Identification of observed and unobserved taste: An application to college-games tickets using contingent valuation

Francisco Rosas †
Universidad ORT Uruguay
& Centro de Investigaciones Económicas (CINVE)
frosas@ort.edu.uy

Santiago Acerenza
Universidad ORT Uruguay
acerenza_s@ort.edu.uy

Peter F. Orazem
Department of Economics
Iowa State University
pfo@iastate.edu

Abstract

Taste of individuals drives their preferences for goods and services. Elicitation of preferences can be performed thorough several approaches which usually focus on identifying observed taste. In this paper, we use contingent valuation of revealed preferences to identify both observed and unobserved taste, and apply this approach to college sports tickets. We use data from a survey performed in Iowa State University and estimate an eight-variate Probit model. The correlation coefficients between sports identify individual's unobserved taste for sports, and show that all of them are complements. Estimation of unconditional willingness to pay show that some sports have room to increase current prices, while others, prices may be reduced if attendance is the objective pursued. Conditional (on purchasing another sport) willingness to pay is aligned with the sign of unobserved taste, such that individuals who purchased one sport are on average willing to pay more for another sport than those who did not. Results can be used by sport's authorities that, by taking into account individual's simultaneous preferences for all sports, generate cross-marketing strategies to increase attendance and revenues. The same framework can be applied to capture observed and unobserved taste for other goods and services.

JEL Codes: C35, D12, D40, Z20
Key words: Sport economics, willingness to pay, multi-variate probit analysis, contingent valuation, unobserved taste

(†) Corresponding author

1. INTRODUCTION

National Collegiate Athletic Association (NCAA) sports is recognized by the popularity of their sports, not only within the college community but also this popularity transcends the college boundaries. This popularity is reflected on several ways such as the hostage of large sport events, media coverage, attendance, advertisements, and donations, among others. As a consequence of that, Universities in general are eager to support Athletics Departments economically, in order to translate this popularity into other dimensions valued by the University such as community recognition, prestige, recruitment, and of course, more revenues. The downside is that, on average, only a couple of the sports generate sufficient revenues to cover their own costs, implying a considerable burden to be bared by University administrators. In an era where the community is aware of budget issues, Universities are facing a problem that is also connected to tuition increases and Federal and State funding, college administrators place a high pressure on expenditures and revenue generation. For example, in 2014 according to ED (2016), Division I-A Universities' spending on sports amounted to $7785 million, and generated revenues for $8220 million.[1] However, the reality is very heterogeneous if we analyse by sport, or if we divide between men's and women's sports, and also if we look at the situation within each University of the division.

Focusing on some sports, namely men's and women's basketball, football, volleyball, wrestling, gymnastics, hockey, and soccer, we can observe the heterogeneity of revenue generation. Within the sports with positive results, men's basketball in 2014 have revenues of $1001 million and total expenses of $703 million, leaving this way a surplus of $298 million; or football whose revenues are $3734 million and expenses are $2127 million, showing an

[1] This includes men's, women's, and coed sports.

interesting surplus of $1607. On the other hand, the subset of sports which do not break even, we have women basketball, volleyball, wrestling, gymnastics, hockey and soccer with significant deficits of $155 million, $66 million, $18 million, $38 million, $15 million and $83 million, respectively. This shows how important for college administrators is the need of maximizing revenues in order to improve financial performance.

One important factor of revenue generation is attendance to college sport games; in fact, attendance performance shows a similar pattern as the overall revenues generated by each sport. Men´s basketball and football (the only profitable sports) have turnouts of 28 and 44 million spectators in 2014 respectively, while women´s basketball have only 8 million attendees during that year (NCAA, 2016).[2] The previous situation opens the possibility of augmenting revenues increasing attendance or increasing prices of sports tickets, and exploiting the possibility that sports fans may show an inelastic demand. In order to know if there is room to increase attendance or prices, individual's willingness to pay for college sports events (games) should be studied in depth.

Willingness to pay for any particular good or service (sports events in this particular case) is determined by the preferences of individuals, which help us to elicit their tastes for sports.

In this study, we compute the willingness to pay for the 8 mentioned sports using data from an experimental setup carried out in Iowa State University (ISU) in 2006, which is suitable for analyzing employing a contingent valuation approach. Moreover, we apply this approach in a way that allows us to identify both individual's observed and unobserved taste for each sport.

ISU is a college which gives a lot of importance to sports. As the dataset collected show, 43% where fans of the college team before attending ISU, 65% play sports in the university and

[2] We do not have access to systematized data on attendance for the rest of the mentioned sports.

68% played varsity sports in high school. Although ISU athletics budgetary picture shows a surplus for 2014 of $7 million, it follows the NCAA Division I-A pattern. While men´s basketball and football show in 2014, respectively, a surplus of $4 million and $18 million, women´s basketball and the rest of the sports showed deficits of $3 million and $12 million, respectively.

We demonstrate that willingness to pay for a given sports can be measured using decisions to accept or reject randomly priced men’s and women’s basketball tickets in a simulated market. The parameters allow us to predict the probability of ticket purchase at any given price, and so we can estimate the reservation prices at which each individual is indifferent between buying a ticket for each of the eight sports. We also compute own- and cross-price elasticities between sports, highlighting the approach used to compute the latter.

2. LITERATURE REVIEW

Reservation prices can be obtained from different types of data. Following Jedidi & Jagpal (2009) the most common ways of getting reservation prices are: a) Self-stated willingness to pay, b) contingent valuation, c) conjoint analysis, d) experimental auctions, and e) information from actual purchase data.

To start with the last one, the main idea of this method is to analyze household survey data. The advantage of this method is that information comes from real purchases and does not suffer from hypothetical bias. Additionally, precise information from price changes in similar products can be exploited.

Self-stated data consists of asking directly how much they are willing to pay for a certain product. The main advantage of the method is the simplicity, nevertheless, suffers a lot of limitations related to self-reporting biases.

Conjoint analysis is divided in two types of methods. First, in the choice-based conjoint methods individuals are asked to choose one hypothetical product profile from several sets of products. Then, the rating-based methods are those in which several hypothetical product types are presented to individuals and they have to rate them.

In experimental auctions individuals use different auction mechanism to reveal their reservation prices. This is one of the main advantages of the method. In addition, the fact that reservation prices are empirically observed in auctions, permits the researchers to leave aside parametric assumptions such as normality of distributions.

Finally, contingent valuation methods consists of dichotomous choice questions of whether to buy or not at a random price that helps to arrive to an estimation of the willingness to pay using statistical methods. This kind of technique helps identifying the whole demands curves because you can get the proportion of people buying at different offered prices. Although it is easy to implement, it requires big samples in order to obtain accurate results.

This paper is based on data from a survey of contingent valuation. Applications to sport economics include estimation of public willingness to pay to attract or retain professional sports teams. Some examples are the documentation of willingness to pay for basketball and baseball venues (Johnson and Whitehead, 2000); for attracting a professional hockey team (Johnson, Groothuis, and Whitehead, 2001); for retaining professional football and basketball teams (Johnson, Mondello, and Whitehead, 2007); and for supporting amateur sports and recreation programs (Johnson et al. 2007). These studies used aggregated willingness to pay to assess whether sports generated sufficient positive externalities to justify government finance.

We use simulated market transactions as are commonly employed in audit studies to elicit willingness to purchase randomly priced tickets of 8 varsity sports of a given University.

Willingness to purchase declines in market price, generating implied demand curves for each sport.

The model selected in this paper will be the multivariate probit. The first paper using the multivariate probit was Ashford & Swoden (1970) which studied the case where an organism has a response to a stimulus in a multi-variate way. In this way, a model where the observed response of an organism is determined by the joint reaction of several underlying physiological systems is presented. Using a variant of probit analysis authors achieve a multi-variate generalization of the probit analysis in order to understand the underlying system behind the observed response of an organism.

The use of multivariate probit used to be scarce on its beginning as a consequence of the difficulty of computation, nevertheless recent computational advances helped to increase its use. Here we name some of the examples of the uses of the model. First of all, the previously mentioned paper of Ashford & Swoden applied the model to a biological system. Additionally, Contoyannis & Jones (2004) use it for studying of a health production function and the parameters of lifestyle equations; Chib & Greenberg (1998) applied it to commuting choice and buying a car choice; Young et al. (2009) uses it in the modeling of the types of claim in a portfolio of insurance policies; and Amemiya (1978) developed a variant of the multivariate probit where one of the variables is partially observed.

In general Maximum Simulated Likelihood is the estimation method selected as opposed to the maximization of the exact likelihood due to the difficulty of finding closed forms to the probabilities maximized.

3. METHODOLOGY

Consider the circumstance where individual *i* faces an exogenously given set of prices for tickets of *J*=8 college sports, denoted asP_i^1, P_i^2, …, P_i^J. The choice of purchasing any one ticket is given by

$$\begin{aligned} B_i^{1*} &= \beta_0^1 + \beta_1^1 P_i^1 + \boldsymbol{Z}_i' \boldsymbol{\delta}^1 - \eta_i^1 \\ B_i^{2*} &= \beta_0^2 + \beta_1^2 P_i^2 + \boldsymbol{Z}_i' \boldsymbol{\delta}^2 - \eta_i^2 \\ &\vdots \\ B_i^{J*} &= \beta_0^J + \beta_1^J P_i^J + \boldsymbol{Z}_i' \boldsymbol{\delta}^J - \eta_i^J \end{aligned} \tag{1}$$

where B_i^{j*} , *j = 1, 2, …* , J represents the latent utility associated with purchasing a ticket of sport *j* at the offered price, P_i^j , holding constant a vector of individual characteristics Z_i and unobservable factors represented by η_i^j. In practice, we do not observe B_i^{j*}, but we do observe the choice of whether or not the individual buys each ticket. For *j = 1, 2, …* , J indexing each sport, let the binary choices be represented by

$$B_i^j = \begin{cases} 1 & if\ B_i^{j*} > 0 \\ 0 & if\ B_i^{j*} \leq 0 \end{cases} \tag{2}$$

which implies that individual *i* purchases the ticket if it yields positive utility. In other words, the unobserved latent utility of purchasing a ticket for sport j is higher than that of not purchasing. We assume the unobserved factors in equation (1) are jointly distributed standard normal with correlation coefficientsρ_{jk}, for $j \neq k$

$$\begin{bmatrix} \eta_i^1 \\ \eta_i^2 \\ \vdots \\ \eta_i^J \end{bmatrix} \sim N\left(\mathbf{0}, \begin{bmatrix} 1 & \cdots & \rho_{1J} \\ \vdots & \ddots & \vdots \\ \rho_{1J} & \cdots & 1 \end{bmatrix}\right) \tag{3}$$

By calculating the joint probabilities that Pr[B_i^1=1, B_i^2=1, …, B_i^J=1], Pr[B_i^1=1, B_i^2=1, …, B_i^J=0], and so on up to Pr[B_i^1=0, B_i^2=0, …, B_i^J=0], we find the following log-likelihood function for the model:

$$L\left(B_i^1, B_i^2, \dots, B_i^J \mid P_i^1, P_i^2, \dots, P_i^J, \boldsymbol{Z}_i, \boldsymbol{\beta_0}, \boldsymbol{\beta_1}, \boldsymbol{\delta}, \boldsymbol{\rho}\right) = \sum_{i=1}^{n} \log\{\Phi_{\mathrm{J}}\,[q_i^1\left(\beta_0^1 + \beta_1^1 P_i^1 + \boldsymbol{Z}_i^{'}\boldsymbol{\delta}^1\right), \dots, q_i^J\left(\beta_0^J + \beta_1^J P_i^J + \boldsymbol{Z}_i^{'}\boldsymbol{\delta}^J\right), Q \cdot \boldsymbol{\rho}]\} \tag{3}$$

where $\boldsymbol{\beta_0} = [\beta_0^1, \dots, \beta_0^J]$, $\boldsymbol{\beta_1} = [\beta_1^1, \dots, \beta_1^J]$, $\boldsymbol{\delta} = [\delta^1, \dots, \delta^J]$, Φ_{J} is the *J*-variate standard normal cumulative distribution function; and for all *j*, q_i^j is conveniently defined as $q_i^j = 2B_i^j - 1$. The term $(Q \cdot \boldsymbol{\rho})$ is the $J \times J$ variance-covariance matrix but defined in terms correlation coefficients. It equals to the dot-product between the $J \times J$ symmetric matrix of correlation coefficients between equations ($\boldsymbol{\rho}$) and a $J \times J$ symmetric matrix (Q) where each entry $Q_{jk} = Q_{kj} = q_j q_k \rho_{jk}$ for $j \neq k$, and $Q_{jk} = 1$ for $j = k$.

Estimation of willingness to pay and/or reservation prices

The parameter estimates from (3) can be used to obtain the estimated willingness to pay for each college sport. This can be done following the contingent valuation approach by Cameron (1988) and Train (2003 pp. 168).

Noting that the reservation price will be the price at which individual *i* is indifferent between purchasing and not purchasing the ticket, we can derive an estimate of the reservation price for each of the *J* sports as the value of P_i^j such that $\Pr\left[B_i^j = 1\right]$ = 0.50. This probability of purchasing represents the demand for the single sport *j* by the individual i. Algebraically, the reservation prices p_i^j implicitly solve the equation for each *j*:

$$\Pr\left[B_i^j = 1\right] = \Phi\left(\hat{\beta}_0^j + \hat{\beta}_1^j p_i^j + \boldsymbol{Z}_i'\hat{\boldsymbol{\delta}}^j\right) = 0.5 \tag{4}$$

where Φ is the univariate standard normal cumulative distribution function and "^" indicates the estimated value of the parameter. Given that the standard normal probability density function is symmetric and centered on zero, the implicit equations for each p_i^j simplifies to

$$\hat{\beta}_0^j + \hat{\beta}_1^j p_i^j + \boldsymbol{Z}_i'\hat{\boldsymbol{\delta}}^j = 0$$
$$p_i^j = -\frac{1}{\hat{\beta}_1^j}\left(\hat{\beta}_0^j + \boldsymbol{Z}_i'\hat{\boldsymbol{\delta}}^j\right) \qquad (5)$$

The natural measure of the reservation price is the one that sets the probability of purchase equal to 0.5 as in equation (4). However, in principle we could evaluate (4) at any common probability of ticket purchase. Because the probability of purchasing a ticket is monotonically decreasing in price, the rank order of individual preferences for men's and women's basketball is invariant to the choice of probability, as we show in Appendix A. Furthermore, as shown by Cameron (1988), a contingent valuation approach will yield numerically identical estimates of the willingness to pay as we derive at the probability of purchase 0.5. A proof is also provided in Appendix A.

Estimation of unobserved correlation between sports

Parameters in matrix $\boldsymbol{\rho}$, say ρ_{jk} for $j \neq k$, measures the unobserved correlation in the errors terms of the demand equations for tickets of sports j and k. A positive estimate implies that unmeasured attributes that increase the likelihood of purchasing a ticket for *j* also increase the likelihood of purchasing a ticket for *k*. A positive and large correlation for any two sports opens the possibility for efforts to cross-market tickets of one sports to purchasers of the other sport, and vice versa. The intuition behind, for example a positive error correlation, is that customers may have an unobserved passion for sports *per s*e, regardless of the sport discipline.

Estimation of conditional probabilities of purchasing a ticket

Important conclusions can be derived from computing the probability of purchasing a ticket for sport *j* conditional on having purchased a ticket for sport *k*. For example, we can document the complementarities (or substitutability) between any pair of sports (*j*,*k*) by comparing the reservation price that solves the conditional probability of purchasing a ticket for sport *j* conditioned on purchasing a ticket for sport *k*, with the reservation price that solves the unconditional probability of purchasing a ticket for sport *j*. We would expect that the former is greater than the latter when there exists complementarity between sports *j* and *k*.

Dropping the subscript *i* that indexes individuals of the sample, we are interested in computing the reservation price for sport *j* that equalizes to 0.5 the expected value of purchasing a ticket of sport *j* ($B^j = 1$) conditional on purchasing a ticket of sport k ($B^k = 1$), and given the observed decision each individual made regarding the remaining sports ($B^{-j} = b^{-j}$). The conditional probabilities computations following Christofides et al. (1997), Greene (2003 pp. 713), and Greene (2009 pp.76) is:

$$\begin{aligned} 0.5 &= E\left[B^j = 1 | B^k = 1, B^{-j} = b^{-j}\right] \\ &= \Pr\left[B^j = 1 \middle| B^k = 1, B^{-j} = b^{-j}\right] \\ &= \frac{\Pr\left[B^j = 1, B^k = 1, B^{-j} = b^{-j}\right]}{\Pr\left[B^k = 1, B^{-j} = b^{-j}\right]} \end{aligned} \tag{6}$$

with $b^{-j} = \{0,1\}$ depending on the individual's decision, and where the last equality results simply from applying the definition of conditional probability as the ratio between the joint probability of the *J* sports, to the joint probability of the conditioning *J*-1 random variables.

Using the expression arising from equation (3) which writes the probabilities of purchasing tickets in terms of the multivariate Normal distribution, we find the reservation price p^j that solves the following highly nonlinear expression, which we denote by p^{jk}:

$$0.5 = \frac{\Phi_{\mathrm{J}}\left[\left(\beta_0^j + \beta_1^j p^{jk} + \boldsymbol{Z}'\boldsymbol{\delta}^j\right), \left(\beta_0^k + \beta_1^k P_i^k + \boldsymbol{Z}_i^{'}\boldsymbol{\delta}^k\right), q^{-j}\left(\beta_0^{-j} + \beta_1^{-j} P^{-j} + \boldsymbol{Z}'\boldsymbol{\delta}^{-j}\right), Q \cdot \boldsymbol{\rho}\right]}{\Phi_{\mathrm{J}-1}\left[\left(\beta_0^k + \beta_1^k P_i^k + \boldsymbol{Z}_i^{'}\boldsymbol{\delta}^k\right), q^{-j}\left(\beta_0^{-j} + \beta_1^{-j} P^{-j} + \boldsymbol{Z}'\boldsymbol{\delta}^{-j}\right), Q \cdot \boldsymbol{\rho}\right]} \qquad (7)$$

For example, the numerator represents the joint decision (demand) from an individual for all the *J* sports, arising from the value that takes the unobserved latent utility from attending each sport at the offered price. The interpretation of the denominator is similar except that is the joint decision but for the *J*-1 sports. In general, as we have *J* sports, this comparison can be made either conditioning on purchasing a ticket for one sport *k* and on the observed decision on the remaining (*J*-1) sports, or any combination of purchasing/not purchasing tickets for the remaining (*J*-1) sports. That is, without loss of generality, the same derivation is valid for analyzing the complementarities or substitutability of any bundle of tickets of the *J* varsity sports.

Importantly, interesting conclusions can be obtained by comparing the reservation price p^j which solves (7) with the unconditional reservation price which solves (5), because this difference quantifies, in terms of willingness to pay, the complementarity or substitutability between any pair of sports *j* and *k*, for $j \neq k$.

Direct and cross-price effects through unobserved taste

Another interesting application derived from this methodology is the calculation of own price and cross price elasticities. Both own- and cross-price elasticites (and marginal effects) can be directly calculated using the likelihood estimation.

The probability of purchasing a ticket for sport *j* represents the estimated demand of individual *i* for sport *j*. Therefore, the own-price elasticities of sport *j*, reflecting the change in the probability of purchasing a ticket for that sport after a change in its own price offered to individual *i* (P_i^j), can be computed as follows:

$$\varepsilon_i^{jj} = \frac{1}{n}\sum_{i=1}^{n} \frac{\partial P\left[B_i^j = 1\right]}{\partial P_i^j} \cdot \frac{P_i^j}{P\left[B_i^j = 1\right]} \tag{8}$$

where *n* is the number of individuals in the sample.

For the cross-price elasticities between two sports, consider de simultaneous joint decision of purchasing tickets for two sports (*j* and *k*) at the offered price P^j and P^k, respectively, and the observed decision regarding the remaining sports (-*j*). Expressed in terms probabilities, we have,

$$\begin{aligned} &\Pr\left[B^j = 1, B^k = 1, B^{-j} = b^{-j}\right] \\ &= \Phi_J\Big[\left(\beta_0^j + \beta_1^j P^j + \boldsymbol{Z}'\boldsymbol{\delta}^j\right), \left(\beta_0^k + \beta_1^k P^k + \boldsymbol{Z}_i'\boldsymbol{\delta}^k\right), q^{-j}\big(\beta_0^{-j} + \beta_1^{-j} P^{-j} \\ &\qquad + \boldsymbol{Z}'\boldsymbol{\delta}^{-j}\big), Q.\boldsymbol{\rho}\Big] \\ &= C \end{aligned} \tag{9}$$

which consists of the joint probability (demand) derived from the fact that, individually, the value of the unobserved latent utilities from attending each sport at the offered prices (P^j, P^k, P^{-j}) is higher than that of not attending in the cases the individual decided to purchase, and is lower in the cases the individual decides not to purchase. The value of C is a scalar between 0 and 1, and is the probability associated with this individual's decision given his or her particular characteristics. In order to find the cross price elasticity of buying sport *j* to changes in the price of sport *k*, let us define $P^{k*} = P^k(1 + 1\%)$ [3]. By substituting P^k by P^{k*} in (9) and leaving everything else unchanged, we can evaluate the induced change in P^j which leaves

[3] We arbitrary set the increment to 1%, but any other one could be defined.

unchanged the probability (C) of purchasing this combination of sports tickets. We denote this solution by P^{j*}, as the price of sport *j* which solves the following non-linear implicit function:

$$\begin{aligned}&\Pr\left[B^{j}=1,B^{k}=1,B^{-j}=b^{-j}\right]\\&=\Phi_{J}\left[\left(\beta_0^{j}+\beta_1^{j}P^{j*}+\boldsymbol{Z}'\boldsymbol{\delta}^{j}\right),\left(\beta_0^{k}+\beta_1^{k}P^{k*}+\boldsymbol{Z}_i'\boldsymbol{\delta}^{k}\right),q^{-j}\left(\beta_0^{-j}+\beta_1^{-j}P^{-j}\right.\right.\\&\qquad\qquad\left.\left.+\boldsymbol{Z}'\boldsymbol{\delta}^{-j}\right),Q.\boldsymbol{\rho}\right]\\&=C\end{aligned} \tag{10}$$

Finally, in order to obtain the cross-price elasticity between *j* and *k*, i.e. the change in demand of sport *j* due to a 1% change in price of sport *k*, we employ equation (8) but where the change in price P^{j} in the denominator (which is exactly equal to $[P^{j*}/P^{j}-1]$) is that induced by the 1% change in P^{k}.

An illustration for the case of two sports implies working with the joint probability $\Pr\left[B^{j}=1,B^{k}=1\right]=C$. The previous equation can be expressed as a restatement of equation (9) but only for two sports, in the following way: $\Phi_2\left[\left(\beta_0^{j}+\beta_1^{j}P^{j}+\boldsymbol{Z}'\boldsymbol{\delta}^{j}\right),\left(\beta_0^{k}+\beta_1^{k}P^{k}+\boldsymbol{Z}_i'\boldsymbol{\delta}^{k}\right),Q.\boldsymbol{\rho}\right]=C$. Applying total differentiation on both sides of the previous equation we obtain: $\mathrm{d}(\Phi_2\left[\left(\beta_0^{j}+\beta_1^{j}P^{j}+\boldsymbol{Z}'\boldsymbol{\delta}^{j}\right),\left(\beta_0^{k}+\beta_1^{k}P^{k}+\boldsymbol{Z}_i'\boldsymbol{\delta}^{k}\right),Q.\boldsymbol{\rho}\right])=0$. Following Greene (2003) and remembering that P^{j} and P^{k} only appear in the decision of sport *j*, *k* respectively, we can compute the total differentiation as:

$$\left[\Phi^{j}\left(\beta_0^{j}+\beta_1^{j}P^{j}+\boldsymbol{Z}'\boldsymbol{\delta}^{j}\right)\beta_1^{j}\right]dP^{j}+\left[\Phi^{k}\left(\beta_0^{k}+\beta_1^{k}P^{k}+\boldsymbol{Z}'\boldsymbol{\delta}^{k}\right)\beta_1^{k}\right]dP^{k}=0 \tag{11}$$

where Φ^{i} are univariate standard normal densities for i=*j*, *k*. Rearranging terms in (11) we can derivate the Marginal rate of substitution between sport *j* and *k* for the probability of purchase C as:

$$MRS_1 = -\frac{dP^j}{dP^k} = \frac{\Phi^{\mathrm{j}}\left[\left(\beta_0^j + \beta_1^j P^j + \boldsymbol{Z}'\boldsymbol{\delta}^j\right)\right]\beta_1^j}{\Phi^{\mathrm{k}}\left[\left(\beta_0^k + \beta_1^k P^k + \boldsymbol{Z}'\boldsymbol{\delta}^k\right)\right]\beta_1^k} \tag{12}$$

Equation (12) allow us to have, for a given pair of prices P^j and P^k, the point at the indifference curve at the joint probability level C. This is illustrated in Panel A of figure 3.1. Additionally, these same prices at certain position in the indifference with MRS_1 are consistent with a level of marginal probability of buying sports *j* and *k* as illustrated by $P(B^j = 1;\ P^j)$ and $P(B^k = 1;\ P^k)$ in panels B and C of figure 3.1, respectively. Appling the exercise previously described can let us move in the indifference curve towards a new point with another marginal rate of substitution(MRS_2)that give us another marginal probability of buying sports *j* and *k* associated to prices P^{j*} and P^{k*} that are illustrated by $P(B^j = 1;\ P^{j*})$ and $P(B^k = 1;\ P^{k*})$. Using this, we can compute the change in the probability of buying sport *k* when we exogenously increase price of sport *j* but we continue in the same indifference curve associated with the joint probability level C. We can also see the change in the marginal rate of substitution. It is important to notice that illustration of figure 3.1 assumes that sports are substitutes; nevertheless, the same exercise is valid for complementary goods. The functional form of the indifference curve in our case will dependent on the signs of the elements of the unobserved correlation matrix $\boldsymbol{\rho}$ that is what will be leading the process by which we can compute the cross-price elasticities. In other words, the sign of elements in matrix $\boldsymbol{\rho}$ should be in line with the sign of the cross-price elasticities.

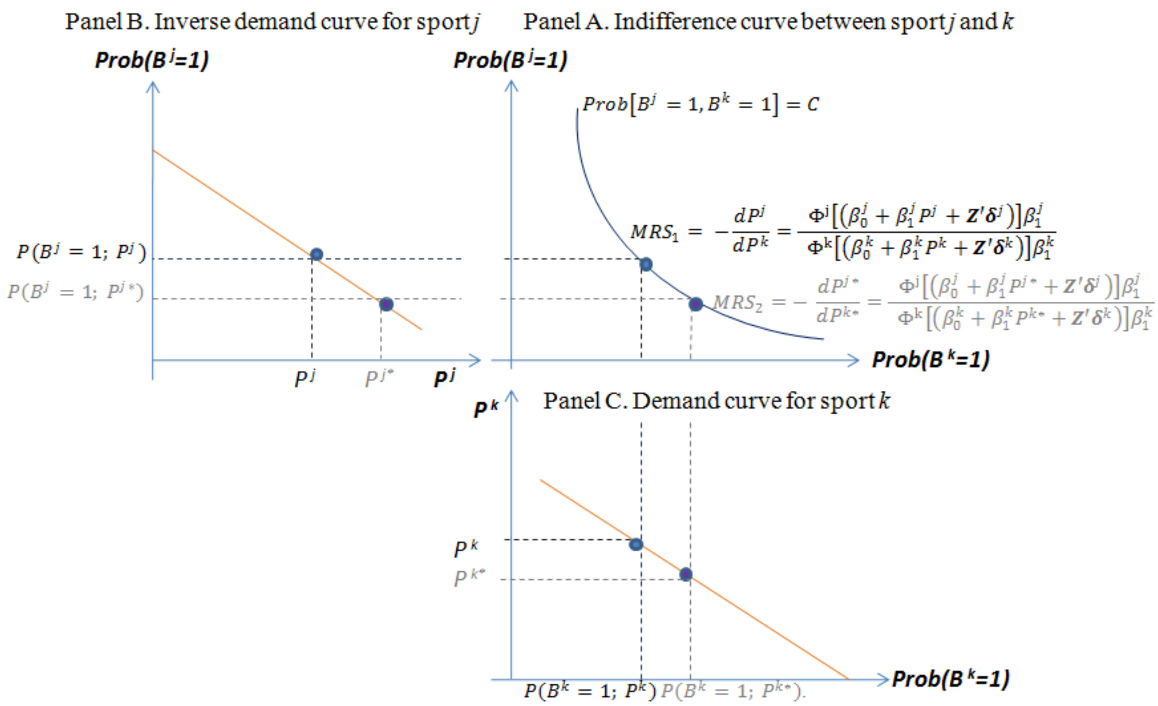

Figure 3.1. Derivation of cross-price elasticities between any pair of sports *j* and *k*

As it will became clear in the estimation subsection, this procedure is applied because the randomization in the experimental design yields that each sport´s demand is not explained (from the statistical significance point of view) by the cross-prices. Therefore, this approach is exploiting the estimated unobserved correlation between sports.

4. DATA

The theoretical model requires that each individual be given the opportunity to purchase or refuse a ticket to attend a college game. We also require sufficient variation in prices to identify the behavioral responses, $\hat{\beta}_1^j$ for $j = 1,2,\ldots,J$ in equation (1). Such data would not be commonly available in market transactions, and so we developed an artificial environment where appropriate data could be obtained. The universe of potential customers in this study is the undergraduate student population at Iowa State University.

In 2007, a random sample of 2000 students was invited by Email to participate in a web-based survey. Of these, 470 (23.5%) provided complete responses, and they represent the working sample for this study. The working sample reflected the attributes of the sample universe well, and so there does not appear to be any systematic relationship between observable student attributes and the likelihood of survey response. A description of the working sample is showed in Table A1 in appendix A.

In addition to demographic information and questions related to participation and interest in sports, each respondent was asked whether they would purchase the next sports tickets: a) a women's basketball ticket at a stated price, b) a men's basketball ticket at a stated price, c) a football ticket at a stated price, d) a volleyball ticket at a stated price, e) a wrestling ticket at a stated price, f) a gymnastics ticket at a stated price, g) a hockey ticket at a stated price, and h) a soccer ticket at a stated price.

Prices were generated from independent random draws from uniform price distributions. The men's basketball price was a whole number drawn from a U(7, 20). Football prices were drawn from a U(13,25). The rest of the sports prices were drawn from a U(1, 10). The actual price charged to students at the time was the median price in each distribution.

Variable definitions and sample statistics are presented in Table A1 on the appendix as mentioned before. Because each respondent received a randomly drawn price, there is no apparent correlation with the control variables Z_i. As a result, we get virtually identical response parameters to the prices regardless of whether or not the Z_i's are included in the estimation. Nevertheless, it is useful to highlight some of the more interesting control variables used in the analysis. To control for overall interest in sports, we included information on whether the individual played sports in high school and whether the individual participated in intramurals in school. To control for demand for Iowa State sports more specifically, we ask whether the individual was an Iowa State fan before coming to college, whether they grew up in Iowa, and whether they had parents or relatives graduate from Iowa State. We control for differences in tastes between men and women by including a gender dummy variable.

A useful check on the success of the price randomization is to plot the probability of a positive purchase response by random price offered. If respondents are behaving as expected rational agents, we should be able to trace out standard demand relationships from data plots. Figures 4.1 shows the fraction of people that bought a ticket of men basketball differencing if they have also bought a college football ticket. While the relationship is not perfectly monotonic, the fitted bivariate relationship between intent to purchase and price satisfies the law of demand: the higher the price, the lower the proportion of people willing to buy the ticket.

An important feature is that there are differences in demand relationships between people who purchased football and those who did not. This descriptive result shows the possibility of complementarities between these two sports. That is, in general, it is expected that willingness to purchase a given sport's ticket might be different conditional on the decision the person made on other sports. For example, in the case of football and men's basketball, those individuals who purchased football tickets are more likely to demand a men's basketball ticket. Therefore, if eagerness to purchase a ticket is positively correlated with willingness to pay for the ticket, it would be interesting to study conditional reservation prices. In fact, this descriptive relation will be tested for all the combinations of sports later in the paper.

Another important issue of the dataset is that each individual receive a price offer (in which the price was randomly distributed) but the order of the offers where always the same. This order was: women basketball, men basketball, football, volleyball, wrestling, gymnastics, hockey and soccer. This data structure could a priori end up biasing the estimations, as a consequence of a priming effect generated by the price offer previous to the actual one. That is, the fact that the individual already received a low (high) price offer and that decided to buy (not buy) that sport might affect the actual decision. In order to test this, we include in table A2 in appendix A as a robustness check, a simulated maximum likelihood estimation but including for each sport the price of the sport that was previously offered. Results showed no significance of the previous sport in almost all cases. Additionally, the inclusion of this previous price did not affect the own price coefficient even in the few cases when the previous price shows significance[4].

The previous results gave us the security that the specification of the decision making process is correct given the structure of the data generated. In this sense, this robustness results let us see two things: a) our model includes enough explanatory variables and adding other prices does not change our main coefficients, and b) the fact that the previous price offer does not affect the decision of buying the actual sports that is being offered at certain price means that the decision of buying sports is not done sequentially but simultaneously. In this sense, the estimated seemingly unrelated probit (multivariate probit) is the appropriate specification, compared to the conditional probits, multinomial probits, multinomial logits[5], conditional probits, and nested logits, because all those show a more sequential kind of decision making.

[4] A similar robustness check was done including all cross-prices and similar results were obtained. Almost no price of sport *j* different from sport *i* was significant to explain decision of buying or not sport *i*.
[5] For a comparison between multivariate probit and multinomial logit see Young et al. (2009).

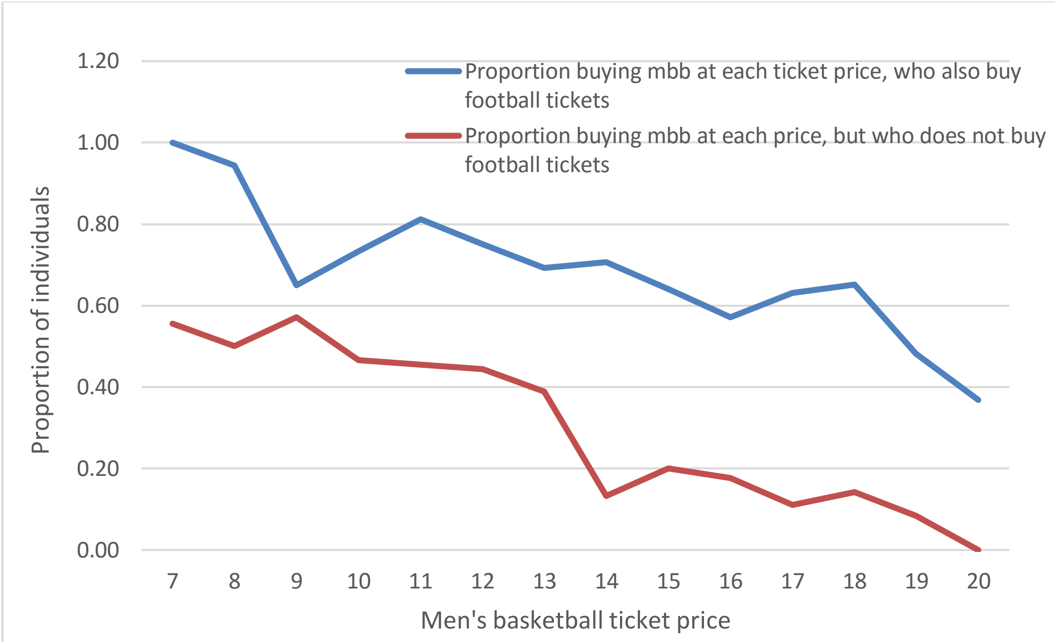

Figure 4.1. Share of individuals (y-axis) who purchased men's basketball tickets as a function of offered price

5. RESULTS

A. Point estimates

System in (2) constitutes a seemingly unrelated 8-variate probit model, and its estimation is conducted by means of simulated maximum likelihood, which is more accurate than conventional maximum likelihood in these cases of multivariate models (Cappellari and Jenkins, 2003). A total of 400 draws from the distribution of likelihood function are taken to compute the simulated likelihood in each iteration of the optimization algorithm. Table 1 shows the estimation results and statistical significance of parameters.

Table 1. Simulated maximum likelihood estimation of seemingly unrelated model in (2). Dependent variable: Binary choice of purchasing ticket of sport *j*, B^j, with $j = 1,2,\ldots,8$.

	1	2	3	4	5	6	7	8
Dep. variable	womenbb	menbb	football	volleyball	wrestling	gym	hockey	soccer
own price	-0.126***	-0.106***	-0.076***	-0.176***	-0.143***	-0.133***	-0.168***	-0.118***
gender	-0.545***	0.084	-0.101	-0.294**	-0.144	-0.672***	-0.209	-0.268*
hi_sport	0.183	0.037	-0.037	0.170	0.634***	0.034	-0.025	0.124

hi_inter	0.052	0.049	0.224	-0.007	-0.229	-0.109	0.036	0.055
isu_inter	0.239	0.315**	0.469***	-0.018	0.318**	-0.065	0.112	0.056
isu_famil	0.004	-0.002	-0.068	-0.112	0.006	0.030	-0.187	-0.053
isu_fan	0.490***	0.425***	0.444***	0.262*	0.335**	0.117	0.119	0.308**
constant	0.653***	1.085***	1.304***	0.709***	0.100	1.046***	1.197***	0.277
N = 449 – Simulation draws = 400 – LogLikelihood = -1,919 – Iterations = 33								

For the above estimation, the Wald test of the joint significance of all estimated parameters in the regression equations yielded that they are jointly statistically significant. Also, the likelihood ratio test was conducted for the joint significance of the ρ_{jk}, and yielded that they are statistically significant. Finally, a test conducted for each equation to analyze the joint significance of the individual characteristics included as explanatory variables, gave mixed results, but for simplicity we opted to maintain the same controls across equations (except the own price)

B. Correlation in unobserved demand for men's and women's basketball

Unobserved correlation between equations j and k in model (2) is captured by parametersρ_{jk}. Estimation results show that for all pairs of sports, there exists a positive correlation, implying that the unobserved factors that prompt individuals to purchase tickets for one sport are positively correlated with those unobserved factors that prompt individuals to purchase tickets of the other sports. Moreover, in all cases (except for the combination of football and soccer) estimations are statistically significant. This empirical result, shown in Table 2, can be used by the Athletic Department to identify the pair of sports with the greatest correlation coefficients in order to generate marketing strategies involving both sports, because those generated for any one sport, may have positive impacts on other sports. Examples of these are women's basketball and volleyball, and men's basketball and football, which may be expected, but also soccer with volleyball, gym and hockey, and volleyball with gym.

Table 2. Estimation of correlation coefficient parameters ρ_{jk}, with j,k = 1,2,…,8.

Dependent variable	women bb	men bb	football	volleyball	wrestling	gym	hockey	soccer
women bb	1.00	0.51	0.20	0.70	0.44	0.59	0.42	0.52
men bb	0.51	1.00	0.60	0.38	0.30	0.33	0.28	0.30
football	0.20	0.60	1.00	0.20	0.23	0.24	0.27	0.12
volleyball	0.70	0.38	0.20	1.00	0.53	0.60	0.52	0.57
wrestling	0.44	0.30	0.23	0.53	1.00	0.45	0.42	0.41
gym	0.59	0.33	0.24	0.60	0.45	1.00	0.55	0.62
hockey	0.42	0.28	0.27	0.52	0.42	0.55	1.00	0.57
soccer	0.52	0.30	0.12	0.57	0.41	0.62	0.57	1.00

Note: All parameters are statistically significant at 5% except for ρ_{38} and ρ_{83}

C. Willingness to pay or reservation prices

Results for this section arise from solving equation (5) for a reservation price for each individual in the sample and for each of the 8 sports. Reservation prices, in this context, are defined as the price level that leaves the individual indifferent between purchasing or not purchasing a ticket for the corresponding sport. For each sport's reservation price, Table 3 show the mean, median and standard deviation over the sample of 449 individuals. When comparing mean and median reservation price for each sport with the actual ticket price charged, we can identify for which sports students are (on average) willing to pay a higher price than that charged at the ticket office. Results show that this is the case of women´s and men´s basketball, football and hockey. They also show that wrestling prices are set roughly at the same level than reservation prices.

Table 3. Estimated unconditional reservation prices of sport tickets, and comparison with actual ticket prices. Based on equation (5)

	women bb	**men bb**	**football**	**volleyball**	**wrestling**	**gym**	**hockey**	**soccer**
Actual price	5.00	13.50	19.00	5.00	5.00	5.00	5.00	5.00
				Estimated reservation price				
Median	7.08	14.46	23.33	4.07	5.17	3.95	6.72	3.42
Mean	7.00	14.67	23.40	4.12	5.01	5.28	6.76	3.31
Std. deviation	3.06	2.62	4.37	1.17	2.83	2.74	0.83	1.76
N	449	449	449	449	449	449	449	449

Interesting conclusions for the Athletic Department may be extracted from the interval over which reservation prices for each sport ranges, because it may give a clearer picture of the proportion of people willing to pay a higher price than the actual ticket price. These cannot be extracted from analyzing the mean or median of reservation prices for each sport. Additionally, while it is clear that a price increase may reduce demand by a quantity that make revenues to decrease, our setup let us compute the exact price that maximize revenues, because we identified the whole demand function for each sport. For the sake of space we leave this analysis for future research.

D. Documenting the complementarity between sports based on conditional probabilities of purchasing

Results in this section arise from numerically solving the nonlinear equation in (7) for a reservation price (p^{jk}) for each individual in the sample who decided to purchase both tickets, say for example, women's basketball (j) and men's basketball (k). The solution of equation (7) for a reservation price is performed by evaluating explanatory variables in vector $\boldsymbol{Z}$ at the observed value for each individual, and B^{-j} at the actual decision each individual made (which, in turn, implies computing the actual value of q^{-j} for each individual). The Bisection methods is employed given that it consists of one equation in one unknown. Following this procedure we compute each pair of reservation prices. The comparison between the conditional to the unconditional (estimated in C.) reservation prices for each sport, allows extracting interesting conclusions about the taste for sports.

Table 4 summarize the results. It can be easily seen that conditional reservation prices are higher that their unconditional counterparts for each combination of sports, which has an appealing economic interpretation. This shows that there is a complementarity between all sports, and also that is in line with the conclusions drawn in subsection B. People that already have a ticket to one sport event have higher willingness to pay for the other sport. For example, individuals who purchased a men's basketball ticket are on average willing to pay 41% more for a women's basketball ticket, compared to the unconditioned situation. This is consistent with the estimation of unobserved taste. When individuals already decide to go to one sport, they are willing to pay more for another sport because they find this satisfactory in line with the unobserved common taste for all sports.[6]

Table 4. Estimated Percent increment of conditional reservation prices of sport tickets relative to unconditional reservation prices. Percentage increment defined as $\frac{(P^{jk}-P^{j})}{P^{j}}$ where P^{jk} is de conditional price of j on k and P^{j} the unconditional price of j. Based on equation (7).

	k	1	2	3	4	5	6	7	8
j	Labels	womenbb	menbb	football	voleyball	wrestling	gymnastics	hockey	soccer
1	womenbb		40.5%	29.4%	48.4%	30.4%	40.9%	31.0%	41.5%
2	menbb	40.7%		37.0%	40.9%	29.7%	39.4%	33.9%	33.6%
3	football	21.8%	20.2%		19.8%	12.5%	20.8%	18.6%	16.4%
4	voleyball	85.4%	79.9%	74.2%		68.7%	78.6%	86.6%	95.1%
5	wrestling	54.4%	43.3%	40.6%	70.7%		67.4%	44.6%	57.4%
6	gymnastics	44.5%	52.0%	32.8%	48.7%	36.7%		46.5%	33.0%
7	hockey	38.6%	39.1%	32.5%	46.7%	23.7%	45.8%		44.6%
8	soccer	156.4%	160.8%	142.7%	169.8%	142.7%	172.7%	170.4%	

Note: Main Diagonal elements have no significant meaning so are dropped out.

In addition to these results, we can check how the variability of individual's choices shift between the conditional and the unconditional case. We compare the coefficient of variation of unconditional reservation prices (standard deviation of P^{j} over median P^{j}) and the coefficient of variation of conditional reservation prices (standard deviation of P^{jk} over median P^{jk}) to conclude about the concentration and variability of prices[7]. This exercise shows that although in the majority of the cases the conditional reservation prices have more variability and are less concentrated than

[6] In appendix A, we show the two tables presenting the median P^{jk} and P^{j} used to construct each cell of table 4.
[7] Detailed results for each sports pair are not shown due to space, but are available from the authors upon request.

unconditional reservation prices, there is no clear pattern, preventing us from extracting generalized conclusions about changes in concentration of prices.

Table 5 below shows an interesting pattern consistent with table 4. When we compute the change on the skewness of conditional prices to unconditional ones, we find that in almost every case the difference is negative, meaning that the bell shaped curve of the distribution of the data after conditioning moves towards the right. Moving towards the right implies that all the distribution of people tend to accept higher prices of sport *j* when they are conditioned on buying sport *k* than in the case where they are not conditioned.

Table 5. Difference of estimated skewness coefficients between conditional and unconditional reservation prices. Each entry is equal to: $\left(s_{jk}^3 - s_j^3\right)$, where s_{jk}^3 (s_j^3) is the skewness coefficient of variable p^{jk} (p^j)

	k	1	2	3	4	5	6	7	8
j	Labels	womenbb	menbb	football	voleyball	wrestling	gymnastics	hockey	soccer
1	womenbb		-0.26	-0.20	-0.37	-0.39	-0.50	0.09	-0.71
2	menbb	-0.83		-0.33	-1.07	-0.42	-0.98	-0.53	-0.64
3	football	-0.76	-0.82		-1.08	-0.87	-1.14	-1.54	-0.93
4	voleyball	-0.36	-0.63	-0.43		-0.12	-0.86	-0.53	-0.80
5	wrestling	-0.54	-0.57	-0.03	-0.46		-0.68	-0.59	-0.16
6	gymnastics	-0.28	-0.69	-0.05	-0.05	-0.35		-0.55	-0.36
7	hockey	-0.53	-0.58	-0.40	-0.95	-0.26	-0.78		-0.95
8	soccer	-0.72	-1.00	-1.19	-5.13	-1.63	-0.95	-0.41	

Note: Main diagonal elements are dropped out because they have no significant meaning.

E. Estimating own price and cross price elasticities

The Slutsky equations establish that the own price elasticity of demand for sport *j* is $\varepsilon_{jj} = \xi_{jj} - \theta_j\eta_j$, where ξ_{jj} is the compensated elasticity (i.e. the change in demand of *j* which leaves the individual in the same level of utility), θ_j is the budget share for the j^{th} sport, and η_j is the income elasticity of demand for the j^{th} sport. Similarly, the Slutsky equations for cross-price elasticities between any two sports *j* and *k*, $j \neq k$, establish that $\varepsilon_{jk} = \xi_{jk} - \theta_k\eta_j$, where ε_{jk} is the cross-price elasticity of demand due to a change in price of sport k, ξ_{jk} is the compensated cross-price elasticity (i.e. the change in demand of j which leaves the individual at the same level of utility after a change in price of k), θ_k is the budget share for the k^{th} sport, and η_j is the income elasticity of demand for the j^{th} sport.

The randomization of the experimental design implies that decisions for each sport are made independently from those for the other sports (which in this set up can be translated in the lack of income effects $\theta_j = 0$), yielding an equality between the uncompensated demand elasticities and their compensated counterparts ($\varepsilon_{jk} = \xi_{jk}$ and $\varepsilon_{jj} = \xi_{jj}$). Having this in mind, we present in this section, results for own-price compensated demand elasticities of women´s basketball, and cross-price compensated elasticities of women basketball (j) after a change in men's basketball price (k).

A 1% change in women's basketball price induces a median reduction of 0.46% in demand for this sport (more precisely, in the probability of purchasing tickets). On the other hand, cross-price elasticity has a median of 0.86, implying that a 1% increase in men's basketball price induces an increase of 0.86% in the probability of purchasing women's basketball tickets. As we observe asymmetry in the distribution of both elasticities across individuals, we decided to report them in medians. Comparing cross- and own-price elasticities, we observe that for women's basketball, individuals are more elastic to changes in the price of another good than with respect to their own price. Moreover, the fact that cross-price elasticity is positive, implies that men and women's basketball sports are substitutes. We infer that while the unobserved correlation between sports (shown in part C) was estimated as positive and can be interpreted as "pure taste for sports" as already noticed by Rosas & Orazem (2014), the estimation of the observed part shows a substitutive behavior between sports.

6. CONCLUSIONS

In this paper using data from a contingent valuation survey performed in Iowa State University, we studied the decision of purchasing or not 8 different sports and the relationship between each other.

We show that the decision of each sport is not done sequentially, but rather simultaneously, because the price of the sport offered previously does not affect the decision of the actual price.

This simultaneous decision context makes the multivariate probit model the appropriate estimation method. The fact that the unobserved correlation coefficients show significance is another proof of it, which showed that sports behave as complements to each other because their estimation was of positive sign in all cases.

Furthermore, we estimated the reservation prices (or willingness to pay) of each sport and compared them with the prices offered. We observed that while in some sports there is toom to increase prices, others should be lowered if attendance is the objective pursued.

We also estimated reservation prices of a sport conditional on buying another sport. In line with the sign of the unobserved heterogeneity, we see that in almost all sports people that already desires another sport is willing to pay more for the sport evaluated. This is reflected by comparing the unconditional to the conditional reservation prices. This result shows the possibility of making product bundles between sports in order to increase attendance and revenues of the University.

Finally, we derive a method for estimating cross-price elasticities for a sport when the prices of the other sport do not appear in the decision equation, by exploiting the unobserved correlation and the multivariate probit framework. We find that women´s basketball demand responds in a more elastic way to changes in prices of men basketball than to changes in its own price. Additionally, we find that both sports behave as substitutes because the value of the cross-price elasticity is positive. We conclude that, while the unobserved correlation between sports shows complementarities (that makes conditional reservation prices bigger than unconditional ones), the price effect shows a substitutive behavior.

REFERENCES

Amemiya, T. "The Estimation of a Simultaneous Equation Generalized Probit Model". *Econometrica*, **46**(5), 1978, 1193-1205.

Ashford, J & Sowden, R. "Multi-variate probit analysis". *Biometrics*, **26** (3), 1970, 535-46.

Cappellari, L, & Jenkins, S. "Calculation of multivariate normal probabilities by simulation, with applications to maximum simulated likelihood estimation". *The Stata Journal*, **6**(2), 2006, 159-89.

Cameron, T. "A New Paradigm for Valuing Non-market Goods Using Referendum Data: Maximum Likelihood Estimation by Censored Logistic Regression." *Journal of Environmental Economics and Management*, **15**(3), 1988, 355-79.

Charles, K. K., & Guryan, J. "Prejudice and Wages: An Empirical Assessment of Becker's The Economics of Discrimination." *Journal of Political Economy*, **116**(5), 2008, 773-809.

Chib, S & Greenberg, E. "Analysis of multivariate probit models". *Biometrika*, **85**(2), 1998, 347-61.

Christofides, L.N., Stengos, T, & Swidinsky, R. "On the Calculation of Marginal Effects in the Bivariate Probit Model." Economics Letters, **54**(3), 1997, pp.203-208.

Contoyannis, P & Jones, A. "Socio-economic status, health and lifestyle". Journal of health economics, **23**(5), 2004, 965-95.

Dietz-Uhler, B., Harrick, E. A., End, & C, Jacquemotte, L. "Sex Differences in Sport Fan Behavior and Reasons for Being a Sport Fan." *Journal of Sport Behavior*, **23**(3), 2000, 219-31.

Greene, W.H."Econometric analysis*",* 2003. Pearson Education India.

Greene, W.H. & Hensher, D. "Modeling ordered choices: A primer". Cambridge, 2010.

Jedidi, K & Jagpal, S. "Willingness to pay: measurement and managerial implications". On *Handbook of Pricing Research in Marketing*, 2009.

Johnson, B. K, Groothuis P. A, & Whitehead, J. C. "The Value of Public Goods Generated by a Major League Team: The CVM Approach.'' *Journal of Sports Economics*, **2**(1), 2001, 6–21.

Johnson, B. K., Mondello, M. J., and Whitehead, J. C. “The Value of Public Goods Generated by a National Football League Team.” *Journal of Sport Management*, **21**(1), 2007, 123–36.

Johnson, B. K., & Whitehead, J. C. “Value of Public Goods from Sports Stadiums: The CVM Approach.” *Contemporary Economic Policy*, **18**(1), 2000, 48–58.

Johnson, B. K., Whitehead J. C, Mason D. S & Walker, G. J. “Willingness to pay for amateur sport and recreation programs.” *Contemporary Economic Policy*, **25**(4), 2007, 553–64

Madrigal, R. “Cognitive and Affective Determinants of Fan Satisfaction with Sporting Event Attendance.” *Journal of Leisure Research*, **27**(3), 1995, 205-27.

National Collegiate Athletic Association (NCAA). Football Attendance. 2016. Available at: http://www.ncaa.org/wps/wcm/connect/public/NCAA/Resources/Stats/football/Attendance/index.html. (Accessed June 2016)

National Collegiate Athletic Association (NCAA). Men’s Basketball Attendance. 2016. Available at: http://www.ncaa.org/wps/wcm/connect/public/NCAA/Resources/Stats/M+Basketball/Attendance/index.html. (Accessed June 2016).

National Collegiate Athletic Association (NCAA). Women’s Basketball Attendance. 2016. Available at: http://www.ncaa.org/wps/wcm/connect/public/NCAA/Resources/Stats/W+Basketball/attendance.html. (Accessed June 2016).

Pan, D. W. & Baker, J. A. W. “Mapping of Intercollegiate Sports Relative to Selected Attributes as Determined by a Product Differentiation Strategy.” *Journal of Sport Behavior*, **22**(1), 1999, 69-82.

Pope, D. G. & Sydnor, J. R. “What’s in a Picture? Evidence of Discrimination from Prosper.com.” *The Journal of Human Resources*, **46**(1), 2011, 53-92.

U.S. Department of Education – ED. The Equity in Athletics Data Analysis Cutting Tool. 2016. Available at: http://ope.ed.gov/athletics/. (Accessed June 2016)

Rosas, F & Orazem, P. “Is Self-Sufficiency for Women’s Collegiate Athletics a Hoop Dream? Willingness to Pay for Men’s and Women’s Basketball Tickets”. *Journal of Sports Economics*, 15(6), 2014, 579-600.

Train, K. “Discrete Choice Methods with Simulation” Cambridge University Press, 2003.

Young, G, Valdez, E & Kohn, R. “Multivariate probit models for conditional claim-types”. *Insurance. Mathematics and Economics*, **44**(2), 2009, 214-28.

APPENDIX A

Table A1. Summary statistics.

Variable	Mean	Standard deviation	Description
Bwbb	57.9%	49.4%	= 1 if buys women's basketball ticket
Bmbb	53.5%	50.0%	= 1 if buys men's basketball ticket
Bf	61.9%	48.6%	= 1 if buys football ticket
Bv	42.2%	49.4%	= 1 if buys volleyball ticket
Bwres	49.2%	50.0%	= 1 if buys wrestling ticket
Bgym	47.9%	50.0%	= 1 if buys gymnastics ticket
Bh	58.1%	49.4%	= 1 if buys hockey ticket
Bs	40.4%	49.1%	= 1 if buys soccer ticket
Pwbb	5.63	2.83	Random price of women's basketball ticket offered
Pmbb	13.65	3.99	Random price of men's basketball ticket offered
Pf	18.83	3.81	Random price of football ticket offered
Pv	5.46	2.86	Random price of voleyball ticket offered
Pwres	5.40	2.93	Random price of wrestling ticket offered
Pgym	5.71	2.89	Random price of gymnastics ticket offered
Ph	5.39	2.81	Random price of hockey ticket offered
Ps	5.56	2.90	Random price of soccer ticket offered
Gender	52.2%	50.0%	=1 if male
hi_sport	67.7%	46.8%	= 1 if played varsity sports in high school
hi_inter	38.3%	48.7%	= 1 if plays intramurals in high school
isu_inter	64.6%	47.9%	= 1 if plays intramurals in school
isu_family	40.8%	49.2%	= 1 if parents, grandparents or siblings attended ISU
isu_fan	42.9%	49.5%	= 1 if was a Cyclone fun before attending ISU
N	471		Number of observations

Notes: ISU = Iowa State University

Table A2. Simulated maximum likelihood estimation of seemingly unrelated model in (2) but including the price offered previously. Dependent variable: Binary choice of purchasing ticket of sport j, B^j, with j = 1,2,…,8.

	1	2	3	4	5	6	7	8
Dependent variable	womenbb	menbb	football	volleyball	wrestling	gym	hockey	soccer
own price	-0.114***	-0.098***	-0.078***	-0.179***	-0.138***	-0.132***	-0.166***	-0.118***
gender	-0.543***	0.109	-0.081	-0.294**	-0.137	-0.675***	-0.209	-0.267*
hi_sport	0.181	0.016	-0.035	0.164	0.632***	0.035	-0.030	0.123
hi_inter	0.057	0.067	0.226	-0.004	-0.227	-0.105	0.034	0.048
isu_inter	0.235	0.287*	0.461***	-0.017	0.314**	-0.067	0.112	0.058
isu_family	0.008	-0.004	-0.058	-0.114	0.008	0.018	-0.188	-0.052
isu_fan	0.492***	0.424***	0.456***	0.265	0.336**	0.127	0.118	0.309**
price previously offered	-	0.057**	0.024	-0.002	-0.019	0.029	-0.003	0.011
constant	0.587***	0.664**	0.984**	0.763**	0.178	0.884***	1.209***	0.220

N=449 - Simulation draws=400- Loglikelihood= 1,919 - Iterations= 34

Table A3 Median unconditional reservation prices of sport j for individuals that buy sport jk.

	k	1	2	3	4	5	6	7	8
j	Labels	womenbb	menbb	football	voleyball	wrestling	gymnastics	hockey	soccer
1	womenbb		8.13	8.13	8.52	8.51	8.51	8.13	8.13
2	menbb	14.79		14.81	14.58	14.79	14.58	14.58	14.74
3	football	24.05	24.46		23.57	24.30	23.12	23.45	23.33
4	voleyball	4.67	4.67	4.36		4.67	4.67	4.32	4.36
5	wrestling	6.35	6.35	6.35	5.80		5.80	6.35	5.76
6	gymnastics	7.44	6.92	7.38	7.38	7.38		7.44	7.63
7	hockey	6.98	6.61	6.74	6.76	6.98	6.98		6.76
8	soccer	3.87	3.76	3.76	3.87	3.81	3.76	3.75	

Note: Main Diagonal elements are 0 because they have no meaning.

Table A4. Median conditional reservation prices for individuals that buy the combination of sport jk.

	k	1	2	3	4	5	6	7	8
j	Labels	womenbb	menbb	football	voleyball	wrestling	gymnastics	hockey	soccer
1	womenbb	7.08	11.42	10.52	12.64	11.09	11.99	10.65	11.50
2	menbb	20.81	14.46	20.30	20.54	19.19	20.32	19.52	19.69
3	football	29.29	29.39	23.33	28.23	27.33	27.94	27.82	27.15
4	voleyball	8.66	8.40	7.59	4.07	7.88	8.34	8.06	8.50
5	wrestling	9.80	9.10	8.92	9.89	5.17	9.71	9.18	9.06
6	gymnastics	10.75	10.52	9.79	10.97	10.09	3.95	10.89	10.15
7	hockey	9.67	9.20	8.93	9.92	8.63	10.17	6.72	9.78
8	soccer	9.92	9.81	9.13	10.45	9.26	10.26	10.13	3.42

Note: Main Diagonal is filled with the median unconditional reservation price for all the sample (not only those who bought sports jk as in table A3).

APPENDIX B

We will prove that the choice of 0.50 probability of purchasing a ticket yield predicted or reservation prices ($\tilde{p}_i^m$ and $\tilde{p}_i^w$) that are mathematically equal to the willingness to pay (W_i^m and W_i^w) calculated following a contingent valuation approach.

Following Cameron (1988) we can use the contingent valuation approach to directly find the willingness to pay for each sport and for each individual. Suppose that the willingness to pay for a men's college-basketball ticket (W_i^{m*}) and a women's college-basketball ticket (W_i^{w*}) can be represented by the following system of latent variables:

$$\begin{aligned} W_i^{m*} &= \alpha_0^m + Z_i\gamma^m - \varepsilon_i^m \\ W_i^{w*} &= \alpha_0^w + Z_i\gamma^w - \varepsilon_i^w \end{aligned} \tag{A1}$$

where α_0^x and γ^x, x={m,w}, are parameters to be estimated, Z_i is a vector of observed individual characteristics and ε_i^x is an error term jointly normal distributed representing unobservable factors:

$$\begin{pmatrix} \varepsilon_i^m \\ \varepsilon_i^w \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^{2\,m} & \rho\sigma^m\sigma^w \\ \rho\sigma^m\sigma^w & \sigma^{2\,w} \end{pmatrix}\right] \tag{A2}$$

But instead of observing the latent variable, we observe a pair binary variables, B_i^m and B_i^w, such that

$$\begin{aligned} B_i^m &= \begin{cases} 1 & if\ W_i^{m*} > P_i^m \\ 0 & if\ W_i^{w*} \leq P_i^m \end{cases} \\ B_i^w &= \begin{cases} 1 & if\ W_i^{m*} > P_i^w \\ 0 & if\ W_i^{w*} \leq P_i^w \end{cases} \end{aligned} \tag{A3}$$

where B_i^m and B_i^w are the dummy variables that equal 1 if individual i purchases the ticket at the offered random price and zero otherwise; P_i^m and P_i^w are the random prices offered to each individual for men's basketball and women's basketball tickets respectively.

By plugging equation (A1) into (A3) we obtain

$$\begin{aligned} B_i^m &= \begin{cases} 1 & if\quad \varepsilon_i^m < \alpha_0^m - P_i^m + Z_i\gamma^m \\ 0 & if\quad \varepsilon_i^m \geq \alpha_0^m - P_i^m + Z_i\gamma^m \end{cases} \\ B_i^w &= \begin{cases} 1 & if\quad \varepsilon_i^w < \alpha_0^w - P_i^w + Z_i\gamma^w \\ 0 & if\quad \varepsilon_i^w \geq \alpha_0^w - P_i^w + Z_i\gamma^w \end{cases} \end{aligned} \tag{A4}$$

This constitutes a bivariate probit model; we use maximum likelihood procedures to estimate parameters α_0^x, γ^x, σ^x, and ρ, x = {m,w}, given our data on the binary variables B_i^m and B_i^w, on the random prices P_i^m and P_i^w, on Z_i and our assumption on error terms. Note that in this model it is required to have a different price offered to each individual in order to identify all the parameters we want to estimate. Manipulating this expression and calculating the joint probabilities Pr[B_i^m=1, B_i^w=1], Pr[B_i^m=1, B_i^w=0], Pr[B_i^m=0, B_i^w=1], Pr[B_i^m=0, B_i^w=0], we obtain the following log-likelihood function

$$L(B_i^m, B_i^w / P_i^m, P_i^w, Z_i, \alpha_0^m, \alpha_0^w, \gamma^m, \gamma^w, \sigma^m, \sigma^w, \rho) = \sum_{i=1}^{i=n} \log\{\Phi_2\, [q_i^m \left(\frac{\alpha_0^m}{\sigma^m} - \frac{1}{\sigma^m} P_i^m + Z_i \frac{\gamma'^m}{\sigma^m}\right), q_i^w \left(\frac{\alpha_0^w}{\sigma^w} - \frac{1}{\sigma^w} P_i^w + Z_i \frac{\gamma'^w}{\sigma^w}\right), q_i^m q_i^w \rho]\} \tag{A5}$$

where Φ_2 is the bivariate standard normal cumulative distribution function, $q_i^m = 2B_i^m - 1$ and $q_i^w = 2B_i^w - 1$. The fact that each individual is offered a different random price allows us to identify the coefficient of P_i^m and P_i^w, which turns out to be the negative reciprocal of the standard deviation of the disturbances distributed jointly normal by assumption. If we define, for x = {m,w},

$$\tilde{\alpha}_0^x = \frac{\alpha_0^x}{\sigma^x}$$

$$\tilde{\alpha}_1^x = -\frac{1}{\sigma^x} \tag{A6}$$

$$\tilde{\gamma}'^x = \frac{\gamma'^x}{\sigma^x}$$

and rearrange terms, the likelihood in (A5) can be written as

$$L(B_i^m, B_i^w / P_i^m, P_i^w, Z_i, \tilde{\boldsymbol{\alpha}}^m, \tilde{\boldsymbol{\alpha}}^w, \tilde{\gamma}^m, \tilde{\gamma}^w, \rho) = \sum_{i=1}^{i=n} \log\{\Phi_2\, [q_i^m (\tilde{\alpha}_0^m + \tilde{\alpha}_1^m P_i^m + Z_i \tilde{\gamma}'^m), q_i^w (\tilde{\alpha}_0^w + \tilde{\alpha}_1^w P_i^w + Z_i \tilde{\gamma}'^w), q_i^m q_i^w \rho]\} \tag{A7}$$

which is the same likelihood function obtained in equation (4). Therefore our estimates of α_0^x and γ'^x, for x = {m,w}, can be transformed according to (A6) to yield the same estimates obtained from the likelihood of (4).

Now, suppose that we calculate the model in (A.1) with the estimated parameters, yielding the following estimated willingness to pay for men's and women's basketball for individual i:

$$\breve{W}_i^m = \breve{\alpha}_0^m + Z_i \breve{\gamma}^m$$

$$\breve{W}_i^w = \breve{\alpha}_0^w + Z_i \breve{\gamma}^w \tag{A8}$$

where "˘"denotes the estimated parameters using the contingent valuation approach. We need to show that the reservation prices $\tilde{p}_i^m$ and $\tilde{p}_i^w$ have a similar form as (A8).

Given that the likelihood in (A7) is the same as the likelihood in (4) we can argue, for x = $\{m,w\}$, the following

$$\widetilde{\breve{\alpha}}_0^x = \frac{\breve{\alpha}_0^x}{\breve{\sigma}^x} = \hat{\beta}_0^x \tag{A9.a}$$

$$\widetilde{\breve{\alpha}}_1^x = -\frac{1}{\breve{\sigma}^x} = \hat{\beta}_1^x \tag{A9.b}$$

$$\widetilde{\breve{\gamma}}^x = \frac{\breve{\gamma}'^x}{\breve{\sigma}^x} = \hat{\delta}'^x \tag{A9.c}$$

that is, the estimated transformed parameters in (A9) are equal to the estimated parameters yielded by likelihood in (4).

According to equation (6) the reservation prices at the 50% probability of purchasing a ticket implicitly solve the following equation:

$$\begin{gathered} \hat{\beta}_0^m + \hat{\beta}_1^m \tilde{p}_i^m + Z_i \hat{\delta}^m = 0 \\ \hat{\beta}_0^w + \hat{\beta}_1^w \tilde{p}_i^w + Z_i \hat{\delta}^w = 0 \end{gathered} \tag{A10}$$

Solving for $\tilde{p}_i^m$ and $\tilde{p}_i^w$ we obtain

$$\begin{gathered} \tilde{p}_i^m = -\frac{1}{\hat{\beta}_1^m}(\hat{\beta}_0^m + Z_i \hat{\delta}^m) \\ \tilde{p}_i^w = -\frac{1}{\hat{\beta}_1^w}(\hat{\beta}_0^w + Z_i \hat{\delta}^w) \end{gathered} \tag{A11}$$

or

$$\begin{gathered} \tilde{p}_i^m = \left(-\frac{\hat{\beta}_0^m}{\hat{\beta}_1^m}\right) + Z_i\left(-\frac{\hat{\delta}^m}{\hat{\beta}_1^m}\right) \\ \tilde{p}_i^w = \left(-\frac{\hat{\beta}_0^w}{\hat{\beta}_1^w}\right) + Z_i\left(-\frac{\hat{\delta}^w}{\hat{\beta}_1^w}\right) \end{gathered} \tag{A12}$$

Substituting (A9.b) into (A12) we obtain

$$\begin{gathered} \tilde{p}_i^m = \left(\breve{\sigma}^m \hat{\beta}_0^m\right) + Z_i\left(\breve{\sigma}^m\, \hat{\delta}^m\right) \\ \tilde{p}_i^w = \left(\breve{\sigma}^w \hat{\beta}_0^w\right) + Z_i\left(\breve{\sigma}^w\, \hat{\delta}^w\right) \end{gathered} \tag{A13}$$

Finally plugging (A9.a) and (A9.c) into (A13) we get our desired result, which completes the proof:

$$\tilde{p}_i^m = \breve{\alpha}_0^m + Z_i\breve{\gamma}^m$$
$$\tilde{p}_i^w = \breve{\alpha}_0^w + Z_i\breve{\gamma}^w \quad \text{(A14)}$$

Even though we can show that the correct probability of purchase to use in the analysis is 0.5, our conclusions are not sensitive to that choice. Because purchase probability is monotonic in price, both reservation price series decrease in purchase probability (Figure A1). So our conclusions regarding relative strength of demand for men's to women's basketball for men and women remain unaffected by choice of purchase probability.

Figure A1. Median reservation prices for men's and women's basketball at different estimated probability of purchase.