ESSAYS ON CUSTOMER PREFERENCE MEASUREMENT

A Dissertation in
Business Administration

by

Li Xiao

© 2013 Li Xiao

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

May 2013
The dissertation of Li Xiao was reviewed and approved* by the following:

Min Ding  
Smeal Professor of Marketing and Innovation  
Dissertation Advisor  
Chair of Committee  

Rajdeep Grewal  
Irving & Irene Bard Professor of Marketing  

Mosuk Chow  
Senior Research Associate  
Associate Professor of Statistics  

Robert Collins  
Associate Professor of Computer Science  
Co-Director of Laboratory for Perception, Action, and Cognition (LPAC)  

Fuyuan Shen  
Associate Professor of Advertising/Public Relations  

Duncan Fong  
Professor of Marketing and Professor of Statistics  
Head of the Department of Marketing  

*Signatures are on file in the Graduate School
ABSTRACT

Customer preference measurement has always been an active area in marketing research. Essay 1 makes first attempt to explore customers’ preferences toward different faces in print advertisement context. It aims to answer three questions that are important to both researchers and practitioners: 1) Do faces affect how a viewer reacts to an advertisement in the metrics that advertisers care about? 2) If faces do have an effect, is the effect large enough to warrant a careful selection of faces when constructing print advertisements? 3) If faces do have an effect and the effect is large, what facial features are eliciting such differential reactions on these metrics, and are such reactions different across individuals? Relying on eigenface method, a holistic approach widely used in the computer science field for face recognition, we conducted two empirical studies to answer these three questions. The results show that different faces do have an effect on people’s attitudes toward the advertisement, attitudes toward the brand, and purchase intentions; and the effect is non-trivial. Multiple segments were found for each key advertisement metric, and substantial heterogeneity in people’s reactions to the ads was revealed among those segments. Implications and directions for future research are discussed.

Essay 2 aims to explore customers’ preferences toward different service innovations. In this essay, we design and validate a mechanism for service firms, called the quasi-patent (qPatent) system. The qPatent system builds upon both principles of the patenting system and unique characteristics of services using state-of-art incentive aligned conjoint analysis. It provides an environment where a firm can incent potential outside inventors to develop service innovations that the firm desires, in a way that innovations addressing the needs of the firm will be protected and rewarded financially based on their market value. We demonstrate the application of the qPatent system in the context of developing a tour package for American tourists visiting
Shanghai, China. It is shown to be capable of generating new service offerings that are more valuable to the firm than existing offerings for various segments of potential customers.

**Key words**: preference measurement; face; facial features; eigenface; service innovation; patent
TABLE OF CONTENTS

List of Figures .................................................................................................................. vii
List of Tables .................................................................................................................. viii
Acknowledgements ......................................................................................................... ix

Chapter 1 Introduction ..................................................................................................... 1

Chapter 2 Essay 1: Explore the Effects of Facial Features in Print Advertising .......... 3
  2.1 Introduction .............................................................................................................. 3
  2.2 Relevant Literature ................................................................................................. 5
     Existing Literature on Face Research ......................................................................... 5
     The Eigenface Method ............................................................................................... 7
     Heterogeneity in Viewers’ Responses to Faces ......................................................... 9
  2.3 Study 1 .................................................................................................................... 12
     Stimulus Advertisements ......................................................................................... 12
     Extraction of Facial Features .................................................................................... 12
     Experimental Procedures ........................................................................................ 15
  2.4 Estimation and Results ........................................................................................... 16
     Effect of Faces on Key Ad Metrics and Effect Size at Aggregate Level .......... 16
     Segmentation Analysis using Facial Features ....................................................... 19
     Effect of Faces on Key Ad Metrics and Effect Size at the Segment Level ...... 24
  2.5 Study 2 .................................................................................................................... 27
     Experimental Design ................................................................................................. 27
     Extraction of Facial Features .................................................................................... 29
     Experimental Procedures ........................................................................................ 30
     Results ....................................................................................................................... 30
  2.6 Conclusions and Discussions ............................................................................... 33

Chapter 3 ......................................................................................................................... 35

A quasi-Patent (qPatent) System for Service Innovation .............................................. 35
  3.1 Introduction ............................................................................................................ 35
  3.2 Literature Review .................................................................................................. 38
     Patent System ........................................................................................................... 39
     Incentive Aligned Mechanism Design .................................................................. 43
     Service Innovation ................................................................................................. 44
  3.3 qPatent System ...................................................................................................... 46
     Innovation Stage ...................................................................................................... 46
     Valuation Stage ........................................................................................................ 56
     Summary and Contrast with the U.S. Patent System ............................................. 61
  3.4 An Empirical Demonstration ................................................................................ 63
     Service Innovation Context .................................................................................... 64
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Customer Needs</td>
<td>64</td>
</tr>
<tr>
<td>Participants</td>
<td>65</td>
</tr>
<tr>
<td>Compensation and Incentive Alignment</td>
<td>66</td>
</tr>
<tr>
<td>System Implementation and Procedure</td>
<td>69</td>
</tr>
<tr>
<td>3.5 Results</td>
<td>69</td>
</tr>
<tr>
<td>Process Statistics</td>
<td>70</td>
</tr>
<tr>
<td>Raw Innovation Statistics</td>
<td>70</td>
</tr>
<tr>
<td>Conjoint Task in the Valuation Stage</td>
<td>72</td>
</tr>
<tr>
<td>Value of Innovation</td>
<td>75</td>
</tr>
<tr>
<td>3.6 Conclusions and Discussions</td>
<td>80</td>
</tr>
<tr>
<td>ROI Considerations</td>
<td>81</td>
</tr>
<tr>
<td>Alternative Incentive Mechanism</td>
<td>82</td>
</tr>
<tr>
<td>Design Variations</td>
<td>83</td>
</tr>
<tr>
<td>Special Cases for qPatent Systems</td>
<td>83</td>
</tr>
<tr>
<td>Extending to Product Innovation</td>
<td>85</td>
</tr>
<tr>
<td>References</td>
<td>86</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

Figure 2-1. Eigenface Decomposition. ...............................................................8
Figure 2-2. Average Face and 15 Eigenfaces ..................................................13
Figure 2-3. Physiognomic Decomposition.......................................................14
Figure 2-4. Average Face and Five Eigenfaces ..............................................29
Figure 3-1. Stages of the qPatent Method.....................................................47
Figure 3-2. Innovation Stage of the qPatent Method........................................52
LIST OF TABLES

Table 2-1. Compare Key Ad metrics at Aggregate Level.........................................................17
Table 2-2. Illustration of Predicted Purchase.............................................................................19
Table 2-3. Result of Segmentation and Variable Selection Using Eigenface Decomposition. .................................................................................................................................21
Table 2-4. Result of Regression on Eigenface Loadings at Segment Level ...............................22
Table 2-5. Using Physiognomic Features to Segment Heterogeneity........................................23
Table 2-6. Using Physiognomic Features to Segment Heterogeneity........................................25
Table 2-7. Result of Regression on Eigenface Loadings at Segment Level ...............................31
Table 3-1. Types of Participants in qPatent Method....................................................................49
Table 3-2. Comparing qPatent with Existing U.S. Patent System...........................................62
Table 3-3. Summary Statistics of Raw Solutions in Innovation Stage ........................................71
Table 3-4. Summary Statistics of FA Choices.............................................................................71
Table 3-5. Summary Statistics of VOCers’ Evaluations...........................................................72
Table 3-6. Categories of New Solutions....................................................................................73
Table 3-7. One-Segment Model Estimate from the Valuation Stage.........................................77
Table 3-8. Subsample Analysis..................................................................................................79
ACKNOWLEDGEMENTS

I wish to thank, first and foremost, my advisor Min Ding, for his invaluable support and guidance throughout the PhD program. I have been very lucky to have the chance to work with him during the five years. He treats me as student, coworker, friend, and family. I would not be where I am today without his patient mentorship. I have learned a lot from him, not only how to do good research, but also and more importantly, how to be a good person. In Chinese, there is a saying “For one day as my teacher, for ever as my father”. I will be grateful to him forever.

I would also like to thank Raj Grewal, Fuyuan Shen, Robert Collins and Mosuk Chow for generously serving on my committee and helping to improve my dissertation along the way. Raj helped me to better shape and position the paper. Fuyuan helped me with advertising literature and empirical design. Bob (Robert) helped me with face processing and other computer science techniques. Mosuk helped me with statistical analysis.

I would also like to thank Ujwal Kanyande, Jing Wang, Gary Lillien, Bill Ross, Eelco Kappe, Lisa Bolton, Meg Meloy, Karen Winterich, Fred Hurvitz and Dave Winterich. Ujwal and Jing invested a lot of time and efforts on idea development and data collection for Essay 2. Gary spared his precious time to peer review both essays, and gave me many great suggestions for improvement. I also owe Gary my gratitude for his help on academic writing and presentations. Bill has always been an intelligent, enthusiastic and pleasant scholar to work with. I’ve got lots of inspirations from communication with him. Eelco kindly helped me to sharpen my research problems. Lisa is the most considerate and caring PhD advisor that a doctoral student can ever possibly have. Meg and Karen helped me a lot with subject recruitment and shared with me their experiences. Fred and Dave shared with me many useful tips to improve my teaching skills.

I also want to thank the entire Penn State marketing department for their generous support and help. Every faculty member is so nice and supportive, who would generously offer
full support whenever I have questions or problems. Our department has the best staff that I have ever met with. Steph Ironside and Elyse Pasquariello have always been of tremendous help for my every administrative request, no matter how small it could be.

I would also like to thank my fellow PhD students. The program has been such a good experience in my life because of them. The days of studying late in office with Hyejin Kim, sharing ideas with Chen Zhou, discussing problems with Huanhuan Shi, will all become my precious memory for the rest of my life.

Finally, last but not the least, I want to give my deepest gratitude to my parents. I would not have been able to pursue my dream and fulfilled my PhD program without their great understanding and generous support. I cannot find words to describe how much gratitude I owe them. I will try to live up to their expectations and be their pride in life.
Chapter 1

Introduction

Customer preference measurement has always been an active area for marketing research. American Marketing Association (AMA) (2007) defines marketing as “the activity, set of institutions, and processes for creating, communicating, delivering, and exchanging offerings that have value for customers, clients, partners, and society at large.” Customers are the most important stakeholder for most companies since they purchase products/service and thus create profit for companies. In order to create value for customers and persuade them into purchasing products/service, the first step is to understand their preferences. Only when customers’ preferences are well measured and understood, companies are able to produce and offer products/services that appeal to customers’ preferences.

Given its importance, customer preference measurement has attracted much attention from both practitioners and academia. A lot of techniques have been invented and proposed to understand and measure customer preferences in the literature. To name just a few. Green and Srinivasan (1990) propose to use conjoint analysis to reveal and quantify customers’ preferences. Conjoint analysis excels on measuring tradeoffs among multi-attributed products and services and revealing customers’ willingness to pay for each level of each attribute, and thus is widely used for new product development. Recently many scholars have proposed to use advanced computer techniques to track and record consumers’ responses so as to better understand their preferences. For example, Hauser et al. (2009) use clickstreams to infer customers’ preferences toward different websites; Texira et al. (2012a, 2012b) use eye tracking and facial expression recognition techniques to infer customers’ preferences toward different video commercials.
The two essays in the present dissertation aim to contribute to the customer preference measurement literature. Essay 1 relies on a computer science technique named eigenface method to explore customers’ preferences toward different human faces in print advertisements. Essay 2 proposes and validates a mechanism for service firms, called the quasi-patent (qPatent) system, which builds upon both principles of the patenting system and unique characteristics of services using state-of-art incentive aligned conjoint analysis.
Chapter 2  

Essay 1: Explore the Effects of Facial Features in Print Advertising  

2.1 Introduction

Print media has been playing an important role for advertising practice. According to eMarketer’s (2012) report, print advertising spending achieved $36 billion in 2011, contributing to 23% of U.S. total media ad spending. In print advertising, it is common practice to have a human endorser present the brand/product and advocate the adoption of the brand/product. A recent GFK MRI report (2001) indicated that only a small portion (less than 10%) of print ads use celebrity endorsers. For many ads that use non-celebrity endorsers, consumers know nothing about these endorsers other than their faces. Although there is a large body of literature on the effects of endorser/spokesperson on advertising effectiveness from various aspects, for example gender (e.g. Gilly 1988), ethnicity (e.g. Wheatley 1971), age (e.g. Freiden 1984), profession (e.g. Ohanian 1991), fame (e.g. Heath et al. 1994), and overall attractiveness (e.g. Bower and Landreth 2001) etc., few studies deal directly with faces. Therefore, the effect of different faces on advertising effectiveness still remains largely unknown. It has not been well studied or empirically demonstrated in advertising context yet. In the present paper, we focus on faces of non-celebrity endorsers, about whom viewers know little but their faces.

Borrowing face recognition techniques from the field of computer science, we empirically test how different faces affect advertising effectiveness in print advertising context. Specifically, we aim to address three questions that are important to both researchers and practitioners.
1) Do faces affect how a viewer reacts to an advertisement in the metrics that advertisers care about? While keeping certain aspects of the spokesperson constant, such as gender, ethnicity, age, etc., will different faces elicit different responses from viewers? Here we focus on three key metrics of advertising effectiveness that advertisers care most about, including attitude toward the ad, attitude toward the brand, and purchase intention (Miniard et al. 1990; Goldsmith et al. 2000). If the answer is no, advertisers should not spend much efforts or money on face selection.

2) If faces do have an effect, is the effect large enough (and/or does it affect the right viewers) to warrant a careful selection of faces when constructing a print advertisement? If the effect of face exists, but the effect is trivial, advertisers should not spend much time or money on face selection. Only if the effect exists and it is substantial, should they be careful selecting which face to appear in the advertisement.

3) If faces do have an effect and the effect is large, what specific facial features are eliciting such differential reactions on these metrics, and are such reactions different across individuals? This question deals with the heterogeneity of the effect of faces on viewers’ responses on key ad metrics. A critical step toward answering this question involves the decomposition of faces. As such, we relied on the eigenface method, which is a widely used approach for face recognition in the computer science field. The eigenface method is “the first really successful demonstration of machine recognition of faces” (Zhao et al. 2003, P412). Many face recognition algorithms were derived from this approach. Eigenface method will be explained in further detail in the next section.
2.2 Relevant Literature

We first provide a review of previous research on the effects of faces and facial features. Then, we introduce the eigenface method and how it works to extract facial features from different faces. Lastly, we review literature indicating that substantial heterogeneity exists in the way people make inferences from faces.

Existing Literature on Face Research

The effect of faces on people’s perceptions and behavior has been studied in multiple disciplines (see Zebrowitz 2006 for a review). In psychology, Todorov et al. (2005) showed that inferences of competence based solely on facial appearance could predict the outcomes of U.S. congressional elections to some extent; Collins and Zebrowitz (1995) found that appearance is significantly related to the type of jobs people hold and partly determines their job status; Pincott (2010) found that women use facial masculinity to infer the attractiveness of males and decide mating preferences. In marketing, Gorn et al. (2008) found that chief executive officers with a baby face are perceived to be more trustworthy and creditable than those with a mature face in a public relations crisis; Solomon et al. (1992) showed that congruency between facial attractiveness of spokesperson and product can achieve favorable communication responses. In computer science, Koda and Maes (1996) showed that faces of virtual agents have an impact on human-computer interaction, and Brahnam (2002) devoted to finding techniques to customize virtual agents’ faces to match the personality of users.

Despite the large body of literature, the effect of face on the effectiveness of print ads has not been empirically demonstrated. The most relevant literature in this domain is Bower’s (2001) study on the overall physical attractiveness of models, not specifically their faces, on product
evaluations. More importantly, the magnitude of a face effect and its practical relevance to practitioners have never been addressed.

An important stream of face research is to identify what specific facial features drive perceptions or behavior. In psychology, researchers have used a physiognomic approach to extract facial features from different faces. Physiognomic method represents a face with a set of facial distances. For example, Berry and McArthur (1985) decomposed a male face into 11 facial distances and found that a face with bigger eyes and a wider chin is judged to be more babyfaced and more trustworthy. Cunningham et al. (1990) decomposed a male face into 26 facial distances and found that a male face with bigger eyes and a longer chin are perceived as more attractive in female’s eyes. Other researchers focused on one feature, facial width-to-height ratio, and found a significant effect on perceived propensity for aggression (Carrie et al. 2009) and cooperative behavior in a trust game (Stirrat and Perrett 2010).

As the above studies illustrate, the choice of physiognomic measurement is somewhat arbitrary. For example, Berry and McArthur (1985) chose 11 physiognomic features, and Cunningham (1986) chose 26, but it is unknown exactly how many features should be studied. Furthermore, due to the model identification issue and concern of multicollinearity, both studies included some interaction terms in their models, such as eye area, eye roundness and chin area, but not all possible interactions. Many interactions might have an effect but are missing from the models, for example the ratio of eye distance to face length, the angle of the eyes to the nostril tip, etc. Based on these limitations, we decided to use a more holistic method, the eigenface method, to deconstruct faces.
The Eigenface Method

Eigenface is a method proposed by Turk and Pentland (1991) for face recognition. It uses principal component analysis (PCA) to project training face images onto a set of dominant eigenvectors. Because each dominant eigenvector looks roughly like a face, these eigenvectors are also called eigenfaces. This method is easy to implement, computationally efficient, achieves good accuracy in face recognition tasks, and thus is widely used in the computer science field (Zhao et al. 2003). Its implementation is briefly described as follows (see more details at Turk and Pentland 1991).

(a) Acquire a training set of face images. Ideally, these face images are all of the same size, in a fixed position, and facing forward under constant lighting. The faces show neutral emotion. Assume that there are M gray-scaled face images in the training set and the size of each image is H × W pixels.

(b) Stack each face image into a HW×1 vector, denoted as \( X_i \). The training set of images is represented by a HW×M matrix, denoted as \( X \).

(c) Calculate the mean \( \bar{X} = \frac{1}{M} \sum X_i \) and subtract the mean from \( X \). The mean vector \( \bar{X} \) could be unstacked back to an H × W matrix and displayed, which looks roughly like a face with blurry contour. This is called the average face.

(d) Use PCA to calculate the eigenvectors and eigenvalues of the covariance matrix. At this step, we obtain M eigenvectors and corresponding M eigenvalues.

(e) Keep K most important/dominant eigenvectors and project each image in the training set onto these K eigenvectors, stacked together and denoted as \( EV \) (HW×K). Thus, each image in the training set \( (X_i) \) can be approximated and represented by a unique loading vector, \( L_i \) (K×1). Each image can be reconstructed as \( X_i \approx EV \times L_i + \bar{X} \). The choice of K depends on how...
many eigenfaces are needed to get a good reconstruction of training images (Turk and Pentland 1991).

By storing the average face and K dominant eigenvectors and given a face image, either an image from training set or a new image, it is very easy to calculate the loading vector of this image. The dissimilarity between two face images can be well represented by the difference between two corresponding loading vectors. An illustration of the eigenface decomposition is shown in Figure 2-1.

*Figure 2-1: Eigenface Decomposition*
With average face and k eigenfaces stored in the database, any face in the training set can be compressed and reconstructed by a k×1 loading vector (L_1, ..., L_k) (Turk and Pentland 1991).

For this type of research, only Branham (2002) made the first attempt to apply the eigenface method to her virtual agent study. However, the faces used in her study were synthetic faces generated by software, not real human faces. To date, no study has applied the eigenface method to real faces and linked eigenface loading features to people’s perceptions or reactions to print advertisements.

Unlike the physiognomic method, which requires a predefined set of facial distances, no prior information is needed for the eigenface method. The extracted eigenvectors contain the major differences among training images, thus it is a holistic measurement. In terms of linking facial features to people’s reaction to faces, the eigenface method involves one drawback in that it is not as intuitive as the physiognomic method. By looking at the eigenfaces, it is very hard to tell exactly how each eigenface is different from the others.

In the present paper, we propose a sequential method. First, we employ the eigenface method to decompose real faces and extract facial features, i.e., loadings on dominant eigenfaces, and analyze the data based on the eigenface loadings. Next, we use the physiognomic method to explain the results intuitively. As such, we combine the advantages of both approaches, keeping the non-ad hoc and holistic nature of the eigenface method and allowing for an intuitive explanation by using the physiognomic method.

**Heterogeneity in Viewers’ Responses to Faces**

All of the previous studies that link facial features to people’s reactions were conducted at an aggregate level, not an individual level, and attempted to use one model to explain how people make inferences from faces. People are treated as homogeneous in these studies, and
individual differences are ignored. One underlying assumption is that people share in common what facial features influence their impressions and how important these facial features are in forming their perceptions. For example, based on Berry and McArthur’s (1985) findings, a consensus trustworthy face must have big eyes and a wide chin. The major theory that backs up this assumption is that the perception of facial features has adaptive value and that trait impressions are based on those facial qualities that demand the greatest attention for the survival of the species, namely, physical fitness, age, and emotional state. Faces that possess features that are indicative of these qualities are believed to have pronounced overgeneralization effects (Zebrowitz 1998). Thus, there is considerable consistency and universal consensus in people’s trait impression of faces (Cunningham et al. 1995).

However, this assumption has been greatly challenged. Ample evidence shows that the heterogeneity among people is not trivial and that it influences their way of making inferences from faces. A comparison between the results from Cunningham’s (1986) and Cunningham et al.’s (1990) studies reveals a gender difference in inferring attractiveness. Solomon et al. (1992) further proposed six distinct types of attractive looks. It is well documented that people attach more favorable impressions to faces from ingroup (i.e., same race, same religion, same social status, etc.) than faces from outgroup (i.e., different race, different religion, lower social status, etc.) (Rule et al. 2010; Hehman et al. 2010). Ekman and Friesen (2003) reported that culture and individual heterogeneity has an effect on both the exhibited face side and the recipient side. Based on these arguments, people should be studied in separate segments defined by such factors as religion, ethnicity, and social status, etc. Hence, multiple segments of people should be identified based on the way they make inferences from faces. Within each segment, people share similar preferences toward faces, while between segments people show substantially different preferences toward faces.
To further complicate the story, the theory of constructive preference states that people’s preferences are inconsistent and labile, varying as a function of a range of task, individual differences, and the contexts in which preferences are elicited and constructed (Payne et al. 1993; Slovic 1995). Coupey et al. (1998) empirically showed that people’s preferences could be easily reversed if the context in which these preferences were elicited was altered. People’s preferences towards faces vary from context to context and from person to person, which means that no significant effect of any single facial feature should be found.

Resolving this controversy for marketing practitioners is important, as quite different marketing strategies might need to be considered to optimize the effects of advertisements on consumer perceptions and behavior. If our results reveal that people are somehow similar and one best face exists, then an advertisement agency’s job is to find a spokesperson with a face that closely resembles the best face for the advertisement. However, if we find that people are heterogeneous in their facial preferences and multiple best faces exist with each appealing to a different segment of consumers, then advertisement agency would want to identify and define the specific target customer segments and then create multiple advertisements, each with a distinct face that closely matches the best face for that specific segment of customers. If consumers are so heterogeneous that too many best faces exist (or, in another words, no best faces actually exist), then advertisement agency would be advised to select a reasonably good face in the ad and focus on other aspects of the advertisement.
2.3 Study 1

**Stimulus Advertisements**

We conducted a focus group with four undergraduate students to select the product categories and stimulus faces to be used in each product category. Five product/service advertisement categories were chosen, including ads for staff recruiting, a restaurant, a real estate agent, a fragrance, and a ski resort. To elicit more a natural advertisement viewing experience from participants, we decided to use real ads. We then used computer technology to replace the face in the real ad with three stimulus faces while holding all other elements for each product category constant.

All stimulus faces were downloaded from the Internet. After screening by focus group, the 15 stimulus faces used in the experiment appeared to be white male faces 30 to 40 years of age, positioned right-side up, facing forward, and showing relatively neutral emotion. And three stimulus faces used for each product category were judged by the focus group participants as reasonably good to appear in the ad.

**Extraction of Facial Features**

As indicated in the previous section, we used a sequential approach, starting with the eigenface method for estimation and segmentation, and then using the physiognomic method to help explain the results. With the eigenface method, we first removed the background noise and extracted the facial image from the original image. Second, we normalized each face and placed it into a 128x128 pixels sized plane. This allowed us to construct the training set of 15 images. Third, we employed the PCA method to obtain the eigenvectors and eigenvalues. After applying a rule of thumb (keeping eigenvectors with corresponding eigenvalues ≥ 1), 14 eigenvectors
remained. These eigenvectors explained over 99% of the variance among training images. The root-mean-square pixel-by-pixel error in representing cropped versions of training face images, i.e., reconstruction errors, was less than two percent. Finally, the loading vector (14×1) of each training image was calculated. Eigenface loadings ranged from -5690.88 to 4165.27, and thus were resized and by dividing by 1000. The resized eigenface loadings are the facial features we extracted from the 15 stimulus images. These loadings were used as independent variables in our subsequent data analysis. The average face and 15 eigenfaces extracted from the 15 training images are shown in Figure 2-2.

![Average Face](image1)

**Average Face**

**15 Eigenfaces**

*Above average face and 15 eigenfaces are trained from 15 stimulus faces that were used in the present study. The contribution of each eigenface to explain variances among 15 faces decreases from left to right and from the upper row to the lower row.*

*Figure 2-2: Average Face and 15 Eigenfaces*
Because the physiognomic method was used only for explanatory purposes, due to concerns on model identification issue, we selected the 11 physiognomic features\(^1\) that have been shown in the literature to have a significant effect on people’s perceptions or behavior: face width, bizygomatic face ratio, eye height, eye width, eye area, eyebrow height, distance between the eyes, nose width, nose area, chin height, and chin area. The measurement and calculation of each physiognomic feature is shown in Figure 2-3.

\(^1\) A total of 13 physiognomic features have been shown in the literature to have a significant effect on people’s perceptions or behaviors, including the 11 physiognomic features, eye roundness, and chin width. Due to multicollinearity, we dropped the latter two features from this study. Eye roundness, calculated by dividing eye height by eye width, is highly correlated with eye height \((r=0.95)\), so we chose to use only eye height in the regression model. Chin height and bizygomatic face ratio are also strongly correlated \((r=0.79)\), so we used only bizygomatic face ratio in the regression model.
The physiognomic measurement of these facial features is adapted from Cunningham (1986) and Stirrat and Perrett (2010). Specifically, 1 = Upper face height, distance from the highest point of the upper lip to the highest point of the eyelids; 2 = Width of face, distance between outer edges of cheekbones at most prominent point; 3 = Bizygomatic face ratio = width of face divided by upper face height; 4 = Eye height, distance from upper to lower edge of visible eye within eyelids at pupil center; 5 = Eye width, inner corner to outer corner of eye; 6 = Eye area, product of eye height and eye width; 7 = Eyebrow height: pupil center to lower edge of eyebrow; 8 = Distance between eyes, distance between pupil centers; 9 = Nose length, measured from bridge at level of inner edge of upper eyelid to nose tip at level of upper edge of nostril opening; 10 = Nose width, width of nose at outer edges of nostrils at widest point; 11 = Nose area, product of nose length and nose width; 12 = Chin height, distance from lower edge of lower lip to base of chin; 13 = Chin width, distance between outer edges of chins at the level of the middle of the mouth; 14 = Chin area, product of chin height and chin length.

Figure 2-3: Physiognomic Decomposition

Two trained raters made the measurement for the purpose of ascertaining reliability. All measurements were taken independently. Each rater used the measurement tool in Adobe Photoshop to measure the distances at accuracy of 0.5 pixels. This measurement is much more accurate than in previous studies, e.g. Berry and McArthur (1985) and Cunningham (1990), which used ruler to measure the distances and achieved an accuracy of 0.5 centimeters. Inter-rater reliability for each physiognomic feature measurement was greater than 0.9, and the overall inter-rater reliability was greater than 0.99. Having demonstrated good inter-rater reliability, the measurements obtained from the two raters were averaged for each feature of each face, and these values were used in the subsequent analysis.

Experimental Procedures

A total of 445 undergraduate students at a major U.S. university participated in this study in an on-campus computer laboratory. Participants received $10 for their participation. Of these
participants, 186 were male (42%) and 259 (58%) were female. Racial groups consisted of 353 Caucasians (79%), 14 African-Americans (3%), 17 Hispanics (4%), 37 East Asians (8%), 11 other Asians (2%), and 13 others (3%). Ages ranged from 18 to 33 years old, with a mean of 19.74 and a standard deviation of 1.63 years.

We applied a between-subject design. Each participant was presented with five advertisements, one for each product category. Within each product category, one stimulus advertisement out of three was randomly assigned to each participant. The sequence of five product categories was also randomized for each participant. After participants reviewed each advertisement, they evaluated the advertisement on attitude toward the ad, attitude toward the brand and purchase intention. Upon completion, they moved to the next ad task until they finished viewing and evaluating five ads. The experiment concluded with a short demographic survey. The average time spent on each advertisement evaluation was 15.3 seconds.

Five items (bad/good, ineffective/effective, uninteresting/interesting, dislike/like, irritating/not irritating) measured attitudes toward the advertisement using a seven-point scale. Attitudes toward the brand were measured with a three-item seven-point scale (unfavorable/favorable, negative/positive, dislike/like). Purchase intention was measured using a two-item seven-point scale (unlikely/likely, improbable/probable) (Miniard et al. 1990). Reliability for the three constructs was 0.88, 0.89, and 0.91, respectively.

2.4 Estimation and Results

Effect of Faces on Key Ad Metrics and Effect Size at Aggregate Level

The present study was intended to answer three questions: a) whether different faces have an effect on advertising effectiveness; b) whether the effect is large; c) what facial features
contribute to differential advertising effectiveness, and whether such contributions are different across individuals. In order to answer the first and second questions, we performed pairwise comparisons on each key ad metric within each product category on an aggregate level. The only difference in the three stimulus ads for each product category is the face appearing on the ad; therefore, any differences in key ad metrics would be attributed to the face used in the ad. For each pairwise comparison, we compared the means and percentages of high rating (rating > 6) participants, as in business practice practitioners care most about the customers who give high ratings to their products or services. The two sets of comparisons showed reasonable convergence, as shown in Table 2-1.

**Table 2-1: Compare Key Ad metrics at Aggregate Level**

<table>
<thead>
<tr>
<th>Attitude toward Ad (Aad)</th>
<th>Mean</th>
<th>Percentage of Aad&gt;6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>face1</td>
<td>face2</td>
</tr>
<tr>
<td>Staff recruiting</td>
<td>3.66 (145)</td>
<td>3.66 (152)</td>
</tr>
<tr>
<td>Restaurant</td>
<td>4.26 (148)</td>
<td>4.51 (155)</td>
</tr>
<tr>
<td>Real Estate Agent</td>
<td>3.71 (128)</td>
<td>3.85 (172)</td>
</tr>
<tr>
<td>Fragrance</td>
<td>3.51 (143)</td>
<td>3.73 (143)</td>
</tr>
<tr>
<td>Ski resort</td>
<td>3.11 (157)</td>
<td>2.99 (130)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Attitude toward Brand (Abrand)</th>
<th>Mean</th>
<th>Percentage of Abrand&gt;6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>face1</td>
<td>face2</td>
</tr>
<tr>
<td>Staff recruiting</td>
<td>3.86 (145)</td>
<td>4.05 (152)</td>
</tr>
<tr>
<td>Restaurant</td>
<td>4.17 (148)</td>
<td>4.41 (155)</td>
</tr>
<tr>
<td>Real Estate Agent</td>
<td>3.76 (128)</td>
<td>3.87 (172)</td>
</tr>
<tr>
<td>Fragrance</td>
<td>3.72 (143)</td>
<td>3.83 (143)</td>
</tr>
<tr>
<td>Ski resort</td>
<td>3.45 (157)</td>
<td>3.35 (130)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Purchase Intention (PI)</th>
<th>Mean</th>
<th>Percentage of PI&gt;6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>face1</td>
<td>face2</td>
</tr>
<tr>
<td>Staff recruiting</td>
<td>3.60 (145)</td>
<td>3.87 (152)</td>
</tr>
<tr>
<td>Restaurant</td>
<td>3.93 (148)</td>
<td>4.08 (155)</td>
</tr>
<tr>
<td>Real Estate Agent</td>
<td>3.39 (128)</td>
<td>3.61 (172)</td>
</tr>
<tr>
<td>Fragrance</td>
<td>3.28 (143)</td>
<td>3.29 (143)</td>
</tr>
<tr>
<td>Ski resort</td>
<td>3.26 (157)</td>
<td>3.12 (130)</td>
</tr>
</tbody>
</table>
Superscript \(^a\) refers to significantly higher than one of the other categories at 0.10 level while subscript \(_a\) refers to significantly lower than one of the other categories at 0.10 level.

Superscript \(^b\) refers to significantly higher than one of the other categories at 0.05 level while subscript \(_b\) refers to significantly lower than one of the other categories at 0.05 level.

Superscripts and subscripts appear in pairs. For example, in attitude toward ad for the restaurant, if mean Aad of Face 1 is significantly lower than that of Face 2 at 0.05 level, mean Aad of Face 2 must be significantly higher than that of Face 1 at 0.05 level.

Numbers in the parentheses following mean values are sample sizes.

To answer the first question, the results show that faces do have an effect on advertising effectiveness. For example, by comparing means, Face 2 in the restaurant category significantly outperformed Face 1 and Face 3 on attitude toward the ad, and significantly outperformed Face 3 on attitude toward the brand and purchase intention. By comparing the percentages of high rating people, Face 2 in the restaurant category outperformed Face 1 and Face 3 on attitude toward brand, and significantly outperformed Face 3 on attitude toward the ad and purchase intention. It is worthwhile noting that for those categories in which a significant effect was not found in the present study, it only means that under the current empirical setting the three specific stimulus faces chosen for the product category did not elicit significant different responses among participants. However, the effect of faces on advertising effectiveness might possibly be revealed for these product categories if changes are made to the empirical setting or stimulus faces.

To answer the second question, the effect is non-trivial. For example, by merely replacing the face in the restaurant ad, advertisers can achieve nearly 10% increase on attitude toward ad, attitude toward brand and purchase intention, and double the number of high rating customers. In order to give a more direct sense of effect size, we converted the purchase intention rating to a predicted purchase measure using Urban and Hauser’s (1993) scale. Results are shown in Table 4-2. By merely replacing Face 1 with Face 2 in the staff recruiting ad, or by replacing Face 3 with Face 2 in the restaurant ad, advertisers can double their percentage of predicted purchase.
outcomes. These results indicate that different faces do have an effect on advertising
effectiveness, and the effect is substantial. Hence, advertisers should carefully select
faces/spokespersons to appear in print advertisements.

Table 2-2: Illustration of Predicted Purchase

<table>
<thead>
<tr>
<th>Predicted Purchase</th>
<th>Percentage face1</th>
<th>Percentage face2</th>
<th>Percentage face3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Staff recruiting</td>
<td>6% (8 out of 145)</td>
<td>12% (17.5 out of 152)</td>
<td>6% (8.6 out of 148)</td>
</tr>
<tr>
<td>Restaurant</td>
<td>10% (14.7 out of 148)</td>
<td>15%b (22.6 out of 155)</td>
<td>7%b (9.6 out of 142)</td>
</tr>
<tr>
<td>Real Estate Agent</td>
<td>4% (5.6 out of 128)</td>
<td>8% (13 out of 172)</td>
<td>7% (10.1 out of 145)</td>
</tr>
<tr>
<td>Fragrance</td>
<td>6% (8.7 out of 143)</td>
<td>5% (7.5 out of 143)</td>
<td>6% (9.6 out of 159)</td>
</tr>
<tr>
<td>Ski Resort</td>
<td>6% (9.7 out of 157)</td>
<td>6% (8 out of 130)</td>
<td>5% (7.4 out of 158)</td>
</tr>
</tbody>
</table>

Urban and Hauser’s (1993) scale was used to convert the purchase intention rating to
predicted purchase, i.e., \(0.9 \times (\text{rating} > 6) + 0.4 \times (\text{rating} > 5 \text{ but} \leq 6)\), using a seven-point scale.

Superscript \(^a\) refers to significantly higher than one of the other categories at 0.10 level
while subscript \(_a\) refers to significantly lower than one of the other categories at 0.10 level.

Superscript \(^b\) refers to significantly higher than one of the other categories at 0.05 level
while subscript \(_b\) refers to significantly lower than one of the other categories at 0.05 level.

Superscripts and subscripts appear in pairs.

Parentheses refer to the number of predicted purchases out of the total sample size.

Segmentation Analysis using Facial Features

We used 14 eigenface loadings as independent variables and the three key ad metrics as
dependent variables. To explore the possible heterogeneity in how people respond to facial
features, we used the finite mixture model proposed by Leisch (2004). Specifically,
Where $y$ is a dependent variable, (in our case attitude toward the ad, attitude toward the brand, or purchase intention) with conditional density $h$; $x$ is a vector of independent variables, in our case, (in our case, a vector of eigenface loadings); $\pi_k$ is the prior probability of segment $k$; $K$ is the predefined number of segments, an assumption to be imposed on the finite mixture model; $\theta_k$ is the segment specific parameter vector for the density function $f$; and

$$\varphi = (\pi_1, \ldots, \pi_K; \theta_1, \ldots, \theta_K)'$$

is the vector of all parameters.

The posterior probability that observations $(x, y)$ belong to class $j$ is given by:

$$P(j|x, y, \varphi) = \frac{\pi_j f(y|x, \theta_j)}{\sum_{k=1}^{K} \pi_k f(y|x, \theta_k)}$$

Considering that we dealt with only 15 faces, although the model including all 14 independent variables is still identified, but the model is very unstable and the convergence is bad. In this case, we propose to add a variable selection procedure to the segmentation model. Given $K$, we used a stepwise greedy search for variable selection (Bishop 1996). We used Akaike information criterion (AIC) as criteria variable selection (Andrews and Currim 2003). In this case, a variable will enter into the model if its coefficient is significant in at least one segment, while a variable will exit out of the model if its coefficient is insignificant in all segments. The heuristic is briefly described as below.

1) Predefine $K$;

2) Start running $K$-segment finite mixture model with empty set $P_0 = \{\emptyset\}$, i.e. only intercept, no $x$'s;
3) Select the next best variable \( x^+ = \arg \min_{x \in P_k} [AIC(P_k + x)] \) to enter into K-segment finite mixture model;

4) Update \( P_{k+1} = P_k + x^+; k = k + 1; \)

5) Remove the worst variable \( x^- = \arg \min_{x \in P_k} [AIC(P_k - x)] \) from K-segment finite mixture model if removing it can help decrease AIC value of the model;

6) Update \( P_{k+1} = P_k - x^-; k = k + 1; \)

7) Repeat 5) and 6) until removing any variable in the model cannot decrease the AIC value;

8) Repeat 3) to 7) until adding any variable into the model or removing any variable from the model cannot help decrease the AIC value, or all 14 variables are included in the model.

We tested for \( K=1,2,3,4 \) and 5, and then used AIC again as criteria to choose the best \( K \). The variable selection and segmentation result is shown in Table 2-3.

**Table 2-3: Result of Segmentation and Variable Selection Using Eigenface Decomposition**

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>Number of Segments</th>
<th>AIC</th>
<th>Selected Eigenface Loadings</th>
<th>Segment Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attitude toward Ad</td>
<td>1</td>
<td>6806.817</td>
<td>7,10,14</td>
<td>45</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>6690.630</td>
<td>1,3,6,12,14</td>
<td>68</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>6689.834</td>
<td>1,3,7,12,14</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>6678.068</td>
<td>1,3,7,8,11,12,14</td>
<td>46</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>6680.619</td>
<td>1,3,6,7,11,12,14</td>
<td>1</td>
</tr>
<tr>
<td>Attitude toward Brand</td>
<td>1</td>
<td>7593.837</td>
<td>13</td>
<td>45</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>7535.990</td>
<td>5,13,14</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>7519.708</td>
<td>5,6,9,13,14</td>
<td>70</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>7526.844</td>
<td>14</td>
<td>89</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>7528.463</td>
<td>14</td>
<td>11</td>
</tr>
<tr>
<td>Purchase Intention</td>
<td>1</td>
<td>8239.266</td>
<td>6,13</td>
<td>45</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>8236.912</td>
<td>5,8,9,12</td>
<td>42</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>8235.599</td>
<td>9,14</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>8236.393</td>
<td>9,10,14</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>8250.164</td>
<td>14</td>
<td>6</td>
</tr>
</tbody>
</table>
We used AIC as model selection criteria. The models displayed here are the best models that achieve the lowest AIC value for a specific number of segments. For each dependent variable, the best model with the lowest AIC value overall is highlighted bold.

The estimation for best model for each key ad metric is shown in Table 2-4.

Table 2-4: Result of Regression on Eigenface Loadings at Segment Level

<table>
<thead>
<tr>
<th>Loadings on Eigenfaces</th>
<th>Attitude toward Ad (Aad)</th>
<th>Attitude toward Brand (Abrand)</th>
<th>Purchase Intention (PI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Seg 1 (146)</td>
<td>Seg 2 (135)</td>
<td>Seg 3 (115)</td>
</tr>
<tr>
<td>L1</td>
<td>0.078&quot;</td>
<td>-0.090&quot;</td>
<td>-0.067&quot;</td>
</tr>
<tr>
<td>L2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L3</td>
<td>0.065&quot;</td>
<td>-0.043</td>
<td>0.025</td>
</tr>
<tr>
<td>L4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L7</td>
<td>0.204&quot;</td>
<td>-0.087</td>
<td>-0.046</td>
</tr>
<tr>
<td>L8</td>
<td>0.104&quot;</td>
<td>-0.175&quot;</td>
<td>0.035</td>
</tr>
<tr>
<td>L9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L11</td>
<td>0.034</td>
<td>-0.255&quot;</td>
<td>0.084</td>
</tr>
<tr>
<td>L12</td>
<td>-0.097</td>
<td>0.104</td>
<td>-0.129&quot;</td>
</tr>
<tr>
<td>L13</td>
<td>-0.067</td>
<td>-0.160&quot;</td>
<td>0.166&quot;</td>
</tr>
<tr>
<td>L14</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* refers to significance at p<0.10; ** refers to significance at p<0.05; a blank means that the particular variable was not chosen to enter the regression.

After obtaining the segmentation labels for each participant for a specific key ad metric, we ran a set of regression analyses. Each key ad metric was regressed on the 11 physiognomic features in each segment respectively. Results are shown in Table 2-5.
Table 2-5: Using Physiognomic Features to Segment Heterogeneity

<table>
<thead>
<tr>
<th>Physiognomic Features</th>
<th>Attitude toward Ad (Aad)</th>
<th>Attitude toward Brand (Abrand)</th>
<th>Purchase Intention (PI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Seg 1 (146)</td>
<td>Seg 2 (135)</td>
<td>Seg 3 (115)</td>
</tr>
<tr>
<td>Face Width</td>
<td>-0.018**</td>
<td>0.009</td>
<td>-0.010</td>
</tr>
<tr>
<td>Bizygomatic Face Ratio</td>
<td>-0.712</td>
<td>-2.445**</td>
<td>2.025**</td>
</tr>
<tr>
<td>Eye Height</td>
<td>-0.155**</td>
<td>-0.067</td>
<td>0.068</td>
</tr>
<tr>
<td>Eye Width</td>
<td>-0.140**</td>
<td>0.023</td>
<td>0.035*</td>
</tr>
<tr>
<td>Eye Area</td>
<td>0.021**</td>
<td>-0.035**</td>
<td>-0.001</td>
</tr>
<tr>
<td>Distance Between Eyes</td>
<td>0.075**</td>
<td>-0.040**</td>
<td>0.025**</td>
</tr>
<tr>
<td>Eyebrow Height</td>
<td>-0.066**</td>
<td>0.100**</td>
<td>-0.045**</td>
</tr>
<tr>
<td>Nose Width</td>
<td>-0.020*</td>
<td>-0.015</td>
<td>0.011</td>
</tr>
<tr>
<td>Nose Area</td>
<td>0.004**</td>
<td>0.006**</td>
<td>-0.002**</td>
</tr>
<tr>
<td>Chin Height</td>
<td>-0.034**</td>
<td>0.054**</td>
<td>0.009</td>
</tr>
<tr>
<td>Chin Area</td>
<td>0.000</td>
<td>0.005**</td>
<td>0.000</td>
</tr>
</tbody>
</table>

* refers to significance at p<0.10; ** refers to significance at p<0.05.
For example, for attitude toward the ad, participants in Segment 2 prefer a narrow face with small eyes, a big nose and a long chin; participants in Segment 3 prefer a round face with a small nose, but they do not care much about eye size and chin length.

Based on above results, we conclude that people make reasonably consistent inferences from facial features, and substantial heterogeneity exists in the way people make inferences.

**Effect of Faces on Key Ad Metrics and Effect Size at the Segment Level**

How do the segmentation results help advertisers understand people’s heterogeneity in making inferences from facial features? In order to address this question, we reran the pairwise comparisons at the segment level. Table 2-6 describes the pairwise comparison on the means for each product category to each key ad metric at the segment level.²

---

² We also ran pairwise comparisons with high rating participants. However, because the number of high ratings at the segment level is small, the percentage difference in magnitude is not significant. Also because the pairwise comparisons on percentages of high rating participants at the segment level show reasonable consistency with pairwise comparisons on the means, we decided not to show the pairwise comparison result on percent of high rating people in this paper.
Table 2-6: Compare Key Ad metrics at Segment Level

<table>
<thead>
<tr>
<th>Attitude toward Ad (Aad)</th>
<th>Seg 1</th>
<th>Seg 2</th>
<th>Seg 3</th>
<th>Seg 4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>face1</td>
<td>face2</td>
<td>face3</td>
<td>face1</td>
</tr>
<tr>
<td>Staff recruiting</td>
<td>(50)</td>
<td>(50)</td>
<td>(46)</td>
<td>(47)</td>
</tr>
<tr>
<td></td>
<td>4.19&lt;sup&gt;ab&lt;/sup&gt;</td>
<td>3.72&lt;sup&gt;a&lt;/sup&gt;</td>
<td>3.42&lt;sup&gt;ab&lt;/sup&gt;</td>
<td>3.37&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Restaurant</td>
<td>(57)</td>
<td>(45)</td>
<td>(44)</td>
<td>(40)</td>
</tr>
<tr>
<td></td>
<td>4.46&lt;sup&gt;b&lt;/sup&gt;</td>
<td>4.98&lt;sup&gt;ab&lt;/sup&gt;</td>
<td>4.33&lt;sup&gt;b&lt;/sup&gt;</td>
<td>3.77&lt;sup&gt;ba&lt;/sup&gt;</td>
</tr>
<tr>
<td>Real Estate Agent</td>
<td>(47)</td>
<td>(56)</td>
<td>(43)</td>
<td>(38)</td>
</tr>
<tr>
<td></td>
<td>3.58&lt;sup&gt;bb&lt;/sup&gt;</td>
<td>4.35&lt;sup&gt;b&lt;/sup&gt;</td>
<td>4.23&lt;sup&gt;b&lt;/sup&gt;</td>
<td>3.77&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Fragrance</td>
<td>(45)</td>
<td>(55)</td>
<td>(46)</td>
<td>(43)</td>
</tr>
<tr>
<td></td>
<td>3.51&lt;sup&gt;a&lt;/sup&gt;</td>
<td>3.84&lt;sup&gt;ab&lt;/sup&gt;</td>
<td>2.57&lt;sup&gt;bb&lt;/sup&gt;</td>
<td>3.40&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Skull resort</td>
<td>(52)</td>
<td>(35)</td>
<td>(59)</td>
<td>(48)</td>
</tr>
<tr>
<td></td>
<td>2.41&lt;sup&gt;b&lt;/sup&gt;</td>
<td>1.99&lt;sup&gt;bb&lt;/sup&gt;</td>
<td>3.73&lt;sup&gt;bb&lt;/sup&gt;</td>
<td>4.15&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Attitude toward Brand (Abrand)</td>
<td>face 1</td>
<td>face2</td>
<td>face3</td>
<td>face1</td>
</tr>
<tr>
<td></td>
<td>(83)</td>
<td>(93)</td>
<td>(94)</td>
<td>(51)</td>
</tr>
<tr>
<td>Staff recruiting</td>
<td>3.79&lt;sup&gt;b&lt;/sup&gt;</td>
<td>4.36&lt;sup&gt;b&lt;/sup&gt;</td>
<td>3.73&lt;sup&gt;b&lt;/sup&gt;</td>
<td>3.91&lt;sup&gt;ba&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>4.16&lt;sup&gt;b&lt;/sup&gt;</td>
<td>4.61&lt;sup&gt;bb&lt;/sup&gt;</td>
<td>4.05&lt;sup&gt;b&lt;/sup&gt;</td>
<td>3.93&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Real Estate Agent</td>
<td>(82)</td>
<td>(106)</td>
<td>(82)</td>
<td>(34)</td>
</tr>
<tr>
<td></td>
<td>3.80</td>
<td>3.84</td>
<td>3.90</td>
<td>3.68&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Fragrance</td>
<td>(90)</td>
<td>(83)</td>
<td>(97)</td>
<td>(45)</td>
</tr>
<tr>
<td></td>
<td>3.75</td>
<td>4.02</td>
<td>3.74</td>
<td>3.73&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Ski resort</td>
<td>(95)</td>
<td>(87)</td>
<td>(88)</td>
<td>(45)</td>
</tr>
<tr>
<td>Purchase Intention (PI)</td>
<td>face1</td>
<td>face2</td>
<td>face3</td>
<td>face1</td>
</tr>
<tr>
<td>------------------------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>Staff recruiting</td>
<td>3.50</td>
<td>3.50</td>
<td>3.63</td>
<td>3.66</td>
</tr>
<tr>
<td></td>
<td>(7)</td>
<td>(7)</td>
<td>(12)</td>
<td>(83)</td>
</tr>
<tr>
<td>Restaurant</td>
<td>4.75(^b)</td>
<td>4.70(^b)</td>
<td>4.50(^b)</td>
<td>3.45(^b)</td>
</tr>
<tr>
<td></td>
<td>(8)</td>
<td>(10)</td>
<td>(8)</td>
<td>(73)</td>
</tr>
<tr>
<td>Real Estate Agent</td>
<td>3.07(^b)</td>
<td>3.38(^b)</td>
<td>4.18(^{bb})</td>
<td>3.82(^b)</td>
</tr>
<tr>
<td></td>
<td>(7)</td>
<td>(8)</td>
<td>(11)</td>
<td>(72)</td>
</tr>
<tr>
<td>Fragrance</td>
<td>3.43(^b)</td>
<td>3.92(^{bb})</td>
<td>1.64(^{bb})</td>
<td>2.99(^b)</td>
</tr>
<tr>
<td></td>
<td>(7)</td>
<td>(12)</td>
<td>(7)</td>
<td>(80)</td>
</tr>
<tr>
<td>Ski resort</td>
<td>3.25(^b)</td>
<td>4.00(^b)</td>
<td>2.00(^b)</td>
<td>3.43(^b)</td>
</tr>
<tr>
<td></td>
<td>(8)</td>
<td>(4)</td>
<td>(14)</td>
<td>(79)</td>
</tr>
</tbody>
</table>

Superscript \(a\) refers to significantly higher than one of the other categories at 0.10 level while subscript \(a\) refers to significantly lower than one of the other categories at 0.10 level.

Superscript \(b\) refers to significantly higher than one of the other categories at 0.05 level while subscript \(b\) refers to significantly lower than one of the other categories at 0.05 level.

Superscripts and subscripts appear in pairs.

Parentheses refer to the number of predicted purchases out of the total sample size.
Results at the segment level reconfirm our answer to question 1 and question 2 – that faces do have an effect on advertising effectiveness, and the effect is non-trivial. To answer question 3, this study found substantial heterogeneity from segment to segment. For example, at aggregate level, the pairwise comparisons on attitude toward the ad for staff recruiting is not significant, but at the segment level, participants in Segment 1 favor Face 1 most and Face 3 least, while those in Segment 2 favor Face 3 significantly more than Face 1 and Face 2. At the aggregate level, on attitudes toward the restaurant ad, we found that participants favor Face 2 significantly more than Face 1 and Face 3, but they are indifferent to Face 1 and Face 3; at the segment level, only participants in Segment 1 behave in a similar manner to the aggregate level. Participants in Segment 2 still favor Face 2 the most and Face 3 the least. Participants in Segment 3 actually favor Face 2 least, but they are indifferent to Face 1 and 2. Participants in Segment 4 favor Face 3 least, but they are indifferent to Face 1 and 2. Hence, ad agencies should not only be careful with selecting faces for their ads, but they also should take heterogeneity among target audiences into consideration. It is ideal if they can make multiple advertisements with each having a different face to appeal to different segments of the target audience.

2.5 Study 2

Experimental Design

A critical concern on Study 1 is the somewhat ad-hoc process of selecting faces and categories. To overcome such ad-hoc aspect, we tried to replicate Study 1, but this time with a more systematic way of selecting faces and categories. For face selection, we restricted to model faces since models are reasonably good looking and thus their faces have a reasonable chance to appear on a print advertisement in real marketing practice. Additionally, in order to make the
eigenface method work, we imposed additional constraints as followed: 1) white male, with age between 25 and 35; 2) no celebrity; 3) no accessories, no obvious moustache; moderate hairstyle; 4) mild smile, frontal view; 5) reasonable resolution. We screened thousands of male model faces from fashion websites such as Macy’s, Brooks& Brothers, Myhabit.com, etc, and finally built up a face database with 60 eligible faces, which satisfy all our constraints. Then we randomly selected 12 faces from the face database and use these randomly selected 12 faces as stimulus faces.

For category selection, we applied the following constraints: 1) general public has some knowledge of the categories; 2) the categories should include both products and services; 3) the categories should differentiate on dimensions like hedonism, utilitarian and technical; 4) the categories should be appropriate to feature a male face. Finally we chose 12 categories. They are beer, restaurant, job search agent, men’s cologne, coffee, computer, hotel, jeans, SUV, sports shoes, camcorder, and car dealership. For each category, we picked a real ad as context to feature all the 12 stimulus faces. For each ad context, we tried to make as little modification as possible in order to mimic the process of real practice. We basically 1) changed the real brand name to a hypothetical one; and 2) replaced the real faces with our stimulus faces. In order to make the stimulus faces to fit well with the chosen ad context, we hired a graphic professional to replace the faces to make sure that the stimulus faces look reasonably natural in the ad contexts and it is difficult to tell that the faces are modified from a quick look at the print ads. We kept four original faces in the ads as baseline, which are restaurants, job search agent, computer, and sports shoes.
**Extraction of Facial Features**

With a similar approach as in Study 1, which has been explained in details in Section 2.3, we extracted eigenfaces and calculated the eigenface loadings of the 12 stimulus faces. We used the 60 faces in our face database as training images. And the loadings on the five most dominant eigenfaces are used as independent variables for the subsequent analysis. The average face and five most dominant eigenfaces extracted from the 60 training images are shown in Figure 2-4.

![Average Face and Five Eigenfaces](Image)

*Figure 2-4: Average Face and Five Eigenfaces*
**Experimental Procedures**

A total of 287 undergraduate students at a major U.S. university participated in this study in an on-campus computer laboratory. Participants received extra credit for their participation. The participants comprise of male (54%) and female (46%). Racial groups consist of Caucasians (75%), African-Americans (1%), Hispanics (4%), Asians (17%), multiracial (2%), and others (2%). Ages ranged from 18 to 30 years old, with a mean of 18.79 and a standard deviation of 1.21 years.

We applied a between-subject design. Each participant was presented with 12 advertisements, one for each product category. Within each product category, one out of 12 stimulus advertisements (or 13 if original face is used for the category) was randomly assigned to each participant out of 12, and each face appeared no more than once. The sequence of 12 product categories was also randomized for each participant. The participants first reviewed 12 advertisements one by one as if they saw them on a magazine in real life. After they finished viewing all the 12 ads, they were asked to evaluate each advertisement that they had just seen on attitude toward the ad, attitude toward the brand and purchase intention one after another. The experiment concluded with a short demographic survey. The average time spent on each advertisement evaluation was 9.1 seconds.

**Results**

To answer the first question about existence of effect, the results show that faces do have an effect on some ad metrics for some categories. For example, by comparing means, Face 1 in the men’s cologne category significantly outperformed Face 12 on attitude toward the ad ($t=4.09$, $df=56$, $p<0.001$), attitude toward the brand ($t=3.73$, $df=56$, $p<0.001$), and purchase intention ($t=4.59$, $df=56$, $p<0.001$). It is worthwhile noting that for sports shoes category, Face 3
significantly outperformed the original face on purchase intention (t=2.39, df=38, p<0.05). This evidence gives merits for a quantitative method of screening of faces over the current advertising practice.

To answer the second question about practical relevance of the effect, the results again show that the effect is none trivial. For example, by merely replacing Face 12 with Face 1 in the cologne ad, advertisers can increase the percentage from 10% to 62% by the percentage of high rating people, whose self reported purchase intention is greater than six on a seven point scale. By merely replacing the original face in the sports shoes ad with Face 3, advertisers can increase the percentage of high rating people from 27% to 61%. Hence, again it shows that advertisers should carefully select faces/spokespersons to appear in print advertisements.

These are the results on the aggregate level. In order to answer the third question of heterogeneity, we ran segmentation analysis to segment participants on their preferences toward facial features, which refer to eigenface loadings in this case.

We used AIC as model selection criteria, and the loadings on the five most dominant eigenfaces as independent variables. The estimation for best model for each key ad metric is shown in Table 2-7.

Table 2-7: Result of Regression on Eigenface Loadings at Segment Level

<table>
<thead>
<tr>
<th>Loadings on Eigenfaces</th>
<th>Attitude toward Ad (Aad)</th>
<th>Attitude toward Brand (Abrand)</th>
<th>Purchase Intention (PI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Seg 1 (135) Seg 2 (53) Seg 3 (99)</td>
<td>Seg 1 (60) Seg 2 (96) Seg 3 (40) Seg 3 (91)</td>
<td>Seg 1 (68) Seg 2 (153) Seg 3 (66)</td>
</tr>
<tr>
<td>L1</td>
<td>-0.002 -0.059** -0.079**</td>
<td>-0.015 -0.061** -0.098** -0.005</td>
<td>0.015 -0.077** 0.049*</td>
</tr>
<tr>
<td>L2</td>
<td>-0.056* -0.011 0.136**</td>
<td>-0.030 0.062 0.086* -0.030</td>
<td>-0.047 0.095** -0.070**</td>
</tr>
<tr>
<td>L3</td>
<td>-0.062 -0.086* 0.125**</td>
<td>-0.062* 0.033 0.242** -0.078</td>
<td>-0.151 0.113** -0.068</td>
</tr>
<tr>
<td>L4</td>
<td>-0.083 -0.066 0.330**</td>
<td>-0.086** 0.136* 0.457** -0.091</td>
<td>-0.205 0.276** -0.162**</td>
</tr>
<tr>
<td>L5</td>
<td>-0.022 -0.114* 0.120*</td>
<td>-0.142** 0.162** 0.048 -0.033</td>
<td>0.011 0.065 -0.090</td>
</tr>
</tbody>
</table>

* refers to significance at p<0.10; ** refers to significance at p<0.05.
Based on above results, we conclude that people make reasonably consistent inferences from facial features, and substantial heterogeneity exists in the way people make inferences.

After obtaining the segmentation labels for each participant for a specific key ad metric, we ran pairwise comparison on the means for each product category to each key ad metric at the segment level. To answer the third question, this study found substantial heterogeneity from segment to segment. For example, for beer ad at aggregate level, Face 6 achieves highest attitude toward the ad while Face 12 achieves the lowest. However, at segment level, for participants in Segment 1, Face 6 achieves the highest attitude toward the ad while Face 12 achieves the lowest, which is consistent with the aggregate level result; at Segment 2, Face 4 achieves the highest, while Face 10 achieves the lowest; at Segment 3, Face 5 achieves the highest and Face 8 achieves the lowest. Similar conclusion can be drawn from the results for attitude toward the brand and purchase intention.
2.6 Conclusions and Discussions

In Essay 1 we empirically demonstrated the effect of different faces on advertising effectiveness for various product categories. To answer the three questions that were raised in Section 2.1, we conclude that a) faces do affect how a viewer reacts to an ad in the metrics that advertisers care about; b) the effect size is substantial; c) people show reasonably consistent preferences toward faces, and substantial heterogeneity exists in how viewers react to advertisements. Moreover, eigenface features are practical for segmenting people based on their preferences toward different faces.

The present work contributes theoretically to the existing literature on that: a) we focused specifically on faces in advertisements and empirically demonstrated the effect of different faces on advertising effectiveness by using real ads and real faces; b) we introduced eigenface method into marketing research, which will hopefully encourage future face studies in marketing; and c) we resolved the controversy over people’s heterogeneity in face preferences and contribute to the face literature in general.

 Practically, the present work has several implications for advertisers and ad agencies: a) the substantial effect of different faces on advertising effectiveness indicates that ad agencies should be careful with selecting faces to appear in the advertisement; b) ad agencies should pay attention to possible heterogeneity in the preferences of the target audience and use different faces to target different customer segments; c) the methods used in the present study provide a new approach for professionals interested in conducting a quantitative study to assist in the screening and selection of print media spokespersons.

In general, there are five directions for future research. Firstly, in the empirical study, we only tested five product categories and used three faces for each product category. It is
worthwhile exploring the face effect in more product categories. Secondly, in the present study, due to data sufficiency issue, we assume that people’s preferences toward faces are constant among categories, i.e., product category free. For example, if a person prefers Face 1 over Face 2 in the real estate agent advertisement, he or she should still prefer Face 1 over Face 2 in the restaurant advertisement. It would be interesting to explore the possible interaction between faces and product categories in the future. Thirdly, the present paper builds a direct link from facial features to advertisement responses. Several behavioral studies have examined the relation between facial features and viewers’ inferences of trait dimensions, such as babyfacedness, trustworthiness, and attractiveness, (e.g. Berry and McArthur 1985; Cunningham 1986) while others focused on the relation between the inferences of these trait dimensions and advertising effectiveness (e.g. Ohanian 1990). Combining these two streams of research suggests that inferences of trait dimensions may serve as mediators between facial features and advertising effectiveness. Future studies might empirically test the mediation role of related trait dimensions. Fourthly, the present paper focuses on static faces with neutral expressions, but future research could go one step further by studying the effect of facial expressions. Lastly, the present study focuses on print ads. Future researchers might consider studying the effect of faces in video ads, given the dominant role of TV advertising on total media ad spending (eMarketer 2012).
Chapter 3

A quasi-Patent (qPatent) System for Service Innovation

3.1 Introduction

The global economy is increasingly dependent on services, which account for about 70% of the aggregate production and employment in the Organizations for Economic Cooperation and Development (OECD) nations (Berry et al 2006), and approximately 80% of GDP in the United States (Bitner et al 2008). Service innovation is critical for service providers to survive and succeed in the market (Kasper et al 1999; Metcalfe and Miles 2000; Andersson 2001). In the United States, services account for around 20% of total business expenditures on research and development, and the share keeps increasing yearly (Pilat 2001). Despite the importance of services, much is lacking on the service innovation front (Metcalfe and Miles 2000; Edwards and Croker 2001; Drejer 2004). Indeed, Menor et al (2002, p135) claim that service innovation remains among “the least studied and understood topics in the service management literature.” Bitner et al (2008) criticize existing new service development (NSD) methods as being *ad hoc*. Most firms, even those that are most successful in providing new services, tend to “fall back on informal and largely haphazard efforts, from brainstorming, to trial and error, to innovation teams” (Thomke 2003, p71). Perhaps because of their *ad hoc* nature, these methods lead to only occasional success in service innovation. Therefore, there is a critical need to develop a systematic and rigorous new service development system that helps service firms innovate.
In contrast, there are many well-documented systematic and rigorous innovation methodologies in the product domain, and such methodologies have also been successfully applied in practice. One of the most important such tool is the patent system (Hauser et al 2006). The patent system allows firms and individuals to extract rents from their product and technological inventions, thus creating a strong incentive to innovate, which also explains the proliferation of new product development (NPD) methodologies (Mansfield 1986; Encaoua et al 2006). It has been well documented in the economics literature that an increase in the amount of patent protection leads to unambiguous increases in the rate of innovation (Lerner 2002). Mansfield (1986) found that even for industries like office equipment, metals and textiles, where the patent system is not as important as in industries like pharmaceuticals and chemicals, about 10% of inventions would not have been developed in the absence of patents.

However, service innovations cannot be protected under the patent system. Utility patents (the most important type of patent, the other two being plant patents and design patents) have three requirements: they must be new, useful, and nonobvious. The “nonobvious” requirement means that an invention must not be obvious to someone skilled in that trade. Due to the non-technological and intangible nature of service innovations (Sundbo 1997; Edwards and Croker 2001), it is practically impossible to satisfy the nonobvious requirement in the patent system.

Huston and Sakkab (2006) proposed that companies rely more on outside inventors to enhance R&D productivity while reducing innovation costs. Now 35% of Procter and Gamble’s innovations originate outside the company, and Lego has had great success involving customers

---

3 In this paper, we make a distinction between products and services following the norm in the literature. Products are defined as solutions that are relatively tangible, are manufactured and can be stored (Nijssen et al 2006), and are used interchangeably with tangible goods (Zeithaml et al 1985), physical goods (Hauser et al 2006) or manufacturing (Drejer 2004). Services in this paper are defined as solutions that are relatively intangible, heterogeneous, inseparable of production and consumption, and perishable (Zeithaml et al 1985; Hauser et al 2006; Nijssen et al 2006). Service innovation in this paper is defined as generation of new service offerings (Leiponen 2005), excluding the new process of delivering existing services. We use it interchangeably with new service development.
in the innovation process. Chesbrough (2010) advocated that service companies use outside inventors as well. However, in the absence of intellectual property protection, potential inventors outside of firms have no incentive to develop service innovations. In addition, the resulting nondisclosure of potential inventions/ideas precludes inventors from building upon each other’s inventions, a critical driver in the development of new product innovation. These severe limitations are compounded by the fact that about 85% of service firms are small businesses (Martin 2001). These small businesses lack personnel within the firm who possess the necessary skills, expertise and experience of innovators (Sundbo and Gallouj 2000; Howells 2001), and they cannot afford to engage professional consultants on a regular basis to make up for it.

To remedy this critical handicap in service innovation, we propose and validate a new service development system, drawing inspiration from three domains: (a) the patent system (see Hauser et al 2006), (b) incentive aligned mechanism design (see Ding 2007), and (c) service innovation (see Zeithaml et al 1985). The system builds upon the principles of patenting with several major differences; hence we call it a quasi-patent (qPatent) system for service innovation. One difference, for example, is that the qPatent is a system at firm level while the existing patent system is at macro level. In addition, this system is designed to accommodate the unique characteristics of services. It provides a platform by which firms can incent outside inventors to create service innovations that address their specific needs, offering the key benefits of a typical patent system while incorporating dimensions most relevant to service innovations. In lieu of incentives embedded in a market level system such as the patent system, the qPatent system develops an incentive mechanism following recent marketing literature on mechanism design and incentive alignment (e.g., Toubia et al 2003; Prelec 2004; Ding et al 2005; Ding 2007; Toubia et al 2007). This incentive mechanism ensures that it is in the best interest of qPatent participants to come up with the most profitable service innovations, based on the needs defined by the firm that is running the particular qPatent system. We organize the rest of this paper as follows. We first
review the literature and practices motivating our work. This is followed by a detailed description of the qPatent system. Next, we describe an empirical implementation of the qPatent system, with a summary of analysis and results. Finally, we discuss general findings, limitations and future research opportunities.

3.2 Literature Review

Our work draws inspiration from three domains of knowledge, namely, the patent system, applied mechanism design (incentive aligned), and service innovation. First, we provide a general overview of the patent system, since it is the framework that we modify and extend. The most critical challenge in designing any system is to ensure that appropriate incentives are in place to motivate participants to do what the system designers want them to do. For this, we rely on recent literature about incentive aligned mechanism design. We follow the tradition in mechanism design literature, aiming to come up with a novel and creative mechanism, basically an incentive structure, that motivates participants to respond truthfully and be highly motivated, behaving in a way that maximizes the system designer’s utility when they try to maximize their own utility (Fudenberg and Tirole 1991; Mas-Colell et al 1995; Toubia 2006; Ding 2007). We would like to demonstrate that our proposed mechanism works, with no intentions to claim that our proposed mechanism is the only or best solution to the problem. This section concludes with a review of service innovation, the substantive domain of our research contribution, where we highlight the unique challenges we intend to address.
Patent System

The patent system has been playing a very important role in motivating product innovations, especially for chemicals, plastics and drugs industries (Hauser et al 2006). By protecting intellectual property, the patent system encourages inventors both inside and outside the firms to create innovations that: (a) have substantial market value, or (b) can serve as building blocks for future inventions with substantial market value, even though such inventions may not lead to valuable new products themselves. Here we provide a quick summary of how the patent system works, and its two main implications.

Patent System Overview

In order to obtain a patent in a specific country, one must file a patent application with that country’s patent office, which in the United States, is the U.S. Patent and Trademark Office (USPTO). In the United States, there are three types of patents: utility patents, plant patents, and design patents. Most patents are utility patents. The word utility means these patents have "useful" processes and products. Design patents protect original designs for articles of manufacture. Design patents cannot be primarily functional; otherwise they would be filed as utility patents. Plant patents are issued for new varieties of asexually-reproducing plants. In almost all innovation contexts, firms are concerned with utility patents. Thus, the rest of the paper only discusses utility patents in the patent system, and are referred to simply as patents.

In order to obtain a patent, an inventor must demonstrate that an invention satisfies the three criteria of new, useful (i.e., utility), and nonobvious. A patent examiner experienced in the domain of invention is assigned to evaluate an application based on these three criteria. An invention can be determined to be new by thoroughly searching the existing patent database and knowledge in the public domain. Whether an innovation is useful can also be relatively easily
judged. The last criterion, nonobvious, however, can be somewhat more subjective. Patent law defines a nonobvious invention as one that is not obvious to people who are skilled in the trade where the invention belongs to.

If a patent application is approved, it is protected for a fixed period of time. In general, a patent is valid for 20 years, starting from the filing of a patent application. The patent applicant, however, must make full disclosure of his or her inventions. Upon the expiration of the patent, others not only can use the invention without cost, but also have access to its complete blueprint from the patent application.

Patents are normally filed by two types of applicants: individuals filing on their own behalf, and firms, including employees, who typically assign all of their inventions to their employers upon employment. The timing of invention is typically determined by inventors, asynchronous with the actual needs of firms and markets.

There is substantial cost associated with filing a patent application, including both application and maintenance fees. The potential reward for a patent holder comes from two main sources. First, the holder can create products based on the invention and sell them directly to customers. Second, the holder can license the patent to other parties. Generally, any dispute on patent infringement is initiated by the patent owner and resolved in a patent court in that country.

However, the current patent system is not perfect, and it suffers some limitations. Ignorance of market needs is one critical criticism of the patent system. The current patent system has very strict requirements on novelty and nonobviousness of innovations from a technological perspective. Therefore, a large percentage of patents is driven by technological advancement, rather than market needs. Higher technological innovativeness does not necessarily result in higher marketability (Kleinschmidt and Cooper 1991; Dodgson et al 2008). Ignorance of market needs and the corresponding lack of commercial viability is a critical cause of high failure rate of patented innovations (Panne et al 2003).
Additionally, in the existing patent system, inventors are not sufficiently incented to create innovations that appeal to the market. When innovators innovate and apply for patents, there is no guarantee that their innovations will be rewarded. The value of their patents or innovations is mainly decided by the innovators, in terms of how much they charge for licensing fees, rather than the market value of the patent (i.e., how much the market really values the innovations). If the patent system can be incentive aligned with market, innovators will make more effort toward generating innovations that appeal to the market (Hauser et al 2006).

Another criticism concerns efficiency issues, mainly time and cost. The pace of innovation heavily influences the growth of output and productivity (Mowery and Rosenberg 1979). It usually takes 2 or 3 years, or even longer for a patent to be granted (Griliches 1998). Technology is progressing so rapidly that an innovation may possibly become obsolete before a patent is issued (Mansfield 1986). Also, it is very costly to obtain a patent. The cost of obtaining patent is estimated at $10,000 to $30,000 per patent (Lemley 2001), and much more if attorneys are involved. Many patentable innovations (from 20% and 40%, varying by industry and firm) are not patented due to efficiency issues.

**Implication One: Encouraging Innovations with Direct Market Value**

For knowledge-based industries such as microelectronics, biotechnology, pharmaceuticals, and telecommunications, success is mainly determined by whether firms are able to recoup investments in innovation and create additional value for stakeholders. To achieve this, others must not be able to copy their inventions freely. The patent system provides such protection for intellectual property by ensuring that nobody can create products based on their inventions for 20 years. The patent system is generally considered to be a prominent and valid policy instrument that encourages innovations.
Patent protection has substantially increased the rate of innovation for many industries. A large portion of inventions, ranging from around 10% for non-knowledge intensive industries to 60% for knowledge intensive industries, would not have been developed in the absence of patents (Mansfield 1986).

Implication Two: Encouraging Innovations That Are Valuable Building Blocks for Other Innovations

The patent system also encourages filing patents for (and thus publishing) innovations that may have no direct market value, what we call intermediate innovations. While this can be considered inefficient, it does provide substantial value to future innovation. Patent systems guarantee that patent holders for intermediate innovations are compensated for any future innovations that include them through licensing fees. This leads many innovators to develop innovations even when potential market applications are unclear, in the hope that future innovations will incorporate them and create direct market value.

There is a real benefit to this implication of the patent system. It serves as a way to aggregate knowledge and expertise, allowing innovators to learn from each other and build upon existing inventions (Scotchmer 1991; Chang 1995) in a way that is consistent with their financial interests. A successful innovation tends to require multiple domains of knowledge. The patent system makes good use of the collective talents of multiple knowledge sources while protecting the intellectual property of each source. At the end of the day, it is likely to substantially increase the number and value of patents that have direct market value.
Incentive Aligned Mechanism Design

The task of designing a useful system to encourage service innovation falls within the broad domain of applied mechanism design. In this type of work, the most critical challenge is to ensure that appropriate incentives are in place to motivate participants do what system designers want them to do. For this, we build upon the recent literature on incentive aligned mechanism design (e.g., Toubia et al 2003; Prelec 2004; Ding et al 2005; Ding 2007; Toubia et al 2007; Dong, Ding and Huber 2010).

In marketing, we have been devising methods to incent participants to perform certain tasks for a long time. The most common task is to reveal their preferences using methods such as conjoint. Normally, such methods are designed with a fixed payment to participants in the end, regardless of their performance. Such a fixed incentive structure makes the task a hypothetical one. This lack of proper incentives causes two problems. First, in hypothetical contexts, when answers are not aligned with payoffs, respondents may not experience strong incentives to expend the cognitive efforts required to provide researchers with accurate answers and/or they may provide more socially desirable answers than what they normally would do in real life. As a consequence, answers collected in hypothetical situations do not correspond to real behavior. This is called hypothetical bias, and is likely to seriously misguide managerial decisions. Second, it is extremely hard to get participants to participate in complex tasks in hypothetical situations, as they quickly lose interest and are not motivated to expend the effort required to fully participate in complex tasks.

To remedy this problem, marketing scholars have developed incentive aligned methods that bring these methods in line with applied mechanism design literature, where the goal is motivating participants to tell the truth (see Ding 2007). Incentive alignment is a set of motivating heuristics designed to ensure that the respondents believe “(1) it is in their best
interests to think hard and tell the truth; (2) it is, as much as feasible, in their best interests to do so; and (3) there is no way, that is obvious to the respondents, they can improve their welfare by ‘cheating’” (Ding et al 2011, p120).

The growing literature has documented convincing evidence that in the incentive aligned context, respondents behave in ways much closer to real life than in the hypothetical context (see for example, Toubia et al 2003; Prelec 2004; Ding et al 2005; Ding 2007; Toubia et al 2007; Lusk et al 2008; Dong, Ding and Huber 2010; Miller et al 2011). In addition, by using the incentive aligned method, researchers can create more complex research designs than with the conventional hypothetical method. In incentive aligned situations, since participants are working to maximize their own benefits, they are more willing to put forth the effort required for more complex procedures (Park et al 2008; Ding et al 2009).

In this paper, we intend to apply incentive aligned mechanism design to service innovation development. That is, we want to develop an incentive aligned system that allows us to ensure that participants in our system respond with honesty and are highly motivated to create valuable service innovations for the entity (firm) implementing the system.

**Service Innovation**

Four specific characteristics well separate services from products: intangibility, inseparability of production and consumption, heterogeneity and perishability (Zeithaml et al 1985). Bateson (1979) held that *intangibility* is the fundamental difference between services and products; services cannot be seen, felt, tasted, or touched in the same way that products can be sensed. *Inseparability of production and consumption* refers to the fact that most services are produced and consumed simultaneously, while products are first produced by manufacturers and then consumed by end users. *Heterogeneity* means that service performance varies substantially
based on who produces and who consumes. Both inseparability and heterogeneity imply that customer participation plays a critical role in the production process. Sundbo and Gallouj (2000) think that customer participation in the production process is the most basic characteristic of services. Perishability means that services cannot be stocked while products can be saved for later use.

As a result of these unique characteristics, service innovation faces the following challenges.

*Lack of Incentives for Innovation Due to Little Intellectual Property Protection*

Unlike technology-focused product innovations, most service innovations are not technological in nature (Edwards and Croker 2001). Sundbo (1997) found that 84% of service innovations are non-technological or primarily non-technological. Due to their non-technical nature and intangibility, it is very hard to protect intellectual property rights for many service innovations. As a result, there is little incentive for inventors to create service innovations as there is no mechanism for them to be rewarded for their innovations.

*Need for User Participation*

Compared to a product, a service is characterized by inseparability of production and consumption and heterogeneity. These two characteristics make it extremely important for service firms to engage their customers before, during, and after innovation (Vargo and Lusch 2004).

*Lack of Internal Personnel for Innovation*

Innovation depends highly on the skills, expertise and experience of innovators (Sundbo and Gallouj 2000; Howells 2001). However, 85% of service firms are small businesses (Martin 2001), which lack skilled creative personnel within the firms to develop new services. Because of their small size, these firms also cannot afford to engage professional consultants to create innovations for them.
3.3 qPatent System

To help address these challenges in service innovation, namely, (a) a lack of incentives for outsiders to innovate, (b) a need for user participation, and (c) a lack of internal resources for innovation, we propose a new mechanism, a quasi-patent (qPatent) system, in this section.

As discussed in the previous section, the most important element in any mechanism design is ensuring that participants will indeed behave in the way the system designers want them to behave. This is achieved by following recent literature on incentive aligned mechanism design in marketing. Specifically, (a) the system must be able to assess the real market value of participants’ ideas and motivate participants to come up with innovations with higher market value; (b) participants should be able to extract rent for their innovations, even if an innovation is only a component of a final new service offering; and (c) the system should protect against infringement and simultaneously allow innovators to build upon each other’s innovations, while rewarding all contributors to the final service innovation.

In addition to the above objectives, a system must satisfy a few additional criteria in order to be useful to a wide range of practices in the field. Specifically, a system should be: (a) valuable, so that innovations generated in the system are implementable and marketable by the firm, and appealing to target customers; (b) focused, meaning that the system gives the firm reasonable flexibility and control, allowing it to focus on the services that the firm desires and addressing the firm’s innovation goals; (c) efficient, meaning that it is fast, cost efficient and easy to implement; and (d) generalizable, so it is applicable to a wide variety of service contexts.

Our proposed qPatent system satisfies all of these criteria. The qPatent system consists of two stages: an Innovation Stage (first stage) and a Valuation Stage (second stage) (Figure 3-1).
Figure 3-1: Two Stages of the qPatent Method

The Innovation Stage is when innovations are generated. The Valuation Stage serves two purposes: (a) to evaluate and screen innovations generated in the Innovation Stage, and (b) to provide appropriate incentives to innovators in the Innovation Stage. We describe the system in detail in the rest of this section. In describing the system, we use the word “solution” to refer to
an innovation that addresses (solves) a particular customer need. Once a solution is officially submitted by an innovator, it becomes a quasi-patent (qPatent). Throughout the rest of the paper, solution and qPatent are used interchangeably. We use “service package” to refer to a package of solutions to a set of related customer needs for that service. We use “qPatent system sponsors” to refer to individuals or organizations, which in many circumstances are firms, that initiate, organize and administer the qPatent system, provide incentives to participants, and eventually benefit from the innovation outcomes.

**Innovation Stage**

This stage is designed to incent outside innovators to work diligently to develop service innovations in a domain chosen by a firm that is administering the qPatent system. It also embeds managers, users, and inventors in a structured process of innovation. We first describe the types of participants in this stage, and then provide detailed operating procedures.

**Participants**

There are four types of participants involved in the Innovation Stage. These are innovators, firm agents (FAs), arbitrators, and voice of the customer (VOCers), as shown in Table 3-1.
### Table 3-1: Types of Participants in qPatent Method

<table>
<thead>
<tr>
<th>Stage</th>
<th>Type</th>
<th>Task</th>
<th>Equivalent in Patent System</th>
<th>Equivalent in Firms</th>
<th>Implementation for Small Firms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Innovation</td>
<td>Innovators</td>
<td>1) Innovate and file qPatent</td>
<td>Patent owners</td>
<td></td>
<td>Outside innovators</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2) File infringement complaint, if any</td>
<td></td>
<td></td>
<td>Standard innovation fee plus incentive alignment</td>
</tr>
<tr>
<td></td>
<td>Firm Agents (FAs)</td>
<td>1) Select and combine innovations into service packages</td>
<td></td>
<td>Firms</td>
<td>Employees, e.g., salespeople, and/or management</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2) Estimate cost</td>
<td></td>
<td></td>
<td>No cost</td>
</tr>
<tr>
<td></td>
<td>Voice of the Customer (VOCers)</td>
<td>Evaluate service packages and provide feedback</td>
<td>Focus group participants, in-depth interviewees</td>
<td></td>
<td>Customer panel, public forum, plus incentive alignment</td>
</tr>
<tr>
<td></td>
<td>Arbitrators</td>
<td>Decide whether a patent has been infringed upon</td>
<td>Patent court</td>
<td></td>
<td>General employees</td>
</tr>
<tr>
<td>Valuation</td>
<td>Users</td>
<td>Choose most preferred service packages in an incentive aligned preference measurement task</td>
<td>Target customers</td>
<td>Customer panel, public forum</td>
<td>Standard fee for survey respondents plus incentive alignment</td>
</tr>
<tr>
<td></td>
<td>System Sponsors</td>
<td>Set up the system and bring various types of participants together</td>
<td>National government</td>
<td></td>
<td>Small firm or third party consulting firm</td>
</tr>
<tr>
<td>Overall</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>One-time setup cost</td>
</tr>
</tbody>
</table>
Innovators are individuals who innovate. They are given a specific set of customer needs within a service context, and are asked to create solutions for each need. Innovators are incented based on whether their innovations are incorporated into the final service package and how well the service package performs overall. This is analogous to being paid a licensing fee as a percentage of revenue or profit if a firm chooses to license a patent. An innovator may file a complaint related to a potential qPatent infringement by another innovator. The arbitrators (described later in this section) evaluate the complaint, and vote yes or no, in favor of or against infringement. Majority opinion determines the outcome of the appeal. If infringed upon, the infringing innovation will be voided. Complaints are private, but infringements, once determined by the arbitrators, are made public.

Firm agents (FAs) are individuals who represent a firm that is interested in selecting and combining innovations, which are then delivered to VOCers as a service package. They essentially act as licensees of various innovations, and then combine these licensed innovations into a package. Each agent may choose to construct a service package either based on a subset of needs or the entire set of needs. In order to address the cost/benefit issue of the service package, we place a constraint that each agent must create one package based on innovations developed by innovators. When an FA decides on a package, s/he must provide a cost estimate for each solution, and take that into consideration in deciding whether s/he wants to include (license) a particular solution or not. Each FA must provide a cost estimate on a specific solution, which is chosen either by oneself or other FAs. An average of the cost estimates across all FAs is assumed to be the cost of that specific solution. FAs also set prices for their service packages, with a fixed markup over the total cost. Other pricing strategies may also be used, if they wish. It is important that the FAs do not know the innovators, otherwise there might be collusion.

Arbitrators are individuals who decide whether a qPatent has been infringed upon by another qPatent after a complaint is filed by the original qPatent’s owner. Majority rule is used to
determine whether an infringement has occurred or not. To make majority rule feasible, the number of arbitrators for a market must be odd.

*Voice of the customer (VOCers)* are individuals (most likely potential customers) who evaluate all service packages and provide feedback. They can also suggest additional needs for future innovations.

*Procedures*

Procedures for implementing the Innovation Stage are illustrated in Figure 3-2. The specific process is described below.
**Figure 3-2: Innovation Stage of the qPatent Method**

*Steps in Rounds 2+ that are different from those in Round 1 are highlighted in grey.*
**Round 1**

**Step 1: Needs formation.** A list of specific customer needs is provided to innovators, from which they will generate innovations. The set of needs is the space within which a service package can create value. Needs should be specific and mutually exclusive so that the service package constructed by the solution to each need is unambiguous. This initial list of needs can be generated in two ways: (a) VOCers are first asked to make a list of needs, which are screened and summarized by the FAs and/or qPatent system sponsors and/or external designers, and then presented to the innovators; or (b) FAs and/or qPatent system sponsors and/or external designers predetermine, a priori, a list of major needs, depending on the goal of service innovation. A focus group also could be used here to identify needs.

Needs can be fixed or flexible during the Innovation Stage. If fixed, the number and specification of needs remains unchanged throughout the Innovation Stage. Otherwise, if flexible, existing needs could be dropped and new needs could be identified by participants (e.g., VOCers and innovators) and added to the list in any round. In this paper, we use fixed needs to make the process more manageable.

**Step 2: Existing solutions.** A set of existing solutions to each need if available is presented as a starting point to innovators. These existing solutions must have been invented already in the market. If an existing solution to a need is not available, we use an approximate solution as starting point to motivate the innovation process. This step prevents innovators from providing trivial solutions, and ensures that innovations are different from existing solutions. This set of existing solutions is also used as the benchmark comparison in the Valuation Stage. If an existing solution is not available and an approximate solution is used, the benchmark for the need in the Valuation Stage is “None”.

**Step 3: Solution generation.** An innovator is asked to come up with at least one solution for each need in the first round. Each innovation must be written in a way that is similar to a
claim in the regular patent system (i.e., a very specific sentence or two). Innovators are informed that the solution should be associated with a specific need. If a solution could potentially be used for multiple needs, the innovator should file this solution for all of the needs to which the solution can be potentially applied. Similar solutions generated by two or more innovators in the same round are jointly owned by these innovators.

*Step 4: Service package construction.* All solutions generated by innovators are presented to the FAs. FAs then pick a subset of solutions for each need as a consideration set and estimate the cost of implementing each solution in the consideration set. Each FA constructs a package with only one solution from the consideration set for each need. The purpose of the screening and selection by FAs is twofold: a) to ensure the implementability of selected solutions, and b) to reduce the information load for VOCers in the following step. The package they construct must be within a cost budget to address the cost/benefit issue. The FAs also determine prices for the service packages based on the price scheme used in the particular qPatent system.

*Step 5: Evaluation and feedback.* The service packages constructed by FAs are presented to VOCers with marked-up prices. VOCers are asked to evaluate each package and provide detailed feedback on what is good or bad about the package and why they like or dislike it and the specific innovations associated with it. VOCers’ feedback helps innovators and FAs gain in-depth insights on what target customers need and favor, thereby helping innovators to come up with better innovations and FAs to construct better packages. By better, we mean more appealing to target customers. Only those innovations chosen by FAs to construct service packages are viewed and evaluated by VOCers.

*Rounds 2-n*

*Step 1: Solution generation.* Innovators are provided information from the previous round, including all innovations generated by innovators, all service packages constructed by
FAAs, and evaluations and feedback from VOCers on each service package. Innovators are asked to innovate as in Round 1, but to avoid infringing other innovators’ intellectual property. They are free to generate a novel solution or build upon their own solutions. They can also build upon another innovator’s solution from previous rounds, but must give credit to the original solution owner. There is no minimum requirement on how many solutions they must generate from the second round on.

**Step 2: Infringement check.** After all innovators provide their solutions for the current round, solutions are disclosed so that they may check for potential infringement. Each innovator checks whether their solutions from previous rounds have been infringed upon by other innovators in the current round. Complaints pertaining to potential infringements are submitted to arbitrators.

**Step 3: Arbitrators’ vote.** Each arbitrator votes anonymously on each complaint. A complaint that receives favorable votes from more than half of the arbitrators is considered successful. Solutions that are deemed to infringe on intellectual property are considered void and deleted and the innovators are punished for infringing other innovators’ solutions. If an innovator receives two infringement rulings, the innovator is disqualified from the NSD process, and loses bonus payments and part of the participation fee. Innovators are punished for filing unsuccessful complaints as well. If an innovator files two unsuccessful complaints (i.e., a majority of arbitrators rule that there is no infringement), the innovator loses the right to file complaints in the future. These two conditions reduce the likelihood of both infringement and frivolous complaints.

**Step 4: Service package construction.** FAs are given information from the previous round, including service packages constructed by all FAs as well as evaluations and feedback from VOCers on each service package. Each FA constructs a new package based on the updated solution pool, including solutions generated in this round and previous rounds.

**Step 5: Evaluation and feedback.** This step is same as in Round 1.
We note here that Step 3 is flexible. It can be concurrent with Steps 4 or 5, or happen anytime after Step 2 and before Step 1 of the next round. Steps 1-5 are repeated in each round until at least one of three possible ending criteria for the Innovation Stage is met: (a) no new solutions are generated; (b) FAs construct the same service packages as in the previous round; or (c) a pre-determined maximum round \(n\) is reached. The maximum round \(n\) could be determined by FAs and/or qPatent system sponsors and/or external designers prior to the beginning of the Innovation Stage, in order to balance quantity/quality of solutions and costs associated with using the qPatent system. Upon completion of the Innovation Stage, the system moves into the Valuation Stage, described next.

**Valuation Stage**

This stage is designed to evaluate and screen innovations from the previous stage, and provide incentives to ensure that innovators in the Innovation Stage will indeed innovate to the best of their abilities. We first describe the types of participants in this stage, followed by the procedure for running it.

**Participants**

There is only one type of participant involved in the Valuation Stage: potential users of service packages in the firm’s target market. Participants in this stage are similar to VOCers in the Innovation Stage. However, they must be different individuals because participants from the Innovation Stage are not allowed to participate in the Valuation Stage to avoid collusion. Furthermore, this sample of participants is used to forecast expected revenues in the target market, so it must be large enough to adequately represent the target market.
**Procedures**

Participants in this stage perform an incentive aligned conjoint task, where the attribute/level space is the innovative solution created during the Innovation Stage. This task must be incentive aligned to ensure the participants in this stage are motivated to respond truthfully. It must also provide appropriate incentives for innovators in the Innovation Stage to motivate them, to the best of their ability, to create the most valuable innovations for firms.

Since most service packages used in such choice-based conjoint analysis were not available in the market at the time of the study, we used the rank order mechanism introduced by Dong, Ding and Huber (2010). This process is described in detail in that paper, so we will not repeat it here; however, we describe how it is implemented in our empirical study later. By using the rank order mechanism, participants in the Valuation Stage have a chance to experience the service package that is based on their inferred rank order and exists in the market. The Valuation Stage allows us to ascertain the performance of multiple service packages in the marketplace, and by doing so, to incent participants in the Innovation Stage.

We suggest two possible ways of constructing the conjoint space based on the innovations generated in the previous stage:

1. Conduct a random drawing to decide which round’s outcome will be used in the second stage, with the later rounds having a higher probability of being selected (e.g., the first round has a 10% chance of being selected, the second round 20%, the third round 30%, etc.). The increasing probability ensures each round’s solutions have the potential to be used (thus motivating innovators at each round) while recognizing that later round innovations are likely to be better. For the selected round, we summarize all the packages chosen by FAs, such that we construct a conjoint space, with each need being an attribute, and all innovations used in the packages to address that need as levels of that attribute. Or,
2. Simply select solutions to each need based on criteria specified by the firm and/or the qPatent system sponsors and/or external designers. The solutions must have been chosen in at least one round by one FA to ensure their implementability, and favored by some VOCers to ensure their market value. For example, we could select four solutions to each of five needs. Then we could construct a conjoint space of five attributes, each with four levels.

We may also supplement the above space with existing solutions to each need, thus creating the final conjoint space. At this point, the firm estimates the cost of offering each solution based on the average cost estimated by FAs during the Innovation Stage, and then constructs a relevant price attribute for each profile. Based on the final conjoint space, a choice-based conjoint can be designed, and responses collected from potential customers can be used to obtain the partworth of each solution in the conjoint space. The value (V) of a solution to a firm is calculated as the difference between the estimated aggregate user valuation of the solution minus the cost of providing it.

**Incentivizing Participants in the qPatent System**

As stated earlier, participants in the Valuation Stage are incentive aligned based on Dong, Ding and Huber (2010). Participants in the Innovation Stage, in particular the innovators, must be incented to exert effort, therefore their outputs must be linked to the performance of their innovations during the Valuation Stage. To do so, we propose the following approach to incentive align the innovators. We note here that participants in the Innovation Stage must know, as clearly as possible, the incentive aligned method used to ascertain the value of the service packages in the Valuation Stage.

We propose a fixed amount of reward, for example $1,000. This reward is divided among those whose solutions are selected, in a proportion equivalent to each solution’s V divided by the
total V for all solutions selected. For solutions that are built upon other solutions, an expert panel is used to determine the contributions of innovators involved in developing those solutions.

We can potentially provide additional incentives for innovators, such as: (a) a fixed bonus for any solution/innovation selected by an FA during any round, equivalent to a fixed license fee (upfront payment); or (b) a minimum number of innovations per round. A bonus for the number of innovations created is not viable, because it is not helpful to simply come up with something regardless of its value as a solution.

FAs do not require special incentives because they are managers of the firm that is sponsoring the qPatent system and should behave in the best interest of their firm. As a matter of fact, any incentive linking them to the outcome might potentially change their behavior, and complicate innovator perceptions of FA decisions, for example, cost estimations, solution selections, etc.

Arbitrators have a relatively objective task. We propose auditing the work of the arbitrators after the Innovation Stage, and not compensating those who are careless or exert little effort toward their work.

Finally, the VOCers are critical to the innovation process, and should be motivated to do their best in providing feedback. To ensure that they do so, we propose that each VOCer be evaluated by FAs on the usefulness of their feedback in helping innovators create better innovations and FAs improve service packages. A five-point scale is used for evaluation, from “not useful at all” to “extremely useful.” Each VOCer receives a bonus based on the usefulness of their feedback.

**Design Parameters to Be Considered**

There are several design parameters that need to be carefully considered when designing the qPatent system, which we discuss below.
Number and Ratio of Participant Types during the Innovation Stage

The ratios of innovators, FAs and VOCers in the Innovation Stage are important. For example, for 20 innovators working on five needs, if there are five FAs, the total number of selected solutions ranges from five (if all FAs choose the same solution) to 25 (if FAs choose completely different solutions). This is acceptable from an innovator perspective, as there is a reasonable chance that at least one solution from each innovator will be selected. If the Innovation Stage has four rounds, the probability increases even more. More FAs, however, means more packages for VOCers to evaluate, so this ratio must be taken into consideration as well. In terms of absolute number, it should be large enough to create learning and competition among innovators, but it should not be so large that it is hard to manage and creates a low FA/innovator ratio. One possibility is to create multiple qPatent markets in the Innovation Stage if a larger number of participants is desired.

An odd number of arbitrators is required to make majority vote rule work. Because the job of arbitrators is simple and objective, three or five arbitrators are sufficient.

Temporal Sequence and Number of Rounds

We recommend four to six rounds, but this depends on: (a) the innovation context, (b) participant abilities, and (c) the potential innovation space.

Each step in a round can be conducted either synchronously (i.e., all innovators are asked to innovate at the same time) or asynchronously (i.e., they are free to innovate throughout the day, and whoever submits a solution first is recognized as the owner of the qPatent). Each of the two has pros and cons. For example, a synchronous innovation process is more time efficient and controllable, while an asynchronous innovation process offers innovators more flexibility and less
time pressure. Which one to use depends on the context of a specific qPatent system implementation.

Cost, Price and Budget

The cost estimate for each solution is critical. Pricing strategies can vary, with the simplest being a markup strategy (cost-plus). A budget constraint might be necessary to ensure no outlandish solutions are suggested.

Format and Substance of Solutions

There are generally two approaches to this issue. One is to impose constraints on what they must specify (e.g., who, what, when, where, etc.), the other is to let them specify whatever they deem necessary to convey their innovations. The first approach is easier to manage in theory, but might induce trivial innovation (i.e., trivial modifications of existing innovations). The second approach might cause more disputes (infringements), but put less burden on the design. The arbitrators can handle this matter of infringement.

Summary and Contrast with the U.S. Patent System

The proposed qPatent system provides a platform that incents outsiders to innovate and identify solutions that address a firm’s specific needs in an environment where firm managers, inventors, and users work together to generate and improve solutions. Such a qPatent system could be especially valuable for the 85% of service firms that are small businesses (Martin 2001).

This is achieved through an incentive aligned mechanism design that is able to: (a) assess the real market value of innovators’ ideas (in the Valuation Stage) and motivate innovators to come up with innovations with higher market value (by linking values revealed in the Valuation
Stage to their payoff); (b) allow innovators to extract rent for their innovations even if those innovations are only components of the final new service offering (through the embedded licensing structure and payment); and (c) protect against infringement and simultaneously allow innovators to build upon each other’s innovations, while rewarding all contributors to the final service innovation (through the mechanism of infringement filing and arbitrators).

In addition, the proposed qPatent system also achieves the four desirable objectives, making it a valuable tool for managers by: (a) generating valuable innovations; (b) focusing innovation on specific customer needs, as decided by a firm (qPatent system sponsor); (c) efficiently generating results within days (or weeks) instead of months or even years and allowing for Internet implementation, which is not only easy, but also increases the pool of likely inventors; and (d) being generalizable, meaning it can be customized by different service firms to address very different needs by changing the system setup and design parameters.

Since the qPatent system is adapted from the existing U.S. patent system, we briefly compare the two in Table 3-2.

Table 3-2: Comparing qPatent with Existing U.S. Patent System

<table>
<thead>
<tr>
<th>Broad Feature</th>
<th>Specific Feature</th>
<th>U.S. Patent (Utility Patent)</th>
<th>qPatent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stages</td>
<td>Number of stages</td>
<td>One</td>
<td>Two</td>
</tr>
<tr>
<td>Domain of Innovation</td>
<td>Who decides which needs to innovate for Needs to be addressed</td>
<td>Innovator decides; unfocused from a firm’s perspective</td>
<td>The sponsor of a particular qPatent system decides; focused innovation Restricted to the needs specified in the system by the system designer</td>
</tr>
<tr>
<td>Patent Users</td>
<td>External to the patent system</td>
<td></td>
<td>Integral part of the system (FAs)</td>
</tr>
</tbody>
</table>
| Learning                  | From other innovators | Slow and sparse | Fast  
|--------------------------|-----------------------|-----------------|-------
| User feedback            | No                    | No              | Yes   
| Firm (potential licensee) feedback | No                    | No              | Yes   |
| Scope                    | Claims                | One or more     | One  
|                          | Stand alone           | In most cases   | Most likely not |
| Patent Approval          | Patent examiner       | None            | None  
| Process                  | Who decides the time  | Innovator       | System designer   
|                          | Synchronous/Asynchronous | Asynchronous    | Synchronous, regular interval   
|                          | Duration              | Years           | Days/Weeks        
| Cost of Filing Patent    | Substantial           | None            | None  
| Infringement             | Patent court, initiated by patent owner | Arbitrator, initiated by the owner of qPatent possibly infringed upon |
| Reward of Innovation     | Sell a product based on a patent | Yes, profit from the sale | No       
|                          | License to others     | Yes, license fee | Yes, share of reward |
|                          | Reward time           | Unknown          | Immediate       
| Licensing Market         | External to the patent system, slow to follow innovation, inefficient (reduced incentives) | Integrated into the system (Valuation Stage) |

### 3.4 An Empirical Demonstration

We now describe an empirical application of the qPatent system, which was conducted to demonstrate that: (a) it is a practical tool that can be easily implemented; and (b) it can facilitate the generation of desirable new service offerings whose market values exceed implementation costs. In this section, we describe the implementation, and in the next section we present the results.
**Service Innovation Context**

We want to use a typical service industry context to demonstrate the utility of the qPatent system. In addition, we want to raise the bar of this “test” application to demonstrate the practicality of the qPatent system. After screening many service contexts, we decided to implement the qPatent system to design a tour package for American tourists who were visiting Shanghai for the first time. International tourism is a critical service industry. According to World Tourism Organization’s (UNWTO) report (2011), international tourism generated $919 billion in export earnings, contributing to around 5% of worldwide gross domestic product (GDP). China is one of the top five tourist destinations. This tour package is positioned as additional activities a tourist can do, to be added upon obligatory must-do activities in Shanghai (e.g., visiting the Bund).

Tourism is a typical service context. By selecting a destination outside United States for American tourists, we were able to test whether the qPatent system worked when participants in the system were spread out across countries and time zones, with language and culture differences. The choice of Shanghai as the destination, however, was simply for the sake of convenience, as one of the authors is currently located there.

**Customer Needs**

To determine the appropriate set of needs for the Shanghai tour package, we conducted qualitative research with potential American tourists from a major American university, and selected five specific and non-overlapping needs. The five needs and associated existing solutions are shown in Table 3-7. The five needs were for American tourists to: (a) discover Shanghai cuisine, (b) learn how Chinese adults stay physically healthy, (c) learn what Chinese people do
for fun/relaxation, (d) learn how much Chinese people value products/brands that are made in the United States, and (e) learn how Chinese people initiate communication when they are attracted to a stranger. For each need, we identified an existing solution by examining the activities offered by a leading Chinese travel website targeting foreign customers (http://www.chinatraveldepot.com). For needs (b) and (e), no existing solutions were available, so we used “None” for the existing solutions as indicated in Table 3-7.

Participants

By design, we needed participants from both Shanghai, China and the United States. Specifically, we required participants with deep knowledge of the types of activities available in Shanghai, along with participants who represented potential American tourists to Shanghai. Our participant recruitment process is described below.

Innovation Stage

Innovators: We recruited 20 Chinese students from a major university in Shanghai to be innovators. These participants possessed a deep knowledge of Shanghai and China. Innovator participants included 12 males and 8 females, ranging in age from 21 to 25 years, with an average age of 23 years.

Firm Agents (FAs): We recruited five Chinese MBA students from the same university in Shanghai to be FAs. These participants were familiar with the tourism market in Shanghai. FA participants included four males and one female, ranging in age from 28 to 35 years with an average age of 31 years.
Arbitrators: We recruited three Chinese undergraduate students from the same university in Shanghai to be arbitrators. Participants included three females, ranging in age from 19 to 22 years with an average age of 20 years.

Voice of the Customer (VOCers): We recruited seven undergraduate students, two PhD students, and four staff members from a major American university to be VOCers. None of the participants had visited China before, and all of them expressed a strong interest in visiting Shanghai. VOCer participants included six males and seven females, ranging in age from 21 to 59 years with an average age of 28 years.

Valuation Stage

Users: We recruited 146 undergraduate students and two staff members from the same American university to be users. None of the participants had visited China before, and all of them expressed a strong interest in visiting Shanghai. Users included 95 males and 53 females, ranging in age from 18 to 56 years with an average age of 21 years.

Compensation and Incentive Alignment

Innovation Stage

Innovators: Compensation for innovators was comprised of two parts. The first part was a participation fee of CNY 200 (approximately $30)\(^4\), paid to every innovator after participating in all four rounds of the Innovation Stage. The second part was the incentive aligned compensation. Participants were informed that they would receive a proportion of CNY 1,000.

\(^4\) Please note that typical hourly pay in China is RMB 20 (approximately $3), compared with $10 in America.
(approximately $150) based on the performance of their solutions. The performance of participant solutions was based on whether an innovator’s solutions were part of the final conjoint space, the market value of solutions, and his or her contribution to final the solutions either directly or indirectly via other innovators’ solutions being built upon his or her solutions. Each innovator received incentive aligned compensation ranging from CNY 6 (approximately $1) to CNY 128 (approximately $20) based on the performance of his or her solutions.

**FAs:** Each FA received a fixed participation fee of CNY 400 (approximately $60) for completing the entire Innovation Stage.

**Arbitrators:** Each arbitrator received a fixed participation fee of CNY 100 ($15) for completing the entire Innovation Stage.

**VOCers:** Compensation for VOCers was comprised of two parts. The first part was a participation fee of $10 for each round in which they participated. VOCers were required to participate in at least two rounds to receive the participation fee so as to ensure consistency in feedback. The second part was the incentive aligned compensation. They received bonuses based on how useful FAs felt their feedback was in helping innovators generate better innovations and FAs improve service packages. Feedback was rated on a five-point scale ranging from “very useful” to “not useful at all.” VOCers received a $30 bonus if their feedback was rated as very useful, $25 for useful, $15 for average, $5 for somewhat useful, and $0 for not useful at all. Out of the 13 VOCers participating in the empirical study, one (8%) provided feedback rated as very useful; two (15%) provided feedback rated as useful; eight (62%) provided feedback rated as average; two (15%) provided feedback rated as somewhat useful; and no participants provided feedback rated as not useful at all.

**Valuation Stage**
Compensation for users during the Valuation Stage had two components. The first part was a participation fee of $10, paid upon completion of the choice-based conjoint task. The second part was the incentive aligned compensation. There was a caveat to implementing the standard rank ordered incentives (Dong, Ding and Huber 2010) which we discuss below, along with our modifications.

The typical incentive aligned conjoint method requires participants to purchase the tour package that best fits their preferences. However, it was unrealistic for us to send one of the participants to Shanghai due to budget constraints and time limitations. As a result, we incented participants by informing them that they would have a chance to experience the service package based on their inferred rank order in a city other than Shanghai and that the name of the city would be revealed at the end of experiment. Specifically, we told them that we would randomly select one individual from all participants in the Valuation Stage as the winner.

The winner’s prize was a paid tour to a large Chinatown and Chinese community in a major city in the United States, in which the participant would be able to experience all the activities in his/her top-ranked tour package (as inferred as from his/her choice responses). In addition, the winner would receive cash equaling $500 minus his/her inferred willingness-to-pay for the tour package. The total prize was worth $500. The city (New York) was revealed two weeks after all participants completed the study, and the winner was announced at the same time. We note that 85% of participants in the Valuation Stage indicated that they believed the researchers would be able to infer their preferences for identical activities in a major Chinatown and Chinese community in the United States based on their responses to the activities included in the Shanghai tour package.
System Implementation and Procedure

Both stages of the study were implemented through a web interface, with the codes (in PHP, available from the authors upon request) and data (in MySQL) stored on a remote server (Linux). With web implementation, the study could be conducted on any computer with an Internet connection, allowing the participants to work in two different locations with dynamic updating of their responses.

We restricted the application to a maximum of four rounds during the Innovation Stage. After three rounds, the innovations showed reasonable convergence, justifying our choice of four rounds. We set the budget for the tour package to be CNY 3,000 ($450), reasonable for a 2-3 day tour package and comparable with the price of actual packages currently available. We priced each FA’s service package with a 20% mark up on cost.

The Innovation Stage ran over a 4-day period, and 24 hours was allocated for each round. First, the participants in Shanghai (innovators, FAs, arbitrators) worked during their daytime based on the sequence described in Figure 2. During their evening (which is daytime in the United States, given the 12-hour time difference), the VOCers in the United States evaluated the packages from that round and provided feedback. The Valuation Stage was conducted in the United States at a later time after the conclusion of the Innovation Stage.

3.5 Results

In this section, we first provide some summary statistics on our empirical application, including statistics on raw innovations, followed by results from the Valuation Stage on whether this application indeed created solutions that were valuable to the firms (qPatent system sponsors).
**Process Statistics**

We set the maximum number of rounds in the Innovation Stage to four, and participants completed all four rounds because none of the ending criteria were met in Rounds 2 and 3. The average time that each innovator spent developing solutions for needs in each round was 1 hour and 8 minutes (specifically 1 hour and 24 minutes in the first round and 1 hour and 2 minutes in Rounds 2-4). Although 17 innovators (85%) indicated that it was not easy to innovate, 17 innovators (85%) indicated that they were engaged, and 15 (75%) indicated that they had fun in the process of developing innovations. Five innovators (25%) felt it was very helpful to view other innovators’ solutions, and five (25%) felt it was helpful. Three innovators (15%) felt it was very helpful to view the FA’s packages, and 10 (50%) felt it was helpful. Finally, seven innovators (35%) felt it was very helpful to view users’ feedback, and 10 (50%) felt it was helpful.

**Raw Innovation Statistics**

Altogether, we obtained 372 raw solutions from the innovators across the five needs and four innovation rounds, with 70 raw solutions for Need 1, 67 raw solutions for Need 2, 83 raw solutions for Need 3, 60 raw solutions for Need 4, and 92 raw solutions for Need 5. Detailed statistics are shown in Table 3-3.
<table>
<thead>
<tr>
<th>Round</th>
<th>Need 1 Shanghai Cuisine</th>
<th>Need 2 Physical Health</th>
<th>Need 3 Fun/Relaxation</th>
<th>Need 4 Made-in-USA</th>
<th>Need 5 Communication</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>32^a (0%)^b</td>
<td>26 (0%)</td>
<td>30 (0%)</td>
<td>29 (0%)</td>
<td>26 (0%)</td>
<td>143 (0%)</td>
</tr>
<tr>
<td>2</td>
<td>9 (11%)</td>
<td>13 (31%)</td>
<td>15 (7%)</td>
<td>14 (29%)</td>
<td>6 (4%)</td>
<td>77 (14%)</td>
</tr>
<tr>
<td>3</td>
<td>14 (50%)</td>
<td>16 (44%)</td>
<td>15 (20%)</td>
<td>5 (20%)</td>
<td>21 (24%)</td>
<td>71 (32%)</td>
</tr>
<tr>
<td>4</td>
<td>15 (40%)</td>
<td>12 (75%)</td>
<td>23 (30%)</td>
<td>12 (33%)</td>
<td>19 (32%)</td>
<td>81 (40%)</td>
</tr>
<tr>
<td>Total</td>
<td>70 (20%)</td>
<td>67 (30%)</td>
<td>83 (13%)</td>
<td>60 (15%)</td>
<td>92 (13%)</td>
<td>372 (18%)</td>
</tr>
</tbody>
</table>

^a Number of raw solutions for the specific need created in the specific round

^b Percentage of raw solutions that were self-reported to be built upon innovations created in previous rounds

Of these 372 raw solutions, 45 raw solutions were chosen by the FAs to construct tour packages. Respectively, nine raw solutions were chosen across the five rounds for Need 1; eight for Need 2; nine for Need 3; nine for Need 4; and 10 for Need 5. Detailed statistics for the FAs’ choices are shown in Table 3-4.

Table 3-4: Summary Statistics of FA Choices

<table>
<thead>
<tr>
<th>Round</th>
<th>Need 1 Shanghai Cuisine</th>
<th>Need 2 Physical Health</th>
<th>Need 3 Fun/Relaxation</th>
<th>Need 4 Made-in-USA</th>
<th>Need 5 Communication</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>[3, 0, 0, 0]^a</td>
<td>[3, 0, 0, 0]</td>
<td>[4, 0, 0, 0]</td>
<td>[3, 0, 0, 0]</td>
<td>[3, 0, 0, 0]</td>
</tr>
<tr>
<td>2</td>
<td>[1, 3, 0, 0]</td>
<td>[0, 3, 0, 0]</td>
<td>[4, 1, 0, 0]</td>
<td>[0, 3, 0, 0]</td>
<td>[0, 5, 0, 0]</td>
</tr>
<tr>
<td>3</td>
<td>[0, 2, 1, 0]</td>
<td>[0, 2, 1, 0]</td>
<td>[2, 0, 2, 0]</td>
<td>[0, 2, 2, 0]</td>
<td>[0, 1, 1, 0]</td>
</tr>
<tr>
<td>4</td>
<td>[1, 0, 0, 2]</td>
<td>[0, 2, 1, 1]</td>
<td>[2, 0, 0, 2]</td>
<td>[0, 1, 1, 1]</td>
<td>[0, 1, 1, 1]</td>
</tr>
</tbody>
</table>

^a The four numbers refer to how many solutions created in each round were selected and combined into packages by FAs for the specific need in the specific round. For example, [3, 0, 0, 0] for Need 1 in Round 1 means that altogether, three solutions to Need 1 were chosen by five FAs to construct packages in Round 1 and some FAs chose the same solutions; out of these three solutions, all three were created in Round 1 and no solutions created in Rounds 2, 3 or 4 were chosen.

Summary statistics for VOCer evaluations are shown in Table 3-5. A general trend of increased evaluation from round to round can be easily seen.
Table 3-5: Summary Statistics of VOCers’ Evaluations

<table>
<thead>
<tr>
<th>Round</th>
<th>FA1</th>
<th>FA2</th>
<th>FA3</th>
<th>FA4</th>
<th>FA5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5.71</td>
<td>4.86</td>
<td>5.57</td>
<td>4.14</td>
<td>5.29</td>
</tr>
<tr>
<td>2</td>
<td>5.36</td>
<td>5.18</td>
<td>5.00</td>
<td>5.64</td>
<td>5.36</td>
</tr>
<tr>
<td>3</td>
<td>5.67</td>
<td>5.25</td>
<td>5.83</td>
<td>5.17</td>
<td>6.08</td>
</tr>
<tr>
<td>4</td>
<td>6.36</td>
<td>5.36</td>
<td>5.64</td>
<td>5.91</td>
<td>5.18</td>
</tr>
</tbody>
</table>

*In each round, packages generated by FAs were evaluated by each VOCer on a seven-point scale. The average scores of VOCers’ evaluations were used as the evaluation of the specific package in the specific round.*

There were 10 infringement complaints submitted by innovators; four of them were determined to be valid, and thus were voided as solutions attributed to the infringing innovators. Three complaints were submitted in Round 2, and one was determined to be an infringement; two complaints were submitted in Round 3 and one was determined to be an infringement; five complaints were submitted in Round 4 and two were determined to be infringements. Only one innovator filed up to two unsuccessful complaints with the second unsuccessful complaint filed in the last round. Also, no innovators were determined to have infringed on other innovators’ solutions two or more times, precluding the removal of any participant from the study for this reason.

**Conjoint Task in the Valuation Stage**

To ensure that the conjoint design for the Valuation Stage was feasible and practical, we needed to identify a small set of solutions for each need. To do so, we used a multi-step approach.

First, we categorized raw solutions for each need by combining similar raw solutions into categories. The categories and the rounds in which they first appeared are shown in Table 3-6. 82% of categories appeared in Rounds 1 and 2, showing reasonable convergence and justifying the use of four rounds in our study.
<table>
<thead>
<tr>
<th>Need 1 Shanghai Cuisine</th>
<th>Need 2 Physical Health</th>
<th>Need 3 Fun/Relaxation</th>
<th>Need 4 Made-in-USA</th>
<th>Need 5 Communication</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Happy hour in the countryside¹;</td>
<td>1. Participate in street exercise with Shanghai residents¹;</td>
<td>1. Go to a theater and watch a traditional Chinese opera¹;</td>
<td>1. Work as volunteer in the Apple store¹;</td>
<td>1. Attend orientation week at a university¹;</td>
</tr>
<tr>
<td>2. Visit a local Shanghai family¹;</td>
<td>2. Visit a local fitness center¹;</td>
<td>2. Go to a tea house and learn how to make Chinese tea²;</td>
<td>2. Attend a public exhibition ⁴;</td>
<td>2. Attend a table-for-six mock-date game at a pub or club⁴;</td>
</tr>
<tr>
<td>3. Self-organized Shanghai cuisine tour¹;</td>
<td>3. Interview with a local Shanghai family ¹;</td>
<td>3. Go to a craft workshop and learn how to make Chinese knots³;</td>
<td>3. Attend an MBA class at a university and play a game with Chinese MBA students¹;</td>
<td>3. Attend a masked ball on cruise²;</td>
</tr>
<tr>
<td>4. Go to a private Shanghai kitchen ¹;</td>
<td>4. Attend a Chinese Kung Fu class¹;</td>
<td>4. Attend a Chinese calligraphy class and learn to use a brush pen to write your name⁴;</td>
<td>4. Arrange a booth at a flea market and sell products made in the USA to Chinese consumers¹;</td>
<td>4. Participate in or act as judge for a &quot;Love Story&quot; English speech contest⁴;</td>
</tr>
<tr>
<td>5. Participate in a cooking contest against Chinese people¹;</td>
<td>5. Try traditional Chinese physical therapy, such as acupuncture and massage¹;</td>
<td>5. Visit a leisure center in the local Shanghai community and learn how to play mahjong¹;</td>
<td>5. Organize a charity auction and sell products made in both China and the USA²;</td>
<td>5. Go to a studio and watch a match making TV show ²;</td>
</tr>
<tr>
<td>6. Learn about the traditional brewing process²;</td>
<td>6. Participate in some unique Chinese sports activities, such as dragon boat racing and tug-of-war²;</td>
<td>6. Go to a studio and watch a live &quot;China’s Got Talent&quot; show²;</td>
<td>6. Watch a Hollywood movie with Chinese people¹;</td>
<td>6. Attend a Chinese wedding ²;</td>
</tr>
<tr>
<td>7. Participate on a TV program about cooking²;</td>
<td>7. Attend a food therapy class⁴;</td>
<td>7. Play a traditional Chinese “Lane Game” with Chinese people¹;</td>
<td>7. Interview Chinese customers and see how they view American infant milk brands versus Chinese brands²;</td>
<td>7. Attend parent-organized blind date party¹;</td>
</tr>
<tr>
<td>8. Attend a cooking class¹;</td>
<td>8. Participate in University sports events and play against Chinese students¹;</td>
<td>8. Watch a Chinese play or a Chinese musical¹;</td>
<td>8. Visit a department store and compare prices between American and Chinese brands¹;</td>
<td>8. Offer free hugs on the street and see how Chinese people respond²;</td>
</tr>
<tr>
<td>9. Go to the famous Shanghai restaurant Lu Bo Lang, where President Clinton was served during his visit to Shanghai¹;</td>
<td>9. Participate in outdoor activities, such as cliff climbing or a bicycle tour²;</td>
<td>9. Stay with a Chinese family during Chinese New Year⁴;</td>
<td>9. Visit international students at Fudan University and ask their opinions²;</td>
<td>9. Visit international students at Fudan University and ask their opinions²;</td>
</tr>
<tr>
<td>10. Try Chinese fast food¹;</td>
<td>10. Attend a health lecture.⁹</td>
<td>10. Go to parks, bars, or KTV. ¹</td>
<td>10. Attend a communication lecture.³</td>
<td>10. Attend a communication lecture.³</td>
</tr>
<tr>
<td>11. Learn to cook famous Shanghai cuisine, like fried dumplings.¹</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

¹,²,³,⁴ Superscripts refer to the round in which the innovation first appeared.
For each category, we identified the solution that was most favored by the FAs and VOCers, and treated them as the best solution within the category. We then screened the best solutions based on the following criteria: (a) the solution was feasible to be offered throughout the year; (b) the solution was realistic as an add-on activity that a real tour company could incorporate into their existing packages; (c) the solution was not obvious, so that it was not a trivial solution; (d) the solution was valuable (not everybody necessarily liked it, but there is a segment of users that found it valuable); and (e) the solution was new (did not exist in typical packages). This screening process is consistent with typical new product/service innovation process (e.g., Urban and Hauser 1983).

Based on the five criteria, we finally chose three solutions for Need 1, three solutions for Need 2, three solutions for Need 3, three solutions for Need 4, and two solutions for Need 5. Together with one existing solution for each need (if available), these solutions defined our conjoint space (see Table 3-7). For each solution included in the conjoint study, we used a title, a picture and a detailed description to help users understand the activity. In addition to these five attributes (representing the five needs, with the solutions representing the levels for the attributes), we added the attribute of price with four levels ($175, $225, $275, and $325). These levels of price were consistent with the actual price ranges of such tour packages offered in Shanghai. The final conjoint space had four levels for Needs 1, 2, 3 and 4, three levels for Need 5, and four price levels.

We used SAS experimental design macros to determine the number and actual profiles of the various Shanghai tour packages for the conjoint study. Given the number of attributes (five needs plus price) and their corresponding levels, a 48-profile design was deemed to be 100% D-efficient. We therefore generated 48 different profiles and divided them into 16 sets with three profiles in each choice set.
Each participant in the Valuation Stage was provided with 16 choice sets in random order. Each choice set had three different Shanghai Tour Packages plus a non-purchase option. Each Shanghai Tour Package included the price and five activities, with each activity addressing a need. For each choice set, each participant was required to choose the Shanghai Tour Package that s/he was most likely to buy for a real Shanghai tour, or choose the non-purchase option if s/he was unlikely to buy any of the three packages. The average time that each participant spent on each choice set was 32 seconds. After they finished the 16 choice sets, they completed an immediate holdout task, which was a choice task with 10 Shanghai tour packages plus a non-purchase option. One week later, they completed a delayed holdout task, which was a choice task with another 10 Shanghai tour packages plus a non-purchase option.

Value of Innovation

In addition to motivating the innovators to work diligently to create solutions for the stated needs, the conjoint analysis in the Valuation Stage can also reveal whether the created solutions indeed had substantial market value for the firms who were running the qPatent system.

We assessed each individual participant’s preferences and willingness-to-pay for each solution by using the normal component mixture model proposed by Allenby et al (1998). Specifically, the probability that the $i$th user chooses the $k$th alternative from the $j$th choice set is given by:

$$p_{ij}^k = \frac{\exp (\beta_i^T x_{ij}^k)}{\sum_i \exp (\beta_i^T x_{ij}^k)} ,$$

where $x_{ij}^k$ describes the $k$th alternative evaluated by the $i$th subject from the $j$th choice set, and $\beta_i$ is a vector of partworths for the $i$th subject. Heterogeneity across participants is modeled with a normal component mixture model given by:
\[ \beta_i = \sum_k \phi_k \text{Normal}(\bar{\beta}_k, \Lambda_k), \]

where \( k \) indicates the number of segments, and \( \phi_k \) is the mass of each segment. Each segment is modeled with a different mean \( \bar{\beta}_k \) and covariance matrix \( \Lambda_k \) (see Allenby et al 1998 for more details about the model).

This specification allows for estimation of individual-level partworths \( \beta_i \), the aggregate or average partworths \( \bar{\beta}_k \), as well as the amount of heterogeneity for each partworth via \( \Lambda_k \). We tested a range of prior values to ensure that the reported results were invariant to the degree of noninformativeness of the specification of the prior. In addition, we assessed the convergence properties of the Markov Chain Monte Carlo analysis (using multiple chains from overdispersed starting values) (Gelman and Rubin 1992) to ensure that the algorithm converged to the target density, as induced by the model specification, before we made marginal summaries of the posterior density.

For each need, the existing solution (or the “None” option) was used as the baseline. Imposing the one-segment assumption on our estimation model, the aggregate willingness-to-pay for each activity compared to the baseline is shown in Table 3-7. The cost of offering each solution is based on the average cost estimated by FAs in the Innovation Stage.

---

5 We estimated the model with multiple segments as well. Likely due to the small sample size (148 users), a one-segment model continued to fit the data better than the two-segment model based on Allenby et al’s (1998) criteria. Log Marginal Density (LMD) for the one-segment model was -1,016.8, which is greater than the LMD of -1,030.9 for the two-segment model. Also, the classification in the two-segment model was quite unbalanced. Only 14 people were classified in Segment 2.
## Table 3-7: One-Segment Model Estimate from the Valuation Stage

<table>
<thead>
<tr>
<th>Solutions</th>
<th>Aggregate Level Estimate</th>
<th>Individual Level Estimate</th>
<th>Cost ($) a</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Posterior Mean</td>
<td>Posterior s.d.</td>
<td>Heterogeneity</td>
</tr>
<tr>
<td>Need 1: Discover Shanghai cuisine</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shanghai style restaurant (baseline)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Happy hour in the countryside</td>
<td>-0.959*</td>
<td>0.194</td>
<td>0.79</td>
</tr>
<tr>
<td>Local Shanghai family</td>
<td>-0.481*</td>
<td>0.170</td>
<td>0.88</td>
</tr>
<tr>
<td>Self-organized cuisine tour</td>
<td>-0.780*</td>
<td>0.227</td>
<td>0.75</td>
</tr>
<tr>
<td>Need 2: Learn how Chinese adults stay physically healthy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None (baseline)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Street exercise</td>
<td>1.074*</td>
<td>0.167</td>
<td>0.89</td>
</tr>
<tr>
<td>Kung Fu class</td>
<td>1.108*</td>
<td>0.214</td>
<td>0.82</td>
</tr>
<tr>
<td>Food therapy class</td>
<td>1.269*</td>
<td>0.195</td>
<td>0.99</td>
</tr>
<tr>
<td>Need 3: Learn what Chinese people do for fun/relaxation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Busiest district (baseline)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Craft workshop</td>
<td>-1.184*</td>
<td>0.192</td>
<td>0.86</td>
</tr>
<tr>
<td>Chinese calligraphy</td>
<td>-0.965*</td>
<td>0.212</td>
<td>0.97</td>
</tr>
<tr>
<td>Mahjong</td>
<td>-1.363*</td>
<td>0.190</td>
<td>0.92</td>
</tr>
<tr>
<td>Need 4: Learn how much Chinese people value Made-in-USA products/brands</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None (baseline)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sell in flea market</td>
<td>0.603*</td>
<td>0.224</td>
<td>0.92</td>
</tr>
<tr>
<td>Charity auction</td>
<td>0.876*</td>
<td>0.187</td>
<td>1.09</td>
</tr>
<tr>
<td>Public exhibitions</td>
<td>0.744*</td>
<td>0.205</td>
<td>1.11</td>
</tr>
<tr>
<td>Need 5: Learn how Chinese people initiate communication when they are attracted to a stranger</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Go to a bar (baseline)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blind-date party</td>
<td>-1.957*</td>
<td>0.210</td>
<td>0.87</td>
</tr>
<tr>
<td>Attend a Chinese wedding</td>
<td>-0.910*</td>
<td>0.203</td>
<td>0.83</td>
</tr>
<tr>
<td>Price</td>
<td>-0.248*</td>
<td>0.037</td>
<td>2.67</td>
</tr>
</tbody>
</table>

* Posterior mean is significant at $|z|>2$, where $z$ value=posterior mean/posterior s.d.  
* Cost of offering each solution is based on the average cost estimated by FAs in the Innovation Stage.  
Log Marginal Density (LMD) of one segment model = log-likelihood-parameters/2ln(observations)= -1,016.8 (Allenby et al 1998)
We now examine the value of the solutions developed using the qPatent system. We consider a solution to be “valuable” if customers’ willingness-to-pay (WTP) for the solution is greater than the cost of offering the solution. It can be easily seen that our proposed qPatent system is capable of generating valuable solutions. For example, with Needs 2 and 4, existing solutions are “None” with zero cost and zero WTP. Therefore, the marginal WTP of the new solution relative to the existing solution is simply the WTP of the new solution. Similarly, the marginal cost of the new solution relative to the existing solution is simply the cost of the solution. For Needs 2 and 4, the WTP of each new solution is greater than the cost of offering it, indicating that these solutions are indeed valuable.

For the other needs for which there were existing solutions, at the aggregate level, no new solution was found to have positive marginal WTP, which means that no new solution outperformed existing solutions at the aggregate level\(^6\). We note that we set the bar high by choosing the baseline comparison to be existing solutions that have been used and favored by the market for a long time. We also note that there is substantial preference heterogeneity at the individual level. For each new solution, there exists a group of participants who have positive WTP (even though the aggregate level mean is not positive), which means that they favor the new solution over the corresponding existing solution. Service is experiential in nature and thus people’s heterogeneity plays an important role in how a service is evaluated (Menor et al 2002; Hauser et al 2006), implying that it is unlikely that a one-size-fits-all offering would be created.

To further explore whether at least some users indeed valued some new solutions more than existing solutions, we conducted the following analysis. For each of the eight new solutions (three solutions for Need 1, three for Need 3, and two for Need 5), we re-estimated the one-segment model for users whose individual WTP estimates were positive for the specific new solutions.

---

\(^6\) In the two-segment model, however, the aggregate level estimates for the “self-organized cuisine tour” and the “attend a Chinese wedding” were significantly positive for segment 2, which means users in Segment 2 favored these two solutions more than the corresponding existing solutions.
solution in the previous analysis (when all users are included in the analysis). We found that three of these eight new solutions significantly outperformed existing solutions in their specific market segments. For the solution “visit a local Shanghai family,” the participants with positive marginal WTP were mainly female, aged 21 years or older, and willing to spend less than $1,000 on an overseas trip, excluding airfare and hotel. For the solution “self-organized Shanghai cuisine tour,” the participants with positive marginal WTP were mainly aged 20 years or younger and were highly willing to take a trip to Shanghai. For the solution “attend a Chinese wedding,” the participants with positive marginal WTP were mainly female, aged 21 years or older and were highly willing to take a trip to Shanghai.

Table 3-8: Subsample Analysis

<table>
<thead>
<tr>
<th>Model</th>
<th>Associated Attribute</th>
<th>Sample Size</th>
<th>Posterior Mean</th>
<th>Posterior s.d.</th>
<th>Posterior Mean of Price</th>
<th>Posterior s.d. of Price</th>
<th>Marginal WTP</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Need 1: Discover Shanghai cuisine</strong>&lt;br&gt;Existing solutions: Go to a Shanghai style restaurant (baseline)</td>
<td>Happy hour in the countryside</td>
<td>23</td>
<td>2.210*</td>
<td>0.496</td>
<td>-0.287</td>
<td>0.194</td>
<td>77</td>
</tr>
<tr>
<td></td>
<td>Visit a local Shanghai family</td>
<td>43</td>
<td>1.390*</td>
<td>0.261</td>
<td>-0.250*</td>
<td>0.107</td>
<td>56*</td>
</tr>
<tr>
<td></td>
<td>Self-organized cuisine tour</td>
<td>38</td>
<td>1.728*</td>
<td>0.404</td>
<td>-0.245*</td>
<td>0.120</td>
<td>70*</td>
</tr>
<tr>
<td><strong>Need 3: Learn what Chinese people do for fun/relaxation</strong>&lt;br&gt;Existing solutions: Busiest district (baseline)</td>
<td>Craft workshop</td>
<td>25</td>
<td>2.290*</td>
<td>0.468</td>
<td>-0.210</td>
<td>0.181</td>
<td>109</td>
</tr>
<tr>
<td></td>
<td>Chinese calligraphy class</td>
<td>28</td>
<td>2.233*</td>
<td>0.445</td>
<td>-0.197</td>
<td>0.161</td>
<td>113</td>
</tr>
<tr>
<td></td>
<td>Learn to play Mahjong</td>
<td>16</td>
<td>2.880*</td>
<td>0.663</td>
<td>-0.278</td>
<td>0.276</td>
<td>104</td>
</tr>
<tr>
<td><strong>Need 5: Learn how Chinese people initiate communication when they are attracted to a stranger</strong>&lt;br&gt;Existing solutions: Go to a bar (baseline)</td>
<td>Blind date party</td>
<td>17</td>
<td>1.877*</td>
<td>0.605</td>
<td>-0.227</td>
<td>0.258</td>
<td>83</td>
</tr>
<tr>
<td></td>
<td>Attend a Chinese wedding</td>
<td>44</td>
<td>1.873*</td>
<td>0.302</td>
<td>-0.225*</td>
<td>0.106</td>
<td>83*</td>
</tr>
</tbody>
</table>

* Posterior mean is significant at $|z| > 2$, where $z$ value = posterior mean/posterior s.d.

* In each model, samples are selected based on the criterion that the individual level estimate for the specific attribute is positive in one component model.

* Only the posterior mean and posterior s.d. of the specific attribute that we selected from the sample are shown here.
Given the results above, we can conclude that the qPatent system can indeed create valuable new solutions for service firms. The solutions generated from the qPatent system tend to create value for service firms when no existing solutions exist. If solutions already exist for a given customer need, it is more difficult for the qPatent system to generate new solutions that will dominate the existing solutions. However, the new solutions may be significantly more appealing to certain customer segments than existing solutions.

3.6 Conclusions and Discussions

We have proposed a new system for service innovation, the qPatent system, to assist in developing valuable new service offerings. Most relevant to the 85% of service firms that are small businesses without the resources to innovate by themselves, the system incent outsiders to innovate and identify solutions for a firm’s specific needs in an environment where firm managers, inventors, and users work together. The incentive mechanism used in the system ensures that participants are rewarded for the market value of their inventions, and facilitates learning and stimulation of creative ideas by implementing a licensing structure and infringement arbitration procedures. In addition, the proposed qPatent system has four desirable features: valuable output, focused innovation, efficient process, and generalizable design.

In an empirical study designed to demonstrate the system and its performance, we found that the qPatent system is indeed capable of generating valuable service innovations with market values much higher than the costs of offering the services. In addition, the system is efficient and easy to implement. The empirical results provide evidence that practitioners should consider using the qPatent system for service innovation.
Like any new mechanism, there are many fruitful research directions that will further enrich and improve the qPatent system proposed and tested here. We discuss five such directions below.

**ROI Considerations**

In our experimental design, we did not explicitly compare the proposed qPatent system with existing service innovation mechanisms such as brainstorming (Thomke 2003). However, our empirical results have addressed the benefits and costs issue to some extent, considering brainstorming as benchmark. For benefits, there is ample evidence showing that the qPatent system outperforms brainstorming. First, Round 1 of qPatent is somewhat similar to brainstorming in real practice and thus can serve as benchmark. We did find an increasing pattern of user evaluation on packages (see Table 5), which indicates the final outcome from Rounds1-4 beats the outcome of Round 1 (benchmark). Second, a large portion of Innovators indicated that feedback from FAs and VOCers was useful for them to create innovations in the Innovation Stage, while such helpful feedback is absent in the benchmark mechanism. Third, assuming the market is efficient, existing solutions can represent the optimized solution generated by conventional methods such as brainstorming and creativity enhancement methods. The fact that innovations generated from our proposed qPatent system beat existing solutions for some segments of people indicated that qPatent system beats the benchmark.

For costs, compared with the benchmark, the qPatent system does not require many additional costs (see Table 1). Additional costs stem mainly from two factors: system setup and incentive alignment for various types of participants. The system setup cost is a one-time cost, which will diminish by running the system again and again. Especially when the qPatent system is run multiple times by a third-party consulting firm, setup cost associated with each use is
trivial. Also, the qPatent system is quite flexible, and thus the cost of adapting the system setup to a particular context is trivial. The incentive alignment cost is associated with market performance of final innovation outcomes, representing a small portion of benefits that firms will reap from the market. The benefit of offering incentive alignment highly exceeds the cost.

**Alternative Incentive Mechanism**

Some unique contexts where qPatent might be used may require modifications to standard incentive alignment mechanisms (such as Ding 2007; Dong, Ding and Huber 2010). In our empirical study on developing a Shanghai tour package for American tourists, we were required to make some changes to the conventional incentive aligned conjoint analysis because it was not realistic to send an American participant to Shanghai. Specifically, we examined preferences toward Shanghai tour packages, but rewarded participants with the opportunity to win a tour package to a large Chinatown in a major city in the United States, with activities similar to the one they chose in the Shanghai tour package.

In addition to the 16 choice sets, each participant in the Valuation Stage was required to complete an immediate holdout task and another delayed holdout task a week later. There were 10 profiles plus a non-purchase option in each holdout task. In the Valuation Stage, 145 users completed the holdout task. The hit rate was 34.5% for the immediate holdout task, and 31.7% for the delayed holdout task using the alternative incentive alignment. Both are significantly higher than the baseline 10% (1/11) hit rate. The hit rates for this modified incentive alignment are comparable to the performance of conventional incentive alignment in conjoint analysis (Dong, Ding and Huber 2010).
**Design Variations**

The qPatent system is flexible, and we encourage users to adopt the set of design parameters that is best suited to their purposes. Earlier, we discussed tradeoffs associated with parameter values, such as the ratio of various participants in the Innovation Stage. Here, we highlight the importance of selecting various designs by discussing the choice between fixed and flexible needs in the system.

In the empirical study, we used a fixed number of needs (determined a priori) to make the study more manageable. It will be worth investigating how flexible needs might be incorporated in the qPatent system. For mature products, such as the tour package used in this study, target customers may have extensive experience with the service category and may already know what they need or want. Thus, it would be easy for firms to identify critical customer needs and innovate for these needs based on focus groups or face-to-face interviews. However, for a novel service category in which customers do not have much experience, customers may not clearly know what they need or want. Conducting focus groups or face-to-face interviews may not be as useful in determining the set of needs. In such categories, new needs might emerge as innovation progresses, and thus such categories might require the system to be flexible in order to accommodate such needs and allow inventors to innovate for them. However, a screening mechanism must be in place so no trivial needs are added to the system.

**Special Cases for qPatent Systems**

There are special cases for qPatent systems that have been used in practice. We briefly discuss two such systems here and contrast them with qPatent. One is the open-source
software development system, and the other is the so-called idea outsourcing method, or “call for solution” approach.

Similar to the qPatent system, the essence of the open source software development system is to expose innovations to the public and allow innovators to build upon each others’ innovations (Von Hippel 2001) while incorporating user feedback into the innovation process. Unlike the qPatent system, it has no built-in assessment of the market value of the innovations, and tangible rewards for innovators are not part of the system. As a result, the best inventors may not be motivated to participate in such a system. In addition, since firms do not play a role in the innovation process, it precludes their ability to determine which innovation best fits firm strategy.

In the idea outsourcing system, firms specify needs and rewards for the winner, and then invite outside innovators to create solutions to address the need. A solution that resolves the need best is selected as the winner and its innovator is rewarded. Unlike the patent system, it has no sophisticated incentives to encourage learning among innovators by protecting intermediate innovations. The winner takes all reward system also discourages some inventors from participating due to perceptions about the likelihood of winning. It also has no interaction mechanism for firms, users, and inventors.

Neither of these two mechanisms can be readily applied for our empirical context (i.e., developing a tour package for American tourists visiting Shanghai). Neither would ensure the implementability and market appeal of final solutions. Furthermore, the open source software development mechanism suffers from time efficiency issues, while idea outsourcing discourages cooperation and collaboration among innovators.

These mechanisms are simplifications (and thus special cases) of the qPatent system. Given the success of these two mechanisms in their respective application domains, it maybe meaningful to explore whether other types of simplifications or trading off some rigor for other benefits (e.g., ease of implementation) are appropriate for other types of contexts. The fact that
the open source software development system and idea sourcing system work very well (despite the fact that they are special cases of qPatent system) also suggests that the qPatent system works even better.

**Extending to Product Innovation**

Although services and products differ on many dimensions, the theme of innovation is similar (Nijssen et al 2006). With some modification, the qPatent system may be used by firms to develop new product offerings. Compared to the patent system that most product innovations have relied upon, a system like qPatent offers several additional benefits: (a) it is a focused system and participants only innovate for the needs specified by the firm; (b) it provides a platform where all stakeholders (inventors, users, firms) interact and improve the final outcome; and (c) it encourages innovations that would not be patentable, and yet are still valuable to a firm. Unlike a typical “call for solution” approach, it also allows inventors to build upon each other’s ideas. Although the modifications are not trivial, we believe a promising direction for future research will be adapting the qPatent system for product innovation as a supplement to the standard patent system and “call for solution” approaches.

In summary, we propose and validate a new service innovation system that is capable of generating valuable new service offerings. We hope the qPatent system will become a powerful addition to the service firm toolbox, especially for the 85% of service firms that are small businesses.
References


VITA

LI XIAO

418B Business Building, Pennsylvania State University, University Park, PA 16802,
Tel: (814) 321-7516, Fax: (814) 865-3015, Email: lixiao@psu.edu

EDUCATION

2008 – June 2013 (expected) PhD Candidate in Marketing,
Smeal College of Business, The Pennsylvania State University,
Chair: Min Ding
Status: Dissertation Proposal Defended (June 2012)

PhD Minor in Computer Science and Engineering,
The Pennsylvania State University

PhD Minor in Statistics,
The Pennsylvania State University

2001 – 2004 Masters of Business Administration, Marketing,
School of Business, Wuhan University, China

1997 – 2001 Bachelor of Engineering, Computer Science and Engineering,
Naval University of Engineering, China

RESEARCH INTERESTS

- Video Data and Applications in Marketing
- Machine Learning and Applications in Marketing
- Innovation, Retailing, Advertising, International Marketing