quarter and each regime are independent of other quarters.

Assume that we observe an arbitrary collection of n RR disclosures, **RR**, where

$$\mathbf{RR} = [\mathrm{RR}_{q_1}, \mathrm{RR}_{q_2}, \cdots, \mathrm{RR}_{q_n}],$$

whose ith element, RR_{q_i}, is equal to the proportion of customers who made one or more purchases within the past four quarters, given that they made one or more purchases one year ago (i.e., five to eight quarters ago). This is the complement of making no purchases over the past four quarters, given that the customer made a purchase prior to the beginning of the year:

$$\mathrm{RR}_{q_i} = 1 - \widehat{E}\left[\prod_{t=0}^{3} \pi^{p}_{B,0}(q_i - t)\right]$$

Similarly, assume that we observe an arbitrary collection of n RBPC disclosures,

$$\mathbf{RBPC} = [\mathrm{RBPC}_{q_1}, \mathrm{RBPC}_{q_2}, \cdots, \mathrm{RBPC}_{q_n}],$$

whose ith element, RBPC_{q_i}, is equal to the proportion of customers who made one or more purchases before the year began, given that they made one or more purchases within the past year. By Bayes' Theorem, this is equal to the probability that the customer made one or more purchases in both periods divided by the probability that the customer made one or more purchases this year, where the probability that the customer made one or more purchases this year depends upon whether or not the customer made one or more purchases before the year began:

$$\begin{aligned}
\mathrm{RBPC}_{q_i} &= \widehat{E}\left(\frac{A_1}{A_1 + B_1}\right), \quad \text{where} \\
A_1 &= \left[1 - \prod_{t=0}^{3} \pi^{p}_{B,0}(q_i - t)\right]\left[1 - \prod_{t=4}^{q_i-1} \pi^{p}_{A,0}(q_i - t)\right] \quad \text{and} \\
B_1 &= \left[1 - \prod_{t=0}^{3} \pi^{p}_{A,0}(q_i - t)\right]\left[\prod_{t=4}^{q_i-1} \pi^{p}_{A,0}(q_i - t)\right]
\end{aligned}$$

Finally, assume that we observe an arbitrary collection of n RBPO disclosures,

$$\mathbf{RBPO} = [\mathrm{RBPO}_{q_1}, \mathrm{RBPO}_{q_2}, \cdots, \mathrm{RBPO}_{q_n}],$$

whose ith element, RBPO_{q_i}, is equal to the proportion of orders within the past year which were made by customers who made one or more purchases before the year began. This is equal to the expected number of orders made this year by customers who purchased one or more times before the year began, divided by the number of orders made this year:

$$\begin{aligned}
\text{RBPO}_{q_i} &= \widehat{E}\left(\frac{A_2}{A_2+B_2}\right), \quad \text{where} \\
A_2 &= \left[\sum_{t=0}^{3}\sum_{s=1}^{13} s\pi^p_{B,s}(q_i-t)\right]\left[1-\prod_{t=4}^{q_i-1}\pi^p_{A,0}(q_i-t)\right] \quad \text{and} \\
B_2 &= \left[\sum_{t=0}^{3}\sum_{s=1}^{13} s\pi^p_{A,s}(q_i-t)\right]\left[\prod_{t=4}^{q_i-1}\pi^p_{A,0}(q_i-t)\right]
\end{aligned}$$

In summary, **AU**/A*, **HAU**/A*, **F**, **RR**, **RBPC**, and **RBPO** are the auxiliary statistics associated with non-parametric auxiliary models, and have been selected so that they converge asymptotically to (finite-valued) true population-level figures. We obtain the repeat purchase auxiliary parameters corresponding to these auxiliary statistics by repeatedly simulating from the true acquisition and repeat purchase processes.

The spend auxiliary model is the true model itself. The parameters of the spend process are identified by the empirical quarterly mean and median of spend amounts over time. The empirical median of spend amounts each quarter is directly disclosed by the firm. The empirical mean of spend amounts each quarter is known, conditional upon the parameters of the acquisition, latent attrition and repeat purchase processes, in conjunction with total quarterly revenues, which is directly disclosed by the firm.

C LARGE-SCALE SIMULATION DETAILS

In this section, we specify the 64 "worlds" or data sets which we generate data from to evaluate the predictive validity of the proposed repeat purchase customer metric pairs. We obtain the parameter values corresponding to each world through a full factorial design, perturbing the parameter values estimated from a canonical data set which has been studied and benchmarked extensively in the marketing literature, the CDNOW dataset (Abe (2009), Fader, Hardie, and Lee (2005b), Zhang, Bradlow, and Small (2014)).

This dataset consists of 2,357 customers who were acquired in the first quarter of 1997 and observed over a period of 1.5 years. We discretize the data at the weekly level, modeling customers' decision to make at least one purchase each week, and summing all spend amounts which occur during purchase-weeks. There is no seasonal pattern in the data, so we do not include a seasonal covariate in the model.

While it is possible to estimate the parameters of the BG/ML model via maximum marginal likelihood, this is computationally expensive because the marginal likelihood expression is not available in closed-form. Instead, we use latent method of moments estimation. We estimate the parameters of a model which only differs in terms of the parametric distribution used to characterize unobserved heterogeneity in the repeat purchase process but whose marginal likelihood is available in closed-form, the Beta-Geometric/Beta-Binomial model (Fader, Hardie, and Shang (2010)). We then match the moments of the purchase and death heterogeneity distributions of the BG/ML with those of the BG/BB[6].

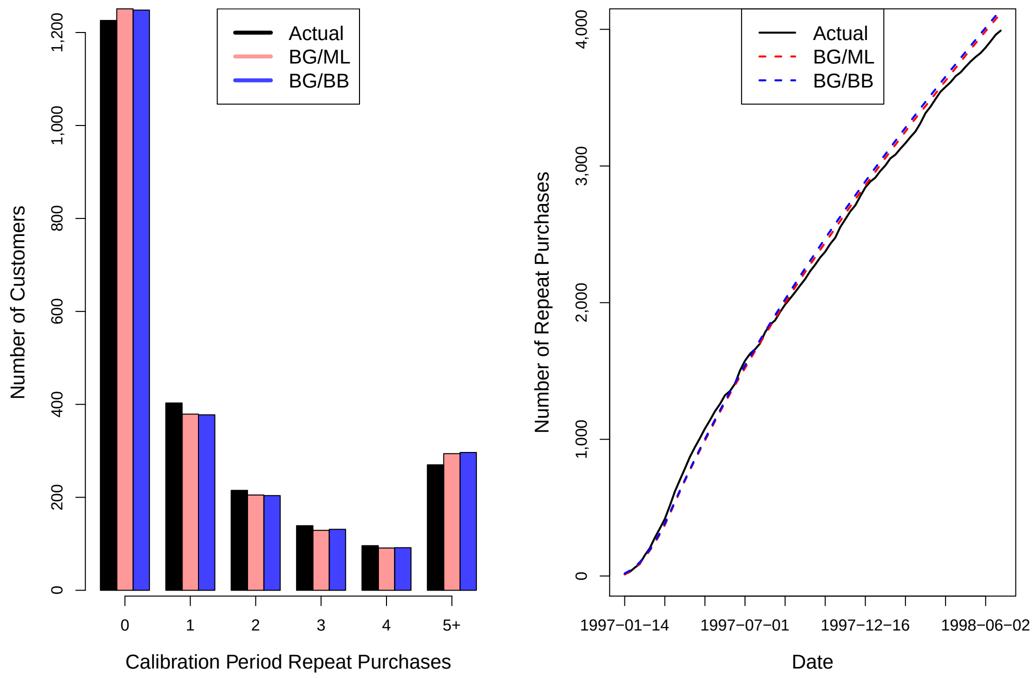
The estimated hyperparameters of the weekly BG/BB model applied to the CDNOW dataset are $\alpha = .58, \beta = 13.13, \gamma = .39$, and $\delta = 9.72$. We match the mean and variance of the heterogeneity distributions with respect to the individual-level purchase and death hyperparameters, q and θ respectively. Because the heterogeneity distributions with respect to θ are the same for the BG/BB model and the BG/ML model, the estimated death hyperparameters for the BG/ML are the same as the estimated death hyperparameters for the BG/BB.

For the purchase process, we simulate a large number (1MM) of samples of q from the beta(α, β) distribution for the BG/BB model. We then apply the logit function to these values of q, computing logit$(q^{(k)}) = \log[q^{(k)}/(1-q^{(k)})]$ for each sample k. The empirical mean and variance of these 1MM logit(q) samples are -4.15 and 3.91, respectively. Therefore, the latent method of moments estimator for the BG/ML model hyperparameters is

$$\mu_p = -4.15, \quad \sigma_p^2 = 3.92, \quad \gamma = .388, \quad \delta = 9.724 \tag{16}$$

The corresponding model fits and forecasts for the BG/ML are virtually identical to those of the BG/BB, as shown in Figure 10.

Figure 10: Predicted vs. Actual Frequency of Repeat Transactions



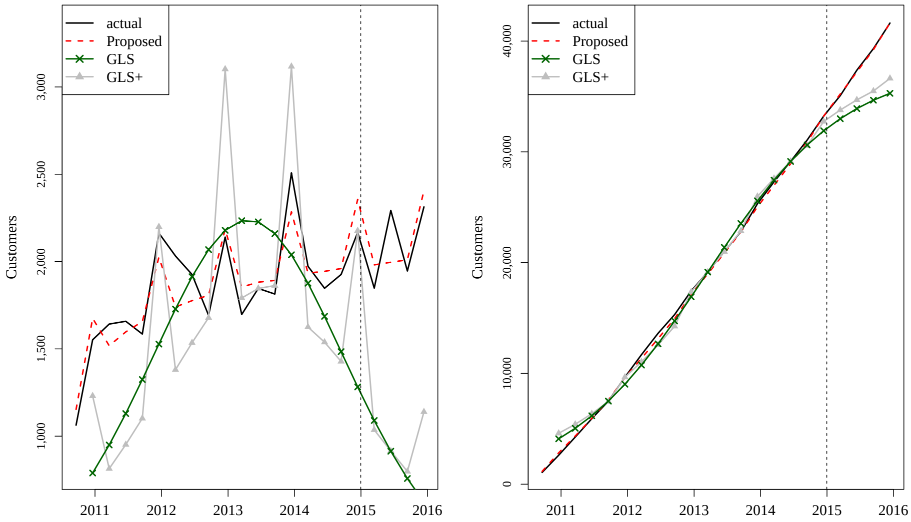
We assume the following two parameter sets for the acquisition process:

$$r_A = .5, \quad \alpha_A \in (5,523.9,\ 15,624.1), \quad c_A = 1.5, \quad \beta_A = \log(2) \tag{17}$$

β_A is set such that the baseline hazard function is doubled during the fourth calendar quarter. These scenarios correspond to a median baseline time until acquisition of 12.5 and 25 years, respectively. By the end of the calibration period, 32.4% and 16.2% of the total prospect pool will have been acquired, respectively.

The 64 data sets are created by varying the parameters of the acquisition, and repeat purchase models along six dimensions relative to the baseline parameters in Equations 16 and 17. We consider high purchase rate and purchase homogeneity scenarios by doubling μ_p and σ_p^2. We also consider high death rate and death homogeneity scenarios by doubling the mean and polarization indices of the beta(γ, δ) distribution, $\gamma/(\gamma+\delta)$ and $1/(1+\gamma+\delta)$, respectively (Sabavala and Morrison (1977)). Finally, we consider a high seasonality scenario in which the baseline purchase rate is doubled during the fourth calendar quarter (i.e., $\beta_p = |\mu_p|$).

Figure 11: Incremental (left) and Cumulative (right) Quarterly Customers Acquired



D GLS MODEL VALUATION AND COMPARISON

In this section, we elaborate upon the model forecasts and valuations provided in the "Empirical Analysis" section. GLS estimates the parameters of the acquisition process using the technological substitution model (TSM). However, seasonality is evident in the fourth calendar quarter. Therefore, we create a variant of the TSM which incorporates covariates using proportional hazards. Using GLS notation (Equation 9 in GLS), denoting the time t vector of seasonal covariates by $\mathbf{x}_A(t)$, and the corresponding vector of parameter estimates by $\boldsymbol{\delta}_A$, the extended expression for the cumulative number of customers N_t at any time t is given by

$$N_t = \alpha \times \left\langle 1 - \exp\left\{ \sum_{i=1}^{t} -\gamma + \log\left[\frac{1+\exp(-\beta-\gamma(i-1))}{1+\exp(-\beta-\gamma i)}\right] \right\} \right\rangle \exp[\boldsymbol{\delta}_A^T \mathbf{x}_A(i)]$$

Parameters are estimated via non-linear least squares, fitting the acquisition parameters to N_t, the cumulative customer additions data. We provide the resulting incremental and cumulative quarterly estimates of customer acquisitions for our proposed method, GLS, and GLS+ in Figure 11, alongside the observed data.

The estimated acquisition parameters for GLS and GLS+ are shown below, using notation from GLS and denoting by β_{Q_4} the quarterly seasonal dummy:

	α	β	γ	β_{Q_4}
GLS	37.823	-2.345	.237	.000
GLS+	39.645	-2.377	.208	.507

The annual repeat rate for the e-commerce retailer is 25.9% as of the end of the calibration period. Therefore, the retention rate for GLS is 25.9%. It is unclear how this process can be easily modified. If customer churn were observed, we could allow the retention rate to vary over time as a function of covariates through a logit formulation. Because churn is not observed, we assume a 25.9% retention rate for GLS+ but counteract this assumption through the repeat purchase process next.

While GLS does not explicitly model future purchases, it does model the dollar margin per active customer, which is assumed to be equal to the trailing four quarter average. Dollar margin per active customer can be decomposed into margin per purchase and the number of purchases per active customer. Assuming that these latter two quantities are constant because their product is constant, future purchases per active customer-quarter under GLS are equal to .495.

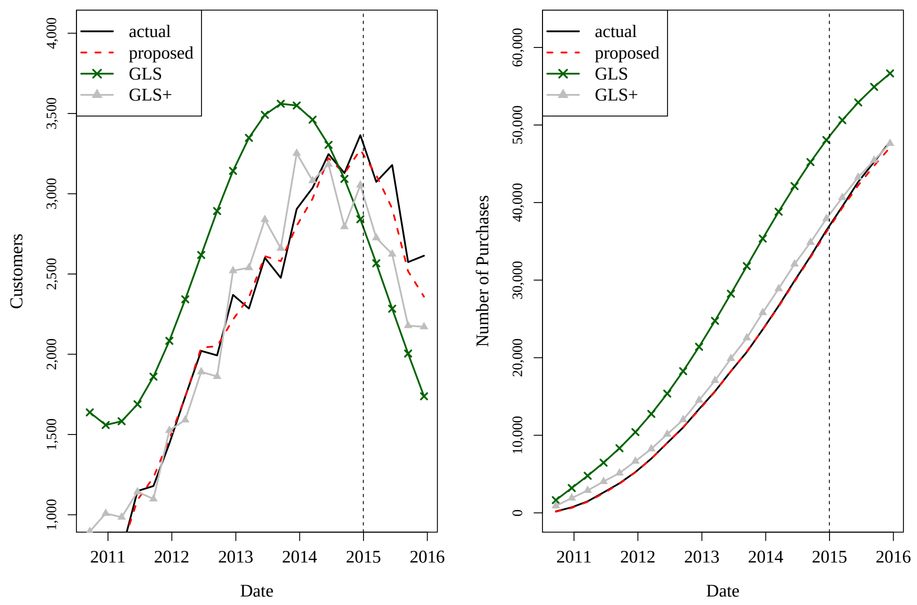
The historical evolution of inferred repeat purchases per active customer is upward-sloping over time, with a seasonal increase in the second calendar quarter. Therefore, we create a variant of the purchase process which allows for both a time trend and seasonality. We estimate the following equation for GLS+ purchases per active customer, denoting by q the number of quarters after the beginning of commercial operations:

$$\widehat{\frac{\text{Purchases}}{\text{Active Customer}}} = .254 + .0116 \times q + .0372 \times \mathbb{1}[Q(q) = 2],$$

where $\mathbb{1}[Q(q) = 2]$ is an indicator variable equal to one if quarter q is within the second calendar quarter and zero otherwise. The incremental and cumulative quarterly estimates of purchases for our proposed method, GLS, and GLS+ are shown in Figure 12, alongside the observed data.

While allowing purchases per active customer to trend over time improves the GLS model's fit and forecast dramatically, it continues to underestimate future purchases.

Figure 12: Incremental (left) and Cumulative (right) Quarterly Repeat Purchases



Using the assumptions made above, dollar sales per purchase is equal to the trailing four quarter average of \$221.80. However, sales per purchase is rising linearly during the calibration period. Therefore, we extend GLS to allow sales per purchase to have a time trend which is estimated with a regression. We estimate the following equation for GLS+ spend per purchase:

$$\widehat{\frac{\text{Spend}}{\text{Purchase}}} = 145.58 + 4.621 \times q.$$

Using the fitted parameter estimates for the GLS and GLS+ customer acquisition, retention, purchases per active customer and spend per purchase, we make future revenue projections. We then convert these revenue projections into overall valuation estimates using the same assumptions that we had made to convert the revenue forecasts of our proposed model into an overall valuation estimate.