Abstract

Demand estimation is important for new-product strategies, but is challenging in the absence of actual sales data. We develop a cost-effective method to estimate the demand of new products based on incentive-aligned choice experiments. Our premise is that there exists a structural relationship between manifested demand and the probability of consumer choice being realized. We illustrate the mechanism using a theory model, in which consumers learn their product valuation through costly effort and their effort incentive depends on the realization probability. We run a large-scale field experiment on a mobile game platform, where we randomize the price and realization probability when selling a new product. The data support our theoretical prediction and the decision effort mechanism. We then estimate a structural model of consumer choice. The structural estimates allow us to infer actual demand with reasonable accuracy using inexpensive choice experiments with small to moderate realization probabilities.

Key words: demand estimation, new product, market research, choice experiment, incentive alignment, external validity, structural modeling.
1 Introduction

Accurate demand estimation is important for new products to succeed, but is challenging in the absence of historical sales data (e.g., Braden and Oren 1994, Urban et al. 1996, Hitsch 2006, Desai et al. 2007, Bonatti 2011). For decades, researchers have spent considerable effort developing market research strategies to estimate product demand before actual launch. Solutions to date can be classified into three categories. Hypothetical approaches ask participants to either state their product valuation or make hypothetical product choices which are then used to infer their product valuation (e.g., Miller et al. 2011). Incentive-aligned approaches further engage respondents by requiring them to actually purchase the product at the price they are willing to pay with a “realization probability” (e.g., Becker et al. 1964, Ding 2007). Test marketing, which can be seen as fully incentive-aligned choice experiments, sells the product in trial markets to gather consumer choice data in real purchase environments.

These solutions are imperfect. Hypothetical approaches are known to generate hypothetical biases (e.g., Frykblom 2000, Wertenbroch and Skiera 2002). Incentive alignment can improve the accuracy of demand estimation compared with the hypothetical approach (e.g., Ding 2007, Miller et al. 2011), but may not recover demand in real purchase settings accurately (e.g., Miller et al. 2011, Kaas and Ruprecht 2006). Test marketing achieves the highest external validity among the three methods (Silk and Urban 1978). However, the gain in external validity comes at a cost. Other things being equal, the higher the realization probability, the more actual products the company must provide for market research. Besides higher operational costs, more products means greater opportunity costs of selling at suboptimal prices – by definition, the company would not know the optimal price before it is able to estimate demand. As a result, existing market research methods often have to trade off external validity against cost control.

In this paper, we try to resolve this cost-validity conundrum by developing a theory-based, cost-effective method to estimate the demand of new products. This method is low-cost because

---

1The company we collaborate with confirmed that it had refrained from test marketing for the same reason.
it relies only on moderate to small realization probabilities. It is effective because it is able to approximate the demand estimation results of test marketing. Figure 1 presents the intended contribution of this paper.

![Figure 1: Intended Contribution of the Paper](image)

The idea is as follows. We posit that consumers must make a costly effort to learn their true product valuation. For example, consumers may spend time inspecting product features, searching through alternative options, or thinking about possible usage scenarios (e.g., Shugan 1980, Wathieu and Bertini 2007, Guo and Zhang 2012, Guo 2016, Kleinberg et al. 2017). Whether consumers are willing to make this costly effort depends on the realization probability. Intuitively, if a consumer knows that her product choice is unlikely to be realized, she will have little incentive to uncover her true product valuation and will make her choice based on her prior belief. On the contrary, if a consumer knows that her product choice is for real, she will want to think about how much she truly values the product and make her choice prudently. As a result, there exists a structural relationship between realization probability and manifested demand. Our proposed demand estimation method thus proceeds in two steps: first, estimate this structural relationship using product choice data under smaller realization probabilities; second, use the estimation results to forecast product demand in actual purchase settings.

We formalize the above mechanism with a theory model, in which consumers decide whether they are willing to purchase a product at a given price and a given realization probability.²

²This choice experiment can be seen as a form of incentive-aligned choice-based conjoint analysis with price
The model predicts that manifested price sensitivity increases with realization probability. To understand the intuition, imagine that the company had offered the product for free. Agreeing to buy the product had been a no-brainer. Now, suppose the company raises the price gradually. As the price approaches a consumer’s prior valuation for the product, she will have a greater incentive to zoom in and think carefully about her true need for the product, and the only change this thinking brings to her decision is to not buy the product. A higher realization probability increases the gravity of the purchase decision and amplifies this negative effect of price on demand. Therefore, it will appear as if consumers are more price-sensitive under higher realization probabilities.

To test the theory model and to evaluate the proposed demand estimation method, we run a large-scale field experiment. We choose the field, as opposed to the lab, in order to minimize factors that may affect external validity other than the realization probability. We collaborate with a mobile soccer game platform. The new product is a new game package that may enhance user performance. We set four realization probabilities: 0, 1/30, 1/2, and 1, where the 0-probability group is designed to capture the effect of hypothetic approaches and the 1-probability group is designed to mirror test marketing. We randomly assign prices and realization probabilities across users exposed to the experiment.

The experiment results support the theory prediction – consumers are more price-sensitive under higher realization probabilities. We rule out a number of competing explanations of this effect using data from a post-choice survey. Moreover, we obtain process measures of consumers’ decision effort. We find that decision effort increases with realization probability, consistent with the behavioral mechanism underlying the theory prediction.

Having validated the theory foundation of the proposed demand estimation method, we develop a structural model of consumer effort choice and purchase decision based on the mechanism developed in the theory. This forms the core of our proposed demand estimation method. More specifically, we estimate the structural model using data from the subsample of smaller being the only product attribute. Marketing practitioners call the hypothetical version of this type of experiment a “Monadic pricing survey.”
realization probabilities (1/30 and 1/2 in the field experiment). To assess the external validity of the proposed method, we use the estimation results to forecast product demand in real purchase settings and compare the forecast against the holdout sample where realization probability equals 1. The structural forecast performs remarkably well. For example, the forecast error in price sensitivity is only 4.49% compared against the holdout sample. Simple extrapolation of data from smaller realization probabilities to actual purchase settings, in contrast, yields forecast errors of around 20%. This suggests that the external validity of the proposed method relies on a detailed, structural understanding of the behavioral process.

The rest of the paper proceeds as follows. We continue in Section 2 with a review of the related literatures. In Section 3, we develop a theory model to illustrate the mechanism and to formulate testable predictions. We then present the field experiment in Section 4 and discuss model-free evidence of the theory in Section 5. In Section 6, we draw on the theory to develop and evaluate a method to estimate new product demand based on structural use of choice data from smaller realization probabilities. We conclude in Section 7 with discussions of future research.

2 Literature Review

Researchers have long been exploring ways to estimate product demand, or equivalently, consumers’ product valuation. The most reliable way to estimate demand is to use actual sales data or test market data (Silk and Urban 1978). These types of data have high external validity because they are observed in real purchase settings. However, actual sales data is not available for new products prior to launch, whereas test market data is costly to obtain. Even in the 1970s, the cost of test marketing could surpass one million US dollars. Furthermore, test marketing can be risky for a firm as it allows competitors to obtain the firm’s product information and respond strategically.

As a result, researchers have developed pre-test-market methods, usually called laboratory or simulated test markets, in which recruited consumers are given the opportunity to buy in a
simulated retail store (Silk and Urban 1978). Pre-test-market methods also have high external validity, because they provide a realistic purchase environment and consumers’ choices are realized for certain (Silk and Urban 1978, Urban and Katz 1983, Urban 1993). However, pre-test-market methods are still costly – the company incurs not only the logistical costs of actual selling, but also the opportunity cost of selling at potentially suboptimal prices. It can even be infeasible as the company may not have enough product samples to sell at the prelaunch stage.

A different approach to estimating product demand is to use hypothetical surveys or hypothetical choice experiments. Marketing researchers have developed hypothetical choice-based conjoint analysis to measure consumers’ tradeoffs among multi-attribute products (see Hauser and Rao 2004, Rao 2014 for an overview), and choice-based conjoint analysis can be augmented to estimate product valuation (e.g., Kohli and Mahajan 1991, Jedidi and Zhang 2002). Economists have used “contingent valuation methods” to estimate people’s willingness-to-pay for public goods (Mitchell and Carson 1989), where participants are asked to either state their valuation directly (open-ended contingent evaluation) or to choose whether they are willing to purchase a good at a given price (dichotomous choice experiments).

These hypothetic methods ask participants to answer questions or make choices without actual consequences. As a result, these methods are riskless, low-cost, and widely applicable to concept testing. However, researchers have often found hypothetical methods unreliable. Both hypothetical open-ended contingent valuation and hypothetical choice experiments are found to over-estimate product valuation compared to actual purchases (Diamond and Hausman 1994, Cummings et al. 1995, Balistreri et al. 2001, Lusk and Schroeder 2004, Miller et al. 2011). This happens due to participants’ lack of incentive to expend cognitive efforts needed to provide an accurate answer, ignorance of their budget constraints, or tendency to give socially desirable answers in hypothetical settings (Camerer et al. 1999, Ding 2007).

A stream of literature tries to derive more reliable demand estimates using data from hypothetic methods, but the results are mixed. One solution is to use “calibration techniques” but the calibration factors vary significantly and are specific to the product and the context (Blackburn et al. 1994, Fox et al. 1998, List and Shogren 1998, Murphy et al. 2005). Cummings and
Taylor (1999) propose a “cheap-talk” design of questionnaire to reduce the hypothetical bias. List (2001) applies this design to a well-functioning marketplace that auctions off sports cards. He finds that the cheap-talk design mitigates the hypothetical bias, but only for inexperienced bidders.

Another stream of research tries to overcome the hypothetic bias by making participants responsible for the consequences of their choices with a probability, called the “realization probability.” Becker et al. (1964) design such a mechanism (hereafter BDM), where a participant is obliged to purchase a product if the price drawn from a lottery is less than or equal to her stated product valuation. The BDM mechanism has been widely used to elicit willingness-to-pay in behavioral decision experiments (e.g., Kahneman et al. 1990, Prelec and Simester 2001, Wang et al. 2007). Wertenbroch and Skiera (2002) compare BDM with hypothetical contingent valuation methods, and find that BDM yields lower willingness-to-pay. Extending the BDM approach, Ding et al. (2005) and Ding (2007) design an incentive-aligned mechanism for conjoint analysis by replacing stated product valuation with inferred product valuation from conjoint responses. Again, participants must adopt the product they chose with a realization probability. The authors show that incentive-aligned choice-based conjoint analysis outperforms its hypothetical counterpart in out-of-sample predictions of actual purchase behavior. Based on this idea, researchers have developed more-advanced incentive-aligned preference measurement methods (e.g., Park et al. 2008, Ding et al. 2009, Dong et al. 2010, Toubia et al. 2012), and confirm that incentive alignment leads to substantial improvement in predictive performance when compared to hypothetical methods.

However, whether incentive alignment can fully recover actual choices remains a question. Notably, Yang et al. (2015a) find that consumers become more price sensitive as realization probability increases in incentive-aligned choice experiments. Our paper reinforces this finding and shows that, indeed, incentive-aligned choice experiments may misrepresent demand in actual purchase settings, although they forecast demand more accurately compared with hypothetical approaches. We propose and empirically validate a theory of decision effort that can explain the bias in incentive-aligned choice experiments. Based on the theory, we develop a
method to correct the bias in incentive-aligned experiments, which allows us to estimate the real demand curve in a cost-effective way.

Our decision effort mechanism emphasizes the idea that consumers need to incur a cost to learn their product valuation. Consumers are often uncertain about product performance and individual preferences (e.g., Urbany et al. 1989, Kahn and Meyer 1991, Ariely et al. 2003, Ofek et al. 2007, Wang et al. 2007). It is costly to evaluate product features (e.g., Wernerfelt 1994, Villas-Boas 2009, Kuksov and Villas-Boas 2010) or to think through one’s subjective preferences (e.g., Shugan 1980, Wathieu and Bertini 2007, Guo and Zhang 2012, Huang and Bronnenberg 2015, Guo 2016). In a recent paper, Kleinberg et al. (2017) show that the notion of costly valuation learning has important implications for mechanism design; in particular, it renders the popular increasing-price auction ineffective. As a result of costly preference learning, consumers face an effort-accuracy tradeoff when making choices instead of maximizing decision accuracy as conventionally assumed (Hauser et al. 1993, Payne et al. 1993, Yang et al. 2015b). Wilcox (1993) shows that increased incentives raise subjects’ willingness to incur decision effort and hence influence decision outcomes. Smith and Walker (1993) survey 31 experimental studies and find that higher rewards shift the experiment results towards the prediction of rational models. They also explain this result with effort theory – that is, higher rewards induce agents to exert more cognitive effort. Yang et al. (2015a) find that consumers’ amount of attention to the choice problem increases with the scale of incentives as measured by the realization probability. There is also evidence from neuroeconomics that, when choosing consumer goods, brain activation is stronger and more widespread in the real choice condition than in the hypothetical condition (Camerer and Mobbs 2017).

In this paper, we further investigate the role of costly decision effort on consumer response in choice experiments, where consumers’ effort incentive depends on the probability of their decisions being realized. This allows us to portray the structural relationship between realization probability and product demand. In the following session, we develop a theory model to describe this mechanism and to form testable predictions.
3 Theory Model

Consider a market with a unit mass of consumers. The true valuation of a new product, \( v_i \), is heterogeneous across consumers, following a distribution \( f(\cdot) \) unknown to the firm and consumers (otherwise there is no need for demand estimation). Consider a representative consumer \( i \). She does not know her true product valuation \( v_i \) \textit{ex ante}. The mean of her prior belief about the true valuation is \( \mu_{0i} \), which can be decomposed as \( \mu_{0i} = v_i + e_i \), where the perception error \( e_i \) follows a distribution \( g(\cdot) \). We assume that \( g(\cdot) \) is continuous and symmetric around 0, is the same across consumers, and that consumers know \( g(\cdot) \) \textit{ex ante}.

The consumer can expend a decision effort to learn about her true valuation of the product. If the consumer devotes effort \( t \), she will know the true value of \( v_i \) with probability \( t \), and her belief about \( v_i \) stays at \( \mu_{0i} \) with probability \( 1 - t \). An example of a choice context this formulation captures is a consumer’s search of whether she already has a product in her possession that is a good substitute for the new product. Alternatively, we can model the decision effort as smoothly reducing a consumer’s uncertainty about her true product valuation, but the qualitative insight of the theory model remains the same. We write the cost of effort as \( \frac{1}{2} c t^2 \) to capture the idea of increasing marginal cost. We assume that consumer \( i \)'s utility from purchase is \( U_i = v_i - p \) and the consumer has a reservation utility of zero. Thus consumer \( i \) will purchase the product priced at \( p \) if and only if \( \mathbb{E}[v_i] \geq p \), where \( \mathbb{E}[v_i] \) denotes the consumer’s expected value of \( v_i \).

The timing of the choice experiment unfolds as follows. In the first stage, the consumer observes the price \( p \) and the realization probability \( r \). She is told that if she chooses “willing to buy,” she will have to pay \( p \) and receive the product with probability \( r \), and will pay nothing and not receive the product with probability \( 1 - r \). In the second stage, the consumer chooses the level of her decision effort, \( t \). In the third stage, the consumer decides whether to choose “willing to buy” based on the outcome of her decision effort. If she is willing to buy, a lottery will be drawn and with probability \( r \) she will pay price \( p \) and receive the product as promised in stage one.

We first derive the optimal effort of the representative consumer. The consumer chooses
effort $t$ to maximize her expected net utility:

$$E[U(t, \mu_0; p, r)] = r \left( tE[(v_i - p)^+] + (1 - t)(\mu_0 - p)^+ \right) - \frac{1}{2} ct^2, \quad (1)$$

where the expectation is taken over consumer $i$’s prior perceived distribution of $v_i$ before she expends any decision effort.

The first-order condition of $\frac{\partial E[U(t, \mu_0; p, r)]}{\partial t} = 0$ yields the optimal effort level:

$$t^*(\mu_0; p, r) = \frac{r}{c} \left( E[(v_i - p)^+] - (\mu_0 - p)^+ \right) \quad (2)$$

The second-order condition is trivially satisfied for this optimization problem. We obtain the following comparative statics results.

**Proposition 1** Suppose $p - \mu_0$ is strictly within the support of $g(\cdot)$. The consumer’s optimal decision effort increases with realization probability $r$, and decreases with the distance between price and her prior belief of her valuation $|p - \mu_0|$. A greater realization probability amplifies the latter effect.

Proof: see the Appendix.

Intuitively, expending effort helps a consumer make a better informed purchase decision based on her true product valuation. The higher the realization probability, the higher the value of this effort. When realization probability equals 1, the consumer makes the same effort as in real purchase decisions. When realization probability equals 0, choices become purely hypothetical with no impact on consumer utility, and the consumer makes no effort to learn her product valuation. Moreover, when product price is extremely low (or high), the consumer may trivially decide to buy (or not buy) regardless of her true valuation, which makes the decision effort unnecessary. On the other hand, when price is closer to a consumer’s prior valuation, making a purchase decision based on the prior belief alone is more likely to lead to a mistake, and the consumer will want to expend more effort to discover her true valuation.

\[\text{The consumer may choose randomly or be pro-social towards the experimentalist and choose truthfully based on her prior belief.}\]
Knowing consumers’ optimal effort decisions, we can derive the “manifested demand” for the product, i.e., the expected fraction of consumers who choose “willing to buy” given price $p$ and realization probability $r$:

$$D(p, r) = \int_{v_i} \int_{e_i} \left[ t^*(v_i + e_i; p, r)\mathbf{1}(v_i \geq p) + (1 - t^*(v_i + e_i; p, r))\mathbf{1}(v_i + e_i \geq p) \right] g(e_i) f(v_i) de_idv_i.$$  \hspace{1cm} (3)

We emphasize the notion of manifested demand, as opposed to estimated demand, to highlight the theoretical effect of realization probability on consumer choices. In other words, even if consumers are behaving truthfully based on their expected product valuation and even if there is no empirical error, manifested demand may still differ from actual demand because consumer choices are not fully realized.

Now we investigate how realization probability affects the manifested demand curve. Let $\frac{\partial D(p, r)}{\partial p}$ denote the local slope of the demand curve at price $p$, measuring consumers’ price-sensitivity at price $p$. To facilitate presentation, we define the slope of the demand curve at the center of the true valuation distribution, $\frac{\partial D(p, r)}{\partial p} |_{p=\mu_v}$, as the “central” slope of the demand curve. The following proposition summarizes our finding.

**Proposition 2** Suppose $g(\cdot)$ is symmetric around 0 and is weakly decreasing on $(0, \infty)$. Suppose $f(\cdot)$ has a unique mode $\mu_v$, and is weakly increasing on $(-\infty, \mu_v)$ and weakly decreasing on $(\mu_v, \infty)$. Then the following results hold.

$D(p, r)$ is weakly decreasing in the realization probability $r$ when $p > \mu_v$, and is weakly increasing in $r$ when $p < \mu_v$. Denote $Z^- (p) = \{ z > 0 : f(p + z) - f(p - z) < 0 \}$, $Z^+ (p) = \{ z > 0 : f(p + z) - f(p - z) > 0 \}$, and $S_g = \{ z > 0 : g(z) > 0 \}$. For $p > \mu_v$, if the (Lebesgue) measure of $Z^- (p) \cap S_g$ is greater than 0, then $D(p, r)$ strictly decreases in $r$. For $p < \mu_v$, if the (Lebesgue) measure of $Z^+ (p) \cap S_g$ is greater than 0, then $D(p, r)$ strictly increases in $r$.

The central slope of the demand curve, defined as $\frac{\partial D(p, r)}{\partial p} |_{p=\mu_v}$, is weakly decreasing in $r$. If $Z(\mu_v) = \{ z > 0 : f(\mu_v + z) + f(\mu_v - z) < 2f(\mu_v) \}$ has a non-empty intersection with the set of $z$ where $g(z)$ is strictly decreasing, then when $r$ increases, the central slope of the demand curve strictly decreases, i.e., the demand curve becomes steeper.
Proof: see the Appendix.

It should be noted that many commonly seen distribution functions satisfy the conditions for the results in the above proposition to hold strictly. We illustrate this fact using normal and uniform distributions, respectively. In the first example, the true valuation \( v_i \) follows a normal distribution \( N(\mu_v, \sigma_v^2) \) and the perception error follows a normal distribution \( N(0, \sigma_0^2) \). Hence the perception error distribution \( g(\cdot) \) is always positive, so that \( S_g = (0, \infty) \). When \( p > \mu_v \), the true valuation distribution \( f(\cdot) \) is strictly decreasing, which implies \( Z^-(p) = (0, \infty) \). When \( p < \mu_v \), \( f(\cdot) \) is strictly increasing, so that \( Z^+(p) = (0, \infty) \). Therefore, the manifested demand \( D(p, r) \) strictly decreases with \( r \) for any \( p > \mu_v \) and strictly increases with \( r \) for any \( p < \mu_v \). We also have \( Z(\mu_v) = (0, \infty) \), which is the same as the set of \( z \) where \( g(z) \) is strictly decreasing. Therefore, the central slope of the demand curve \( \frac{\partial D(p, r)}{\partial p} \bigg|_{p=\mu_v} \) strictly decreases with \( r \).

In the second example, the true valuation \( v_i \) is uniformly distributed on \([\mu_v - \sigma_v, \mu_v + \sigma_v]\) and the perception error is uniformly distributed on \([-\sigma_0, \sigma_0]\). It follows that \( S_g = (0, \sigma_0) \). When \( p > \mu_v \), \( Z^-(p) = (|p - (\mu_v + \sigma_v)|, \|p - (\mu_v - \sigma_v)|) \), so manifested demand \( D(p, r) \) strictly decreases with \( r \) if and only if \( \sigma_0 > |p - (\mu_v + \sigma_v)| \), that is, if and only if \( \mu_v + (\sigma_v - \sigma_0) < p < \mu_v + (\sigma_v + \sigma_0) \). When \( p < \mu_v \), \( Z^+(p) = (|p - \sigma_v|, \|p - (\mu_v + \sigma_v) - p) \), manifested demand \( D(p, r) \) strictly decreases with \( r \) if and only if \( \sigma_0 > |(\mu_v - \sigma_v) - p| \), that is, if and only if \( \mu_v - (\sigma_v + \sigma_0) < p < \mu_v - (\sigma_v - \sigma_0) \).

Figure 2 presents how the demand curve changes with realization probability \( r \) assuming \( f(\cdot) \) and \( g(\cdot) \) are both uniform distributions.\(^4\) We can see that the demand curve rotates at \( p = \mu_v \) as realization probability changes. Specifically, demand increases with realization probability for any price below \( \mu_v \) and decreases with realization probability for any price above \( \mu_v \). As realization probability increases, the demand curve becomes steeper, and consumers appear to be more price sensitive.

\(^4\)The figure is plotted under the condition that \( \sigma_0 > \sigma_v \).
4 Field Experiment

We run a field experiment to validate the prediction and the mechanism of the theory and to evaluate the proposed demand estimation method. We choose the field experiment approach to minimize threats to external validity such as the decision context. This allows us to identify the effect of realization probability on the external validity of demand estimation methods.

We collaborate with a top mobile platform of soccer games in China. Founded in 2013, the platform currently hosts 80,000 daily active users, generating 2 million US dollars in monthly revenue. In the game, each user manages a soccer team and the goal is to win as many times as possible. A team’s likelihood of winning depends on the number of high-quality players it enlists. The new product we sell in the field experiment is a “lucky player package” that consists of six high-quality players. This player package had never been sold on the game platform prior to the experiment.

We want to randomize realization probability and price. We set four different realization probabilities: 0, 1/30, 1/2, and 1. The 0-probability group is designed to replicate hypothetical surveys, and the 1-probability group is meant to mirror the actual purchase setting. We add two interim realization probability groups because the proposed demand estimation method needs at least two realization probability levels for empirical identification and we choose only
two for a conservative evaluation of the method’s predictive power. We assign a 1/2-probability group to observe the effect of moderate realization probabilities. In addition, we create a 1/30-probability group because, in many experiments, the rule-of-thumb is to recruit 30 subjects per condition. For future applications of the proposed demand estimation method using 30 subjects per condition, this realization probability can be more tangibly interpreted as one out of the 30 subjects buying the product for real, which makes the experiment looks more trustworthy than using a smaller realization probability.

We set five price levels, measured as 1600, 2000, 2400, 2800, or 3200 “diamonds,” which is the currency used in the game. Users need to pay real money to obtain diamonds. The exchange rate is about 1 US dollar for 100 diamonds. We discuss with the company to make sure this price range is reasonable and at the same time the gap between prices is large enough to elicit different purchase rates. The five price levels, combined with the four realization probabilities, lead to 20 conditions for the experiment. Once a user enters the experiment, she is randomly assigned to one of the conditions.

Each condition presents the user with a screen of the choice task. (Figure A1 in the Appendix presents the screen for the 1/30-probability group.) On this screen, we inform the user that she has a chance to purchase a lucky player package at price $p$ and ask her to choose between “willing to purchase” and “not willing to purchase.” For the 0-probability group, we inform the user that this is a hypothetic survey and no actual transaction will take place. For the 1-probability group, we notify the user that she has the chance to actually purchase the package. For the interim probability groups, we explain that if the user chooses “willing to purchase,” a lottery will be drawn and there is probability $r$ that she will actually receive the player package and will be charged the price $p$ automatically. If the user chooses “not willing to purchase” or does not win the lottery, she will not receive the player package and will not be charged anything. Users can click on the player package and see the set of players contained therein (see Figure A2 of the Appendix). They can also click on each player and see what skills the player has. After making the purchase decision, the user will be directed to a follow-up survey, which we designed to obtain auxiliary data to test the theory.
Table 1: Number of Users by Realization Probability and by Price

<table>
<thead>
<tr>
<th>Probability</th>
<th>Entered the Experiment</th>
<th>Completed the Choice Task</th>
<th>Completed the Survey</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1355</td>
<td>1095</td>
<td>882</td>
</tr>
<tr>
<td>1/30</td>
<td>1355</td>
<td>920</td>
<td>708</td>
</tr>
<tr>
<td>1/2</td>
<td>1355</td>
<td>922</td>
<td>723</td>
</tr>
<tr>
<td>1</td>
<td>1355</td>
<td>895</td>
<td>671</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Price</th>
<th>Entered the Experiment</th>
<th>Completed the Choice Task</th>
<th>Completed the Survey</th>
</tr>
</thead>
<tbody>
<tr>
<td>1600</td>
<td>1084</td>
<td>774</td>
<td>599</td>
</tr>
<tr>
<td>2000</td>
<td>1084</td>
<td>757</td>
<td>575</td>
</tr>
<tr>
<td>2400</td>
<td>1084</td>
<td>764</td>
<td>589</td>
</tr>
<tr>
<td>2800</td>
<td>1084</td>
<td>761</td>
<td>603</td>
</tr>
<tr>
<td>3200</td>
<td>1084</td>
<td>776</td>
<td>618</td>
</tr>
</tbody>
</table>

The experiment took place from 12AM, December 2, 2016 to 12PM, December 4, 2016. We randomly selected half of the platform’s Android servers, and all users on these servers automatically entered the experiment once they accessed the game during the period of the experiment. A total of 5,420 users entered the experiment, 271 in each condition. Among these users, 3,832 (70.7%) completed the choice task. Among those who completed the choice task, 2,984 (77.87%) filled out the survey. Table 1 reports the number of users assigned to each probability and price group, and the number that completed the choice task or the survey. We notice higher completion rates in the 0-probability group. However, reassuringly, for all groups with positive realization probabilities, neither completing the choice task nor completing the survey is significantly correlated with the assigned realization probability ($Corr = -0.0163, p = 0.2993$ and $Corr = -0.0195, p = 0.3080$, respectively) or the assigned price level ($Corr = 0.0023, p = 0.8660$ and $Corr = 0.0265, p = 0.1013$, respectively).

For each user who completed the choice task, we collect data on her characteristics at the time of the experiment, including the number of diamonds the user has (Diamond) and the VIP level of the user (VIP). The VIP level is determined by how much money the user has spent in the game. Table 2 presents the summary statistics. We can see that Diamond is a highly right-skewed variable, hence we convert it into a new variable $\text{Log-Diamond} = \log(\text{Diamond}+1)$ and will use this new measure in subsequent analysis.
Table 2: Summary Statistics of User Characteristics

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Std Dev</th>
<th>Median</th>
<th>Min</th>
<th>Max</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diamond</td>
<td>3134.44</td>
<td>5498.09</td>
<td>1614.00</td>
<td>0</td>
<td>150969</td>
<td>3832</td>
</tr>
<tr>
<td>Log-Diamond</td>
<td>7.09</td>
<td>1.64</td>
<td>7.39</td>
<td>0</td>
<td>12</td>
<td>3832</td>
</tr>
<tr>
<td>VIP</td>
<td>3.00</td>
<td>3.10</td>
<td>2.00</td>
<td>0</td>
<td>15</td>
<td>3832</td>
</tr>
</tbody>
</table>

Notes. The sample consists of all users who completed the choice task.

5 Reduced-Form Evidence

In this section, we present reduced-form evidence of the theory prediction and of the decision effort mechanism, using data from the field experiment.

We first examine aggregate-level demand. By demand, we mean the proportion of users who chose “willing to purchase” out of those who completed the choice task in each condition. Figure 3 shows how aggregate-level demand changes with price under each realization probability. We see a pattern – as the realization probability increases, demand seems to decrease faster with price; in addition, the overall level of demand seems to decrease with realization probability.

To verify these observations, we fit a linear demand curve for each realization probability group by regressing individual-level purchase decisions on price. The dependent variable Purchase equals 1 if the user chooses “willing to purchase” and 0 if the user chooses “not willing to purchase.” For the ease of presentation, we normalize the five price levels to 4, 5, 6, 7, 8 respectively in this regression and subsequent analysis. Table 3 presents the estimated price coefficient and intercept of the demand curves. The slope of demand curve (absolute value of the price coefficient) increases with the realization probability, consistent with the prediction of our theory.

We further examine how individual-level purchase decisions are jointly influenced by price and realization probability. Column (1) of Table 4 shows that the likelihood of purchase decreases with price, as expected. It also decreases with realization probability, an effect we will comment on later. In column (2), we add the interaction term of price and realization probability, and this interaction term has a significantly negative coefficient. This means that users are more price sensitive with higher realization probabilities, consistent with the theory.
Figure 3: Realization Probability and Demand Curve – Aggregate-Level Evidence

Prediction. In column (3), we further control for user characteristics, namely, Log-Diamond and VIP. Having more diamonds and having lower VIP status are associated with higher purchase rates. Again, users become more price sensitive as realization probability increases.

A comment on the level of demand is in order. According to our theory model, for any given price, the level of demand decreases with realization probability only when prices are higher than users’ average product valuation. As a further test of the theory, in the post-choice survey we ask users to rate how they perceive the price of the product on a scale from 1 (very low) to 5 (very high). Indeed, the answers confirm that users view the price as relatively high; the mean answer is 3.99, significantly higher than the neutral level of 3 (\(t = 52.83, p < 0.001\)).

So far, data support our theory prediction that the demand curve becomes steeper as the realization probability increases. Next we examine whether the change in the slope of the demand curve is driven by the decision effort mechanism we propose. We need a measure
Measuring decision effort is difficult (Bettman et al. 1990), and we try to do so using two measures. First, our experiment setting allows us to gauge how much a user has learned about the product. More specifically, in the post-choice survey, we ask each user to answer “which of the following soccer players was not included in the player package.” If a user has carefully thought about her valuation of the player package, presumably she should know its content. We let the effort measure equal 1 if the user provides the correct answer (there is only one correct answer), and 0 otherwise. As a second proxy of decision effort, we draw upon the classic measure of decision time (Wilcox 1993). We record decision time as the number of seconds it takes from the point the user first arrives at the choice task page to the point she makes a choice. The decision time variable is highly right-skewed with some extremely large values, hence we take a log transformation of it for subsequent analysis. Table 5 reports the summary statistics of these effort measures.

As a direct mechanism test, we regress the two measures of decision effort on realization probability and price. Table 6 presents the result. For both measures of effort, users’ effort input increases with realization probability, consistent with Proposition 1. Effort also decreases with price, although the effect is insignificant. The negative effect of price on effort echoes the survey result that users perceive the price of the player package as relatively high. As price increases from an already-high level, not to buy becomes a clearer decision regardless of a user’s true product valuation, which makes effort less needed. This result is again consistent
Table 4: Price Sensitivity Increases with Realization Probability

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Purchase</td>
<td>Purchase</td>
<td>Purchase</td>
</tr>
<tr>
<td>Price</td>
<td>-0.0372***</td>
<td>-0.0219***</td>
<td>-0.0218***</td>
</tr>
<tr>
<td></td>
<td>(0.00556)</td>
<td>(0.00753)</td>
<td>(0.00752)</td>
</tr>
<tr>
<td>Probability</td>
<td>-0.226***</td>
<td>0.0304</td>
<td>0.00998</td>
</tr>
<tr>
<td></td>
<td>(0.0191)</td>
<td>(0.0844)</td>
<td>(0.0838)</td>
</tr>
<tr>
<td>Price × Probability</td>
<td>-0.0427***</td>
<td>-0.0391***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0135)</td>
<td>(0.0134)</td>
<td></td>
</tr>
<tr>
<td>Log-Diamond</td>
<td>0.0217***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00499)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VIP</td>
<td>-0.0159***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00255)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.773***</td>
<td>0.681***</td>
<td>0.574***</td>
</tr>
<tr>
<td></td>
<td>(0.0349)</td>
<td>(0.0464)</td>
<td>(0.0587)</td>
</tr>
<tr>
<td>N</td>
<td>3832</td>
<td>3832</td>
<td>3832</td>
</tr>
<tr>
<td>adj. $R^2$</td>
<td>0.044</td>
<td>0.046</td>
<td>0.058</td>
</tr>
</tbody>
</table>

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Summary Statistics of Effort Measures

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Std Dev</th>
<th>Min</th>
<th>Max</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Having the Correct Answer</td>
<td>0.55</td>
<td>0.50</td>
<td>0</td>
<td>1</td>
<td>2984</td>
</tr>
<tr>
<td>Log Decision Time</td>
<td>2.91</td>
<td>2.65</td>
<td>-1</td>
<td>10</td>
<td>3832</td>
</tr>
</tbody>
</table>

The variable “Having the Correct Answer” is recorded for all users who completed the survey. Decision time is recorded for all users who completed the choice task.

with Proposition 1.

6 Evaluating the Proposed Demand Estimation Method

In this section, we use data from the field experiment to evaluate the proposed demand estimation method. The core of the method is a structural model of consumer product choice based on the decision effort mechanism we propose. We estimate the structural model drawing on choice data from the 1/2-probability and 1/30-probability groups, leaving data from the 1-probability group as the holdout sample. We then use the structural estimates to forecast demand in actual purchase settings (i.e., settings where realization probability equals 1), and compare the
Table 6: Effort Increases with Realization Probability

<table>
<thead>
<tr>
<th></th>
<th>(1) Effort as Correct Answer</th>
<th>(2) Effort as Decision Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability</td>
<td>0.0550** (0.0227)</td>
<td>0.274** (0.109)</td>
</tr>
<tr>
<td>Price</td>
<td>-0.00314 (0.00638)</td>
<td>-0.0375 (0.0303)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.552*** (0.0402)</td>
<td>3.039*** (0.191)</td>
</tr>
<tr>
<td>N</td>
<td>2984</td>
<td>3832</td>
</tr>
<tr>
<td>adj. $R^2$</td>
<td>0.001</td>
<td>0.002</td>
</tr>
</tbody>
</table>

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

forecast with demand in the holdout sample. To assess the value of having a theory-based model, we also compare the structural forecast with simple extrapolation of demand from the 1/2-probability and 1/30 probability groups.

6.1 A Structural Model of Consumer Product Choice

The structural model captures the same behavioral process as the theory model of Section 3 but operationalizes it to match the empirical context. For a conservative evaluation of the proposed demand estimation method, we strive to keep the structural model parsimonious.

We let user $i$’s true valuation of the product be

$$v_i = b_0 + b_1 \text{Log-Diamond}_i + b_2 \text{VIP}_i + e_{vi},$$

where $e_{vi}$ represents the unobserved heterogeneity in consumers’ true product valuation, which follows a normal distribution $N(0, \sigma^2_v)$. Recall that $\text{Log-Diamond}_i = \log(\text{Diamond}_i + 1)$, where $\text{Diamond}_i$ is the number of diamonds user $i$ has at the time of the experiment. $\text{VIP}_i$ denotes the VIP level of user $i$ at the time of the experiment, which is determined by how much this user has spent in the game. For the ease of interpreting the parameter estimates, we scale both $\text{Log-Diamond}_i$ and $\text{VIP}_i$ to $[0, 1]$ by dividing each variable by its maximum value. We
conjecture that a user with more diamonds at hand is likely to have a higher willingness-to-pay for the product. The sign of VIP is \textit{a priori} ambiguous. A user who has spent a lot may be more likely to spend on the new product out of habit, or less likely to spend because she already owns high-quality players contained in the player package.

User $i$’s prior belief about her product valuation, $\mu_{0i}$, follows the normal distribution $N(v_i, \sigma^2_{0i})$, where the prior uncertainty $\sigma_{0i}$ is operationalized as

$$\sigma_{0i} = \exp(a_0 + a_1 \text{VIP}_i). \quad (5)$$

We use the exponential function here to guarantee that $\sigma^2_{0i}$ is positive. We expect VIP to have a negative coefficient because, other things being equal, more spending arguably means greater experience with the game and hence less uncertainty about product valuation.

Knowing her prior mean valuation of the product $\mu_{0i}$ and her prior uncertainty $\sigma_{0i}$, user $i$ can derive her optimal level of effort in the same way as in the theory model:

$$t_i = \min \left\{ \frac{r_i}{c_i} \left( \mathbb{E}[(v_i - p_i)^+] - (\mu_{0i} - p_i)^+ \right), 1 \right\} \quad (6)$$

where the expectation is taken over consumer $i$’s belief that $v_i \sim N(\mu_{0i}, \sigma^2_{0i})$. $p_i$ and $r_i$ are the price and realization probability that user $i$ is randomly assigned in the experiment. We restrict effort $t_i$ to be no larger than 1 because it is defined as the probability that the consumer will learn her true valuation (see Section 3). We further operationalize user $i$’s effort cost $c_i$ as

$$c_i = \exp(c_1 + c_2 e_{ci}), \quad (7)$$

where $e_{ci} \sim N(0, 1)$. The exponential transformation guarantees that effort cost is positive. The $e_{ci}$ term allows effort cost to be heterogeneous among consumers.

Given her effort level $t_i$, with probability $t_i$, user $i$ learns her true product valuation $v_i$ and should buy the product if $v_i \geq p_i$. With probability $1 - t_i$, user $i$ retains her prior belief and should buy if $\mu_{0i} \geq p_i$. We assume that users have a response error when making purchase
decisions, and the response error follows i.i.d. standard Type I extreme value distribution. It follows that user $i$’s probability of choosing “willing to buy” is given by the standard logit formula:

$$\Pr(Buy_i = 1) = t_i \frac{\exp(v_i - p_i)}{1 + \exp(v_i - p_i)} + (1 - t_i) \frac{\exp(\mu_i - p_i)}{1 + \exp(\mu_i - p_i)}.$$  

(8)

The log-likelihood function of the observed purchase decision data is

$$LL = \sum_{i=1}^{N} \left[ 1(Buy_i = 1) \log \Pr(Buy_i = 1) + 1(Buy_i = 0) \log (1 - \Pr(Buy_i = 1)) \right].$$  

(9)

The above formulation of the log-likelihood function does not rely on actual data on consumer effort choices. Instead, it calculates effort choices based on model parameters following the process described in the theory model. Recall that we do have measures of effort from the field experiment. We could in theory incorporate these measures to derive additional moments for the estimation. Again, for a conservative evaluation of the proposed demand estimation method, we deliberatively avoid relying on effort data for model calibration. In fact, we would like the model to forecast well in the absence of effort data, which will lower the data requirement and broaden the applicability of the proposed demand estimation method.

### 6.2 Estimation Procedure

The structural model is estimated using the simulated maximum-likelihood estimation approach (Train 2009). For a given set of parameter values, we calculate the purchase probability of each user averaged over a large number of pre-simulated random draws, and then calculate the log-likelihood by summing up the log-likelihood of each user. The estimated parameter values are found by maximizing the simulated log-likelihood. The standard error is estimated using the inverse of Hessian matrix at the estimated parameter values. We present the detailed estimation procedure in the Appendix.

We use data from conditions where realization probability equals 1/30 or 1/2 to estimate the model parameters. We leave the 1-probability condition as the hold-out sample to assess
the forecast ability of the proposed demand estimation method. We do not use data from
the 0-probability condition in estimation for two reasons. First, our theory predicts that the
consumer can make any choice decision in this purely hypothetic setting. Thus we need to make
further tie-breaking assumptions to interpret data from this condition. For instance, we could
estimate an additional parameter that captures consumers’ tendency to act on their true beliefs
when indifferent. The identification of this parameter, however, still has to rely on information
from the 1/2-probability and 1/30-probability groups. Second, we include the 0-probability
group in the field experiment to assess how hypothetical surveys perform compared with other
demand estimation methods within the same empirical context. Application of the proposed
demand estimation method, however, does not require data from the 0-probability group. We
exclude this group from the estimation to keep the method “lean” in terms of data requirement.
Nevertheless, we verify that including the 0-probability group does not change the estimation
results significantly.

6.3 Identification

The parameters we need to estimate are the constant and coefficients in users’ true valuation
\((b_0, b_1, b_2)\), the constant and coefficient in users’ prior uncertainty \((a_0, a_1)\), the parameters deter-
mining users’ effort cost \((c_1, c_2)\), and the standard deviation of users’ unobserved heterogeneity
in true valuation \((\sigma_v)\). \(b_0\) is identified from the overall level of demand. \(b_1, b_2\) are identified from
the exogenous variation in users’ characteristics (Log-Diamond and VIP) and their difference
in purchase tendency. \((a_0, a_1)\) and \((c_1, c_2)\) together determine users’ optimal effort and thus
determine the difference in intercepts and slopes of demand curves under different realization
probabilities. \((a_0, a_1)\) can be separately identified from \((c_1, c_2)\) because it not only enters into
the formulation of optimal effort, but also determines the variance of prior beliefs and hence
determines the slope of the demand curve when no effort is expended. \(a_1\) is separately identified
from how the gap in demand curves differs for users with different VIP levels. Since every user
only makes one purchase decision in our data, the unobserved heterogeneity \(\sigma_v\) cannot be iden-
tified using the systematic difference in users’ behavior. The estimated value actually measures the part of heterogeneity in valuation that cannot be captured by \( b_1 \text{Log-Diamond}_i + b_2 \text{VIP}_i \) and is identified from the slope of the demand curve.

### 6.4 Estimation Results

Table 7 reports the parameter estimates and their standard errors. Users’ true valuation of the product, not surprisingly, increases with the amount of currency they own in the game \((b_1 > 0, p < 0.001)\). Users’ true valuation of the product also decreases with the VIP level \((b_2 < 0, p = 0.08)\). One explanation is that users with higher VIP levels tend to have spent more in the game and, as a result, are more likely to have staffed their teams with high-quality players already, so that the new player package we sell is of less value to them. In addition to these observed variations, there is significant unobserved heterogeneity in users’ true valuation \((\sigma_v > 0, p < 0.10)\). The magnitude of this unobserved heterogeneity is nontrivial given that Log-Diamond and VIP are both normalized to \([0, 1]\) in the estimation. Moving on to prior uncertainty, as expected, users with higher VIP levels are more certain about their valuation of the product \((a_1 < 0, p = 0.06)\). Finally, the effort cost parameter \(c_1\) is positive and significant \((p < 0.01)\), but the heterogeneity term \(c_2\) is almost zero. These results suggest that decision effort is costly, and similarly costly to all users in this empirical context.

To put the estimation results in context, we calculate each user’s optimal effort level \((t^*_i)\) and total cost of effort \((c_it^*_i^2/2)\) based on the parameter estimates. Table 8 presents the mean and standard deviation of these measures by experiment condition. Both the optimal effort level and effort cost increase with realization probability. They equal zero in the hypothetical condition by definition. In the actual choice condition with realization probability equal to 1, users on average spend an effort of 0.447 out of a normalized range of 0 to 1. This translates into an average effort cost of $12.77 in the actual choice condition. Recall that the average price of the product in the experiment is $24. Roughly speaking, users on average incur a cost equivalent to 53% of product price to learn their product valuation in the actual choice
### Table 7: Estimation Results

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>S.E.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$b_0$</td>
<td>-3.522</td>
<td>1.196</td>
</tr>
<tr>
<td>$b_1$</td>
<td>14.387</td>
<td>1.824</td>
</tr>
<tr>
<td>$b_2$</td>
<td>-8.812</td>
<td>5.036</td>
</tr>
<tr>
<td>$\sigma_v$</td>
<td>0.645</td>
<td>0.370</td>
</tr>
<tr>
<td>$a_0$</td>
<td>4.872</td>
<td>0.343</td>
</tr>
<tr>
<td>$a_1$</td>
<td>-3.109</td>
<td>1.672</td>
</tr>
<tr>
<td>$c_1$</td>
<td>3.310</td>
<td>1.026</td>
</tr>
<tr>
<td>$c_2$</td>
<td>1.534e-07</td>
<td>0.775</td>
</tr>
</tbody>
</table>

Prior uncertainty

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>S.E.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_0$</td>
<td>4.872</td>
<td>0.343</td>
</tr>
<tr>
<td>$a_1$</td>
<td>-3.109</td>
<td>1.672</td>
</tr>
</tbody>
</table>

Effort cost

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>S.E.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$c_1$</td>
<td>3.310</td>
<td>1.026</td>
</tr>
<tr>
<td>$c_2$</td>
<td>1.534e-07</td>
<td>0.775</td>
</tr>
</tbody>
</table>

**N**: 1842

Log-Likelihood: -1238.69

Notes. The sample consists of conditions in which realization probability equals 1/30 or 1/2. Prices are normalized to {4, 5, 6, 7, 8}. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

6.5 Forecasting Demand in Real Purchase Settings

Based on the parameter estimates, we simulate the purchase decision of each individual assuming realization probability equals 1 in the structural model (see the Appendix for details of the simulation). The simulation results form our forecast of demand in real purchase settings. We
compare the forecast against actual demand in the hold-out sample, that is, the 1-probability group we have set aside. To put the forecast in context, we also compare it with actual demand in the 1/30-probability and 1/2-probability groups. For the ease of visualization, we fit a linear demand curve based on the forecast and for each of these probability groups.

Figure 4 presents the comparison. Consistent with prior findings from the literature, the hypothetical approach (i.e., the 0-probability group) performs poorly; it overestimates demand considerably and it underestimates the degree of price sensitivity. Incentive alignment (i.e., the 1/30-probability and 1/2-probability groups) improves forecast accuracy, especially if realization probability is higher (1/2 as opposed to 1/30). The structural forecast generates a demand curve the closest to the actual demand curve of the hold-out sample.

Figure 4: Actual and Forecast Demand Curves

A natural question at this point is whether one can forecast demand using simple extrapolation methods instead of a complex structural model – one can use data from the two interim probability groups and extrapolate to the case where realization probability equals 1. To answer
this question, we examine two extrapolation methods. The first method, linear extrapolation, estimates an individual-level linear regression model of purchase decisions as a function of price, realization probability (1/30 or 1/2), their interaction terms, as well as observed user characteristics (Log-Diamond and VIP). The estimates then allow for extrapolation of purchase decisions to the case of realization probability being 1. The second method is based on the same idea but replaces the linear model with a logit model to capture the binary nature of the purchase decision.

We plot the fitted demand curve based on linear extrapolation in Figure 4. The logit version is very close and is omitted from the figure for the clarity of presentation. This fitted demand curve is closer to actual demand than the raw demand curves in the two interim probability groups. However, it performs notably worse than structural extrapolation. This is true although these simple extrapolation methods use exactly the same data as the structural forecast. Structural forecast performs better here because it uses the data in a better way by imposing a theoretically sound and empirically valid behavioral process.

We formally quantify and compare the forecast performance of the various demand estimation methods presented in Figure 4. The first column of Table 9 reports estimated price sensitivity, the metric we have focused on throughout the paper. The hypothetical method performs the worst. It generates a price sensitivity estimate nearly 70% lower than its actual value in real purchase settings. Incentive alignment with realization probabilities of 1/30 and 1/2 perform better, but still produce noticeable forecast errors. The forecast error reduces to around 20% for linear and logit extrapolations, and is only 4.62% for the structural forecast.

Besides price sensitivity, Figure 4 suggests that the various forecast methods may have overestimated the level of demand. We perform a likelihood ratio (LR) test to determine the overall fit of forecast demand to actual demand. For each forecast method \( k \in \{\text{Prob} = 0, \text{Prob} = 1/30, \text{Prob} = 1/2, \text{Linear, Logit, Structural}\} \), its likelihood ratio is calculated as \( LR_k = -2[LL_{\text{Pooled}} - (LL_{\text{Actual}} + LL_k)] \), where \( LL \) represents the log-likelihood of a linear demand curve based on observed purchases or simulated purchase probability. The likelihood ratio follows a chi-square distribution with degrees of freedom equal to the difference in the
Table 9: Performance Comparison across Demand Estimation Methods

<table>
<thead>
<tr>
<th>Method</th>
<th>Price Sensitivity</th>
<th>Likelihood Ratio (vs. Actual)</th>
<th>Optimal Price</th>
<th>Profit Loss</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prob=0 (Hypothetical)</td>
<td>0.0197</td>
<td>197.43***</td>
<td>$76.58</td>
<td>100%</td>
</tr>
<tr>
<td>Prob=1/30 (Incentive Alignment)</td>
<td>0.0305</td>
<td>28.52***</td>
<td>$40.87</td>
<td>71.84%</td>
</tr>
<tr>
<td>Prob=1/2 (Incentive Alignment)</td>
<td>0.0413</td>
<td>21.77***</td>
<td>$32.72</td>
<td>22.96%</td>
</tr>
<tr>
<td>Linear Extrapolation</td>
<td>0.0512</td>
<td>16.95***</td>
<td>$28.45</td>
<td>8.19%</td>
</tr>
<tr>
<td>Logit Extrapolation</td>
<td>0.0543</td>
<td>15.87***</td>
<td>$27.43</td>
<td>5.77%</td>
</tr>
<tr>
<td>Structural Extrapolation</td>
<td>0.0625</td>
<td>2.93</td>
<td>$23.81</td>
<td>0.58%</td>
</tr>
<tr>
<td>Prob=1 (Actual Demand)</td>
<td>0.0655</td>
<td>0</td>
<td>$22.12</td>
<td>0</td>
</tr>
</tbody>
</table>

* $p < 0.10$, i.e., $LR > \chi^2(0.9, 2) = 4.6052$;
** $p < 0.05$, i.e., $LR > \chi^2(0.95, 2) = 5.9915$;
*** $p < 0.01$, i.e., $LR > \chi^2(0.99, 2) = 9.2103$.

The number of free parameters, which is 2 in our case. The second column of Table 9 reports the likelihood ratio of each forecast method relative to actual demand. We cannot reject the null hypothesis that the structural forecast coincides with actual demand, whereas all the other methods significantly deviate from actual demand at the $p < 0.01$ level.

To see the practical value of the proposed demand estimation method, we calculate the optimal price implied by the actual demand and by the various forecast methods, respectively. Suppose the fitted demand curve is $D(p) = \alpha_0 + \alpha_1 p$ and the marginal cost of production is $mc$, then the optimal price is $p^* = \frac{mc}{2} - \frac{\alpha_0}{2\alpha_1}$. The third column of Table 9 presents the optimal price implied by the coefficients of each demand curve, assuming a marginal cost of 0 (which is a reasonable assumption considering the digital nature of the product we sell in the field experiment). Furthermore, by plugging the optimal price recommended by each method into the actual demand curve, we can calculate the expected demand and thus the expected profit if that price is charged. Comparing the expected profit to the optimal profit in the actual demand condition, we obtain the percentage profit loss associated with each method. The last column of Table 9 presents the results. By recommending an excessively high price, the hypothetical method leads to no demand and a 100% profit loss in this particular empirical setting. Incentive alignment performs better. Simple extrapolations based on incentive-alignment data introduces further improvement, with logit extrapolation reducing the profit loss to 5.77%. However,
structural extrapolation takes the forecast accuracy to yet another level. It cuts the profit loss to 0.58%, which is about one tenth of the loss under logit extrapolation.

To summarize, our proposed demand estimation method forecasts actual demand reasonably well. It forecasts actual demand significantly better than hypothetic surveys, and incentive-aligned choice experiments of moderate to small realization probabilities. Moreover, it forecasts actual demand significantly better than simple extrapolations of these incentive-aligned choice experiments to real purchase settings. We have strived to keep the structural model parsimonious for this first test of the proposed demand estimation method. The method’s forecast accuracy may further improve if we enrich the structural model by, for instance, allowing for more sources of consumer heterogeneity.

Finally, it is worth noting that the proposed demand estimation method represents significant savings in market research costs compared with test marketing. For a conservative assessment, let us abstract away from the higher logistic overhead of test marketing and focus solely on the opportunity cost of selling products at suboptimal prices. Suppose a sample size of \( n \) purchase decisions is required. To run test marketing, the company must prepare to sell \( n \) products at suboptimal prices. To implement the proposed demand estimation method, let us assume the company gathers \( n/2 \) observations from the 1/30-probability and 1/2-probability groups each. It follows that the company needs to prepare \( n/60 \) products for the 1/30-probability group and \( n/4 \) product for the 1/2-probability group. This translates into a 73% savings in products required compared with test marketing. The company may be able to save even more by optimizing the allocation of sample size across probability conditions, and by choosing the probability values wisely.

7 Concluding Remarks

In this paper, we proposed a theory-based, cost-effective method to estimate product demand prior to launch. The proposed method draws on data from incentive aligned choice experiments, but imposes structure on the data via a decision effort mechanism. This method allows us to rely
on moderate to small realization probabilities to form reasonably accurate forecast of demand in actual purchase settings.

There are several ways to extend this research. First, the current study chooses realization probabilities somewhat arbitrarily for a first test of the proposed method. As mentioned in the previous section, it will be interesting to investigate the optimal choice of realization probabilities, as well as the size of each probability group. Second, we focus on price as the only product attribute for a clean illustration of our proposed method. It will be rewarding to extend the method to settings of multi-attribute products. Last but not least, formally modeling the role of decision effort on choices may shed new light on the the design of choice experiments in other contexts.
References


Appendix

A.1 Proof of Proposition 1

Proof. Consumer \(i\) observes \(\mu_{0i}\), \(g\), \(p\), and \(r\) when choosing her optimal effort level. She also knows that \(v_i = \mu_{0i} - e_i\), and thus \(v_i - p \geq 0\) is equivalent to \(e_i \leq \mu_{0i} - p\). Rearranging Equation (2), the consumer’s optimal effort level is

\[
t^*(\mu_{0i}; p, r) = \frac{r}{c} \left( \int_{-\infty}^{\mu_{0i} - p} (\mu_{0i} - e_i - p)g(e_i)de_i - (\mu_{0i} - p)^+ \right)
\]

where the second case in (A1) is derived from the fact that \(\mu_{0i} - p = \int_{-\infty}^{\infty}(\mu_{0i} - p - e_i)g(e_i)de_i\) which holds because \(\int_{-\infty}^{\infty} g(e_i)de_i = 1\) by definition and \(\int_{-\infty}^{\infty} e_i g(e_i)de_i = 0\) by assumption.

First, consider the case of \(\mu_{0i} - p < 0\). When \(e_i < \mu_{0i} - p\), the first term of the integrand in (A1), \(\mu_{0i} - e_i - p\), is positive. Because \(g(\cdot)\) is continuous, as long as \(\mu_{0i} - p\) is strictly within the support of \(g(\cdot)\), there exists \(e_i \in (-\infty, \mu_{0i} - p)\) such that \(g(e_i) > 0\) and that the integral in (A1) is positive, which implies that \(t^*(\mu_{0i}; p, r)\) increases with \(r\).

Meanwhile, we obtain \(\frac{\partial t^*(\mu_{0i}; p, r)}{\partial (\mu_{0i} - p)} = \frac{r}{c} \int_{-\infty}^{\mu_{0i} - p} g(e_i)de_i\) and \(\frac{\partial^2 t^*(\mu_{0i}; p, r)}{\partial (\mu_{0i} - p)^2} = \frac{1}{c} \int_{-\infty}^{\mu_{0i} - p} g(e_i)de_i\). Both terms are positive as long as \(\mu_{0i} - p\) is strictly within the support of \(g(\cdot)\). This means \(t^*(\mu_{0i}; p, r)\) decreases with \(|\mu_{0i} - p|\), and the effect is is amplified when \(r\) increases, as long as \(\mu_{0i} - p\) is strictly within the support of \(g(\cdot)\).

Second, consider the remaining case of \(\mu_{0i} - p \geq 0\). When \(e_i > \mu_{0i} - p\), the first term of the integrand in (A1), \(e_i - (\mu_{0i} - p)\), is positive. Because \(g(\cdot)\) is continuous, as long as \(\mu_{0i} - p\) is strictly within the support of \(g(\cdot)\), there exists \(e_i \in (\mu_{0i} - p, \infty)\) such that \(g(e_i) > 0\) and that the integral in (A1) is positive, which implies that \(t^*(\mu_{0i}; p, r)\) increases with \(r\).

Meanwhile, we obtain \(\frac{\partial t^*(\mu_{0i}; p, r)}{\partial (\mu_{0i} - p)} = \frac{r}{c} \int_{-\infty}^{\mu_{0i} - p} (-g(e_i))de_i\) and \(\frac{\partial^2 t^*(\mu_{0i}; p, r)}{\partial (\mu_{0i} - p)^2} = \frac{1}{c} \int_{\mu_{0i} - p}^{\infty} (-g(e_i))de_i\). Both terms are negative as long as \(\mu_{0i} - p\) is strictly within the support of \(g(\cdot)\). This means
$t^*(\mu_0; p, r)$ decreases with $|\mu_0 - p|$, and the effect is amplified when $r$ increases, as long as $\mu_0 - p$ is strictly within the support of $g(\cdot)$. ■

A.2 Proof of Proposition 2

Proof. Based on equation (3) and (A1)

$$D(p, r) = \int_p^\infty \int_{p-v_i}^\infty g(e_i)de_i f(v_i)dv_i + \int_p^\infty \int_\infty^{p-v_i} t^*(v_i + e_i; p, r)de_i f(v_i)dv_i +$$

$$\int_\infty^{p-v_i} \int_p^\infty (1 - t^*(v_i + e_i; p, r))de_i f(v_i)dv_i$$

$$= \int_p^\infty \int_{p-v_i}^\infty g(e_i)de_i f(v_i)dv_i +$$

$$\int_p^\infty \int_{-\infty}^{v_i + e_i - p} \frac{r}{\epsilon} \int_{-\infty}^{v_i} (v_i + e_i - \tilde{e}_i - p)g(\tilde{e}_i)d\tilde{e}_i de_i f(v_i)dv_i +$$

$$\int_\infty^{p-v_i} \int_{p-v_i}^\infty \left(1 - \frac{r}{\epsilon} \int_{v_i + e_i - p}^{\infty} (\tilde{e}_i - (v_i + e_i - p))g(\tilde{e}_i)d\tilde{e}_i \right)de_i f(v_i)dv_i \quad (A2)$$

Notice that in the second and the third integrals, the first inner layer is to calculate $t^*$ and the integral element is $\tilde{e}_i$, whereas the second inter layer’s integral element is $e_i$ and it determines the value of $\mu_0_i$.

We first calculate $\frac{\partial D(p, r)}{\partial r}$ to investigate how $D(p, r)$ changes with $r$,
Based on equation (A2),

\[
\frac{\partial D(p, r)}{\partial r} = \frac{1}{c} \int_{\infty}^{\infty} \int_{p}^{-v_i} \int_{v_i + e_i - p}^{\infty} (v_i + e_i - \tilde{e}_i - p) g(\tilde{e}_i) d\tilde{e}_i g(e_i) f(v_i) dv_i - \\
\frac{1}{c} \int_{-\infty}^{\infty} \int_{-\infty}^{v_i - e_i - p} \int_{v_i + e_i - p}^{\infty} (\tilde{e}_i - (v_i + e_i - p)) g(\tilde{e}_i) d\tilde{e}_i g(e_i) f(v_i) dv_i. \quad (A3)
\]

From (A3) to (A4), we substitute the integral element to \( x = v_i + e_i - p - \tilde{e}_i \) in the first part and substitute the integral element to \( x = \tilde{e}_i - (v_i + e_i - p) \) in the second part. From (A4) to (A5), we substitute the integral element to \( y = e_i - (p - v_i) \). Then we replace \( y \) with \( -y \) in the first part and get (A6). From (A6) to (A7) we substitute the integral element to \( z = v_i - p \), and then get (A8) by replacing \( z \) with \( -z \). Based on our assumption that \( g(e) = g(-e), \forall e \), we further get (A9).
Define $H(z) = \frac{1}{c} \int_0^\infty \int_0^\infty xg(y + x)dxg(y + z)dy$. Then

$$\frac{\partial D(p, r)}{\partial r} = \int_0^\infty H(z)[f(p + z) - f(p - z)]dz. \quad (A10)$$

We first consider the case of $p > \mu_v$. Since $p + z - \mu_v > p - z - \mu_v$ and $p + z - \mu_v > \mu_v - p + z$ for any $z > 0$, and hence $|p + z - \mu_v| > |p - z - \mu_v|$. Recall that we assume $f(\cdot)$ has a unique mode $\mu_v$, and $f(\cdot)$ is weakly increasing on $(-\infty, \mu_v]$ and weakly decreasing on $[\mu_v, \infty)$. Given that $p + z$ is further away from the mode $\mu_v$, we have $f(p + z) - f(p - z) \leq 0$. Noticing that $H(z) \geq 0$, we get that

$$\frac{\partial D(p, r)}{\partial r} = \int_0^\infty H(z)[f(p + z) - f(p - z)]dz \leq 0. \quad \text{Since} \ f(\cdot) \text{cannot be constant throughout the real line, then for any } p \in (\mu_v, \infty), \text{there must exist } z > 0 \text{ such that } f(p + z) - f(p - z) < 0. \text{ Denote } Z^-(p) = \{z > 0 : f(p + z) - f(p - z) < 0\}, \ S_g = \{z > 0 : g(z) > 0\}. \text{ If the (Lebesgue) measure of } Z^-(p) \cap S_g \text{ is greater than 0, then the integral } \int_0^\infty H(z)[f(p + z) - f(p - z)]dz \text{ is strictly negative.}$$

Similarly, we can prove that when $p < \mu_v, f(p + z) - f(p - z) \geq 0$, so

$$\frac{\partial D(p, r)}{\partial r} = \int_0^\infty H(z)[f(p + z) - f(p - z)]dz \geq 0, \text{ and it is positive when } p \text{ satisfies the condition that } Z^+(p) \cap S_g \text{ has a positive (Lebesgue) measure, where } Z^+(p) = \{z > 0 : f(p + z) - f(p - z) < 0\}. \text{ Now we calculate } \frac{\partial}{\partial r}\left(\frac{\partial D(p, r)}{\partial p}|_{p=\mu_v}\right) \text{ to see how the central slope of demand curve changes with } r. \text{ It is easy to see that } \frac{\partial}{\partial r}\left(\frac{\partial D(p, r)}{\partial p}|_{p=\mu_v}\right) = \frac{\partial^2 D(p, r)}{\partial p \partial r}|_{p=\mu_v}. \text{ According to } (A10) \text{ and the}}$$

A1: If $f(\cdot)$ is constant throughout the real line, $\int_{-\infty}^{\infty} f(v)dv = 0 \text{ or } \pm \infty$, which conflicts with $\int_{-\infty}^{\infty} f(v)dv = 1$. 

A-4
theorem of integration by parts, we have

\[
\frac{\partial^2 D(p, r)}{\partial p \partial r} = \frac{\partial}{\partial p} \int_0^\infty H(z)[f(p + z) - f(p - z)]dz
\]

\[
= \int_0^\infty H(z)d_z[f(p + z) + f(p - z)]
\]

\[
= H(z)[f(p + z) + f(p - z)]|_{z=0}^\infty - \int_0^\infty [f(p + z) + f(p - z)]dH(z) \quad (A11)
\]

\[
= -2H(0)f(p) - \int_0^\infty [f(p + z) + f(p - z)]dH(z) \quad (A12)
\]

\[
= 2f(p)\int_0^\infty dH(z) - \int_0^\infty [f(p + z) + f(p - z)]dH(z)
\]

\[
= \int_0^\infty [2f(p) - f(p + z) - f(p - z)]dH(z) \quad (A13)
\]

The reasoning from \([A11]\) to \([A12]\) is as follows. By definition of p.d.f, \(\int_{-\infty}^{\infty} g(z)dz = 1\), so we must have \(\lim_{z \to \pm \infty} g(z) = 0\). Then \(\lim_{z \to \infty} g(y + z) = 0\) for any \(y > 0\), and therefore \(\lim_{z \to \infty} H(z) = 0\). Similarly, \(\lim_{z \to \pm \infty} f(z) = 0\), and thus \(\lim_{z \to \infty} f(p \pm z) = 0\).

By \([A13]\),

\[
\frac{\partial}{\partial r} \left( \frac{\partial D(p,r)}{\partial p} \bigg|_{p=p_v} \right) = \int_0^\infty [2f(\mu_v) - f(\mu_v + z) - f(\mu_v - z)]dH(z).
\]

Recall that \(g(\cdot)\) is assumed to be symmetric around 0 and is weakly decreasing and non-constant on \((0, \infty)\). Then according to the definition of \(H(z)\), \(H(z)\) is weakly decreasing and non-constant on \((0, \infty)\). Given our assumption that \(f(\cdot)\) is weakly increasing on \((-\infty, \mu_v)\) and weakly decreasing on \((\mu_v, \infty)\), \(2f(\mu_v) - f(\mu_v + z) - f(\mu_v - z) \geq 0\). Then \(\frac{\partial}{\partial r} \left( \frac{\partial D(p,r)}{\partial p} \bigg|_{p=p_v} \right) \leq 0\), and it is guaranteed to be negative when \(Z(\mu_v) = \{z > 0 : f(\mu_v + z) + f(\mu_v - z) < 2f(\mu_v)\}\) has a non-empty intersection with the set of \(z\) that \(H(z)\) is strictly decreasing, which is the same as the set of \(z\) that \(g(z)\) is strictly decreasing. ■
A.3 Screenshots from the Field Experiment

Figure A1: Screenshot of the Choice Task (1/30-Probability, Price=2800 Diamonds)

Figure A2: Content of the Player Package
A.4 Details of Structural Estimation and Extrapolation

We first draw three $N \times K$ matrices of random numbers that are independent and identically distributed, following the standard normal distribution, where $N$ is the number of individuals, and $K = 100$ is the number of iterations we will perform to simulate the average purchase probability of each individual. The three matrices are denoted as $e_1, e_2, e_3$. The draws are actually quasi-random: we generate a two-dimensional Halton set with three columns, each column of which are evenly distributed numbers on $[0, 1]$, take the first $N \times K$ elements of each column, and then converting the numbers to standard normal distribution by taking the inverse of normal c.d.f. of them. Since Halton set is more evenly distributed on $[0, 1]$ compared to direct random draws of the uniform distribution on $[0, 1]$, the random draws created in this way leads to better convergence performance compared to direct random draws and requires less number of draws \textsuperscript{(Train 2009)}.

We also generate two $1 \times T$ vectors, which are the Gauss-Hermite quadrature nodes and weights over $[-1, 1]$, where $T = 1000$. They will be used to calculate the expectation of a function of a normally distributed random variable.

Given these draws, we can calculate the average purchase probability of each individual under a given set of parameters, and then calculate the log likelihood of the observed data. The objective function is the sum of log likelihood of the observed data.

Given a set of parameter values $(b_0, b_1, b_2, a_0, a_1, c_1, c_2, \sigma_v)$, we perform $K$ iterations of calculation. Within each iteration, the steps are as follows.

1. Simulate each individual’s true valuation $v_i = b_0 + b_1 \text{Log-Diamond}_i + b_2 \text{VIP}_i + \sigma_v e_{1ik}$, for $i = 1, \ldots, N$, where $e_1$ is an $N \times K$ matrix of i.i.d. standard normal draws which we have created at first, and $e_{1ik}$ is the element $(i, k)$ of it.

2. Calculate each individual’s prior uncertainty $\sigma_{0i} = \exp (a_0 + a_1 \text{VIP}_i)$, for $i = 1, \ldots, N$.

3. Simulate each individual’s prior belief $\mu_{0i} = v_i + \sigma_{0i} e_{2ik}$, for $i = 1, \ldots, N$, where $e_2$ is also an $N \times K$ matrix of i.i.d. standard normal draws. Thus we have $\mu_{0i} \sim N(v_i, \sigma_{0i}^2)$.
4. Calculate $\mathbb{E}[(v_i - p_i)^+]$, $i = 1, ..., N$, where $p_i$ is the price assigned to $i$. The expectation is taken over each individual $i$’s belief about the distribution of $v_i$, which is $N(\mu_0i, \sigma_0^2)$. To get better convergence performance, we use Gauss-Hermite quadrature method. That is, if a random variable $Y \sim N(\mu, \sigma^2)$, $\mathbb{E}[f(Y)] \approx \frac{1}{\sqrt{\pi}} \sum_{j=1}^T w_j f(\mu + \sqrt{2}\sigma x_j)$, where $x_j, w_j$ are the Gauss-Hermite quadrature nodes and weights over $[-1, 1]$. In our case, $\mathbb{E}[(v_i - p_i)^+] \approx \frac{1}{\sqrt{\pi}} \sum_{j=1}^T w_j \cdot (\mu_0i + \sqrt{2}\sigma_0i x_j - p_i)^+$. 

5. Simulate each individual’s effort cost $c_i = \exp(c_1 + c_2 e_3i)$, where $e_3$ is the standard normal matrix that we have drawn.

6. Calculate each individual’s effort level $t_i = \min \left\{ \frac{\tau_i}{c_i} (\mathbb{E}[(v_i - p_i)^+]) - (\mu_0i - p_i)^+), 1 \right\}$, where $\tau_i$ is the realization probability assigned to $i$. $c_i, \mathbb{E}[(v_i - p_i)^+], \mu_0i$ have been simulated in previous steps.

7. Calculate each individual’s purchase probability $Pr_k(Buy_i = 1) = t_i \frac{\exp(v_i - p_i)}{1 + \exp(v_i - p_i)} + (1 - t_i) \frac{\exp(\mu_0i - p_i)}{1 + \exp(\mu_0i - p_i)}$.

After the $K$ iterations, for each individual, we average over the iterations to get the individual’s purchase probability $Pr(Buy_i = 1) = \frac{1}{K} \sum_{k=1}^K Pr_k(Buy_i = 1)$, and then calculate the sum of log likelihood $LL = \sum_{i=1}^N \left[ 1(Buy_i = 1) \log Pr(Buy_i = 1) + 1(Buy_i = 0) \log (1 - Pr(Buy_i = 1)) \right]$. Being able to calculate the simulated log likelihood given a set of parameter values, we search over the parameter space and find the set of parameter values that maximizes the simulated log likelihood. We restrict the value of $\sigma_v$ and $c_2$ to be positive, since they represent the standard deviations of a normal distribution and a log-normal distribution. Only the data of $r = 1/30, 1/2$ conditions are used to perform the estimation.

Given the parameter estimates, we follow the same steps as described above to calculate the purchase probability of each individual in the $r = 1$ condition: first draw random numbers, and then go over $K$ iterations to simulate each individual’s purchase probability under the estimated parameter values. Lastly, we aggregate the purchase probability and plot the demand curve of the structural forecast.