

**MODELAGEM DA DEMANDA COMBINANDO DADOS DE  
PREFERÊNCIA DECLARADA E REVELADA**

**DEMAND MODEL COMBINING STATED AND REVEALED  
PREFERENCE DATA**

**Luciana Londero Brandli**

Engenheira Civil, Professora Adjunta

UNIJUI/UPF

Departamento de Tecnologia/FEAR

Rua São Francisco, 501. Caixa postal 560. CEP 98700-000, Bairro São Geraldo

Tel. 55 3332 0512, brandli@unijui.tche.br

**Luiz Fernando M. Heineck, Dr**

Engenheiro Civil, Professor Titular

UFSC

Engenharia de Produção

Campus Universitário, Bairro Trindade, Florianópolis, CEP 88040 800

Tel. 48 331 7045, Freitas8@terra.com.br

**RESUMO**

As técnicas de preferência revelada e declarada têm contribuído muito para o desenvolvimento da literatura econométrica na tentativa de determinar as variáveis que influenciam a decisão do indivíduo num processo de escolha. Este artigo combina dados de preferência, com o objetivo de obter as vantagens da complementaridade das forças e franquezas de ambas os tipos de dados. A abordagem envolve a estimação de um modelo apenas com dados de PR, apenas com dados de PD e combinando dados de PR e PD. A aplicação é feita no mercado habitacional, onde se observa, através da literatura, que a maioria dos trabalhos de escolha do consumidor têm se restringido a apenas uma das abordagens. As funções utilidades obtidas mostram a importância relativa do atributo, a tendência de

comportamento através dos sinais e o seu nível de significância. A análise dos resultados dos modelos indica diferenças e similaridades entre o comportamento dos atributos.

**Palavras-chave:** preferência revelada, preferência declarada, mercado habitacional.

## ABSTRACT

The revealed and stated preference methods have been contributing a lot for the development of the econometric literature in the attempt of determining the variables that influence the individual decision in a choice process. This article combines preference data, with the objective of obtaining the advantages of the complementarity of the forces and frankness of both types of data. The approach involves the estimate of a model only with RP data, only with SP data and combining RP and SP data. The application is in the housing market, where it is observed, through the literature, that most of the papers of the consumer's choice has restricted the only one approaches. The utility functions obtained show the relative importance of the attributes, the tendency of behavior through the signs and its significance statistical. The results analysis of the models indicates differences and similarities about the attribute's behavior.

**Key-words:** revealed preference, stated preference, housing market.

## 1. INTRODUÇÃO

Os modelos econométricos de escolha discreta envolvem as medidas das relações econômicas e têm como ocupação principal o comportamento individual, na tentativa de determinar as variáveis que influenciam a decisão do indivíduo num processo de escolha.

As técnicas de preferência revelada (PR) e declarada (PD) têm contribuído em muito para o desenvolvimento da literatura econométrica, na qual, na grande maioria das vezes, são usadas de forma independente para a estimação de uma variedade de aplicações de escolha discreta. Existe, no entanto, um crescente número de estudos que exploram ambos os tipos de dados em um mesmo modelo, na perspectiva de complementaridade e melhora da capacidade de informação dos modelos. De acordo com Louviere, Hensher e Swait (2000), muitos destes

estudos podem ser encontrados nas áreas de transporte, marketing e análise ambiental durante a última década.

Para o caso específico da habitação, a maioria dos trabalhos de escolha do consumidor restringe-se a uma das abordagens (ver TIMMERMANS; MOLIN; NOORTWIJK, 1994). Uma aplicação da análise combinada é encontrada nos trabalhos de Earnhart (2001; 2002), que examina as escolhas habitacionais de indivíduos através de dados reais de aquisição de imóveis e de possibilidades de compras de imóveis hipotéticos.

Este artigo combina dados de preferência declarada e revelada com o objetivo de obter as vantagens da complementaridade das forças e fraquezas de ambas os tipos de dados. Para tal, a abordagem envolve a estimação de um modelo apenas com dados de PR; a estimação de um modelo apenas com dados de PD; e a estimação seqüencial com dados de PR e PD. O objeto de estudo é o mercado habitacional para estudantes universitários.

O maior requisito das pesquisas PR e PD é a estimação de uma função utilidade, na qual pode-se prever a demanda de cada indivíduo por cada alternativa de escolha dentro de um conjunto de alternativas relevantes que caracterizam aquele bem. Estas escolhas podem ser atuais ou hipotéticas conforme descreve a seção seguinte.

O artigo está estruturado como segue. A partir desta introdução, apresenta-se uma breve caracterização sobre os dois tipos de dados de preferência e sobre o modelo conjunto. Na seção seguinte, descreve-se a estrutura metodológica da estimação combinada. É então apresentado um estudo aplicado à demanda habitacional de jovens que saem da casa dos pais para estudar, utilizando dados de PR e PD. O artigo finaliza com algumas considerações sobre os modelos gerados, as diferenças e similaridades de cada um.

## 2. DADOS DE PREFERÊNCIA E SUAS FORMAS

Os dados de preferência podem ser de duas formas: dados de preferência revelada e dados de preferência declarada. A técnica da preferência revelada é baseada nas observações das escolhas reais do indivíduo. As respostas usualmente configuram escolhas simples ou diferentes escolhas ao longo de um determinado período de tempo (BRADLEY, 1994). Classificada por Adamowicz, Louviere e Williams (1994) como método indireto, a limitação desta forma de obtenção das informações refere-se à identificação das decisões que os usuários tomariam caso defrontassem com situações que não tenham vivenciado no passado (SENNA; TONI; LINDAU, 1994). Para Kroes e Sheldon (1988), pode ser difícil de se obter variação suficiente nos dados de PR para examinar todas as variáveis de interesse. É a partir dessas limitações que se assume a importância do papel da preferência declarada.

O uso da técnica da preferência declarada é frequentemente denominado de análise experimental, escolha declarada, conjoint analysis ou análise de *trade-off*. Esta técnica se insere nos métodos chamados diretos (ADAMOWICZ; LOUVIERE; WILLIAMS, 1994). Seu princípio básico é apresentar ao entrevistado um conjunto de opções hipotéticas das quais ele escolhe uma. Esta escolha feita pelo indivíduo representa a sua preferência pelos atributos de uma alternativa sobre as outras (ADAMOWICZ; LOUVIERE; WILLIAMS, 1994).

Senna, Toni e Lindau (1994) afirmam que a técnica da preferência declarada pode ser vista como o método para identificação das preferências dos usuários ou possíveis ações em resposta às mudanças nas condições da oferta. Como explicam Ben-Akiva e Morikawa (1990), é um modelo de escolha que representa a decisão entre mudar para uma nova alternativa ou manter a escolha existente. Na visão de Hensher (1994), o estudo da PD é um resultado potencial, ou seja, o resultado com maior probabilidade de ocorrência dada uma combinação de atributos.

A crítica comum a esta técnica é a natureza hipotética das questões e conseqüentemente, das escolhas das pessoas (BROWNSTONE; BUNCH; TRAIN, 2000; EARNHART, 2002), que podem incorrer em respostas não representativas do comportamento que as pessoas teriam caso a escolha fosse real.

### **3. MODELO COMBINANDO DADOS DE PREFERÊNCIA**

Na literatura econométrica encontra-se um número cada vez maior de estudos que estimam modelos combinando dados de preferência declarada e revelada. Além disto, existem muitos trabalhos que têm utilizado o enfoque híbrido, incluindo questões na mesma pesquisa que contemplam ambas as técnicas de preferência (HENSHER, 1993 apud SENNA; TONI; LINDAU, 1994). No entanto, esta abordagem não necessariamente contempla a análise combinada dos dados apresentada neste artigo, a partir da estimação conjunta dos parâmetros de PR e PD, e fundamentada por Morikawa (1989), Ben-Akiva e Morikawa (1990), Hensher (1992) e Bradley e Daly (1991; 1994).

Exemplos desses estudos podem ser encontrados em Wardman (1998), Souza (1999) e Souza et al. (2000), que apresentam métodos para comparação de dados de PR e PD. Wardman (1988) compara os resultados do modelo da PR e da PD para avaliação do tempo de transporte para diferentes segmentos da amostra. Souza (1999), por sua vez, determina a função utilidade de supermercados na visão de seus clientes potenciais usando dados de PR e PD. A integração dos dois tipos de dados é feita pelo autor confrontando os resultados das duas funções utilidades originadas de cada tipo respectivamente. O autor repete esta abordagem na modelagem da escolha portuária (SOUZA et al., 2000).

A principal defesa da abordagem conjunta refere-se ao fato de que estes dois tipos de dados (preferência revelada e declarada) possuem diferentes vantagens e desvantagens. Como comenta Morikawa (1989), os dados de preferência declarada podem não ser válidos para

previsões, mas podem ajudar a identificar as preferências subentendidas que determinam a escolha atual. Por outro lado, embora os dados de preferência revelada representem o comportamento atual, eles podem não conter informações que auxiliem no entendimento das preferências do consumidor.

Assim, a análise que explora ambos os tipos de dados simultaneamente pode usá-los de forma compensatória, melhorando a qualidade dos resultados finais (MORIKAWA, 1989; BRADLEY; DALY, 1991). Isto resultaria, segundo Earnhart (2002), num modelo econométrico com maior poder explanatório.

Além disto, como esta análise pode não apresentar similaridades no processo de decisão, mas mostrar diferenças entre a preferência revelada e a declarada, ela permite que sejam indicados quais os parâmetros similares no processo e quais os que fazem com que o comportamento atual difira das intenções declaradas (EARNHART, 2002). Isto possibilita ao pesquisador um maior conhecimento sobre as variações nas respostas escolhidas e nos níveis dos atributos das alternativas de escolha (HENSHER, 1994).

#### **4. ESTRUTURA METODOLÓGICA DOS MODELOS COM DADOS COMBINADOS**

Para a combinação dos dois tipos de dados é necessário que ambas as abordagens reflitam o mesmo processo de escolha baseado em atributos. Isto implica que os modelos devem representar os mesmos *trade-offs* (ADAMOWICZ; LOUVIERE; WILLIAMS, 1994).

Inicialmente, vamos apresentar o entendimento da função utilidade. A seleção de uma alternativa sobre as outras implica que a utilidade  $U_i$  de um objeto é maior que a utilidade  $U_j$  de outro. A função utilidade é aleatória uma vez que pode apenas analisar a probabilidade de escolha de uma alternativa sobre outra.

Sendo as funções utilidades associadas respectivamente às alternativas  $i$  e  $j$ :

$$U_i = V_i + \xi_i \quad (1)$$

$$U_j = V_j + \xi_j \quad (2)$$

onde:

$V_i$ ,  $V_j$  são os componentes determinísticos.

$\xi_i$ ,  $\xi_j$  são os componentes de erro randômico. Representam uma parcela aleatória ou desconhecida da função utilidade que captura a dispersão das escolhas e os fatores não controláveis ou desconhecidos do pesquisador tais como atributos relevantes não considerados, erros de medição dos atributos, erros de percepção do entrevistado e diferenças não observáveis na avaliação (BOVY; STERN, 1990 apud SCHMITZ, 2001).

A probabilidade de escolha de uma alternativa sobre outra é:

$$\Pr\{i\} = \Pr \{V_i + \xi_i > V_j + \xi_j ; \forall i \in C_n\} \quad (3)$$

onde:

$C_n$  é o conjunto de escolhas de  $n$  indivíduos.

Assumindo que os erros têm uma distribuição de Gumbel com fator escalar  $\mu$ , a probabilidade de escolha para a especificação do Modelo Logit Multinomial (MLM) é:

$$\Pr\{i\} = \frac{\exp^{\mu V_{in}}}{\sum_{i \in C_n} \exp^{\mu V_{in}}} \quad (4)$$

A função utilidade também pode ser representada pela expressão:

$$U = \beta x + \alpha w + \xi \quad (5)$$

Logo:

$$\Pr\{i\} = \frac{\exp^{\mu(\beta x + \alpha w)}}{\sum_{i \in C_n} \exp^{\mu(\beta x + \alpha w)}} = \frac{\exp^{\mu(V)}}{\sum_{i \in C_n} \exp^{\mu(V)}} \quad (6)$$

onde:

$\mu$  é o fator escalar

$\beta$ ;  $\alpha$  são os parâmetros, coeficientes das variáveis

$x$ ;  $w$  são as variáveis que influenciam a escolha

O fator escalar desempenha um papel crucial no processo de combinação de dados de preferência, por isto é importante entender a sua função no MLM. A expressão (6) mostra que o fator escalar e os parâmetros de um modelo de escolha são multiplicados e inseparados. Então, não é possível identificar o fator escalar dentro de um conjunto de dados em particular, mesmo sabendo-se que o fator escalar afeta o valor dos parâmetros, de tal forma que quanto maior a escala maior os parâmetros.

Como esta escala ( $\mu$ ) e os parâmetros da utilidade ( $\beta$ ;  $\alpha$ ) são confundidos e não podem ser separados, o fator escalar normalmente é considerado igual a 1 no MLM.

No entanto, quando existem dois tipos de dados, seu valor deverá ser estimado, pois na comparação de dois conjuntos de dados não se pode determinar se as diferenças observadas na função utilidade de cada um são resultado das diferenças na escala, nos parâmetros ou em ambos.

De fato, se os dois conjuntos de dados refletem a mesma função utilidade (mesmos atributos), mas tem diferentes fatores escalares,  $\mu_1$  e  $\mu_2$ , os parâmetros estimados vão diferir nos dois modelos gerados (por exemplo,  $\mu_1\beta_1$  e  $\mu_2\beta_2$ )

Na combinação de dados PR e PD, a questão é se  $\mu_1\beta_1 = \mu_2\beta_2$ . Assim,  $\beta_1 = (\mu_2/\mu_1)\beta_2$  e o gráfico de um vetor parâmetro de PR sobre outro de PD seria uma reta passando pela origem, com inclinação igual a  $\mu_2/\mu_1$ , onde  $\mu_2$  refere-se aos parâmetros de PD e  $\mu_1$  aos parâmetros de PR, como mostra a Figura 1.

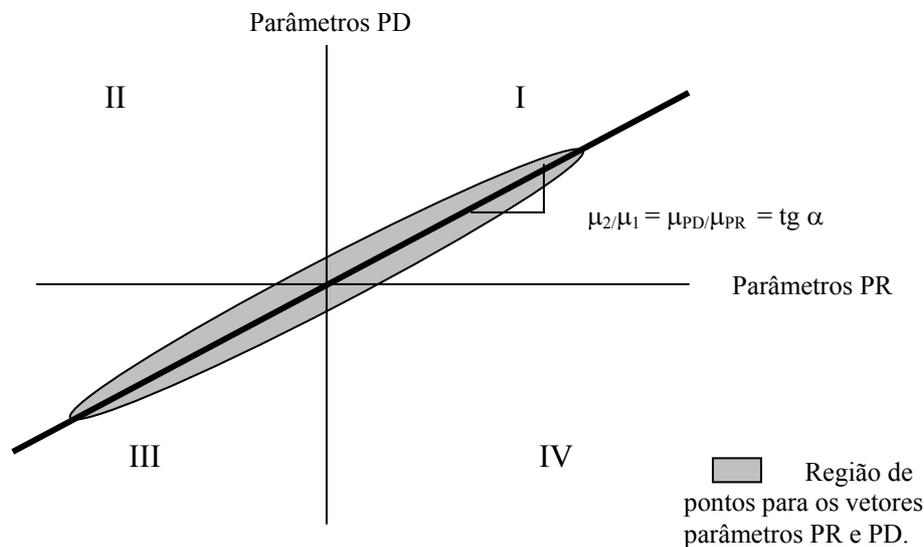


Figura 1: Teste visual para igualdade dos vetores parâmetros de dois conjuntos de dados

**Fonte:** Louviere, Hensher e Swait (2000, p.235)

Este gráfico serve para identificar visualmente a igualdade dos parâmetros entre os dados. Por exemplo, se o parâmetro for positivo para ambos os dados PR e PD ele será plotado no quadrante I, se for negativo em ambos, será plotado no quadrante III, e se obtiver sinais opostos em ambos os modelos será plotado nos quadrantes II ou IV.

O fator escalar no MLM é inversamente relacionado com a variância do termo de erro:

$$\text{Var}(\xi)^2 = \pi^2/6\mu^2 \quad (7)$$

A estrutura comportamental representada pela função utilidade é aplicável a ambos os tipos de dados de preferência. A resposta de cada indivíduo - se atual ou hipotética - é assumida por ter uma função utilidade (BRADLEY; KROES, 1990). A abordagem usual da análise conjunta tem sido testar se os coeficientes da utilidade dos modelos de escolha da PR e PD são proporcionais (SWAIT; LOUVIERE; WILLIAMS, 1994).

A função utilidade na preferência revelada é:

$$U^{PR} = \beta x^{PR} + \alpha w^{PR} + \xi^{PR} \quad (8)$$

Onde:

$x^{PR}$ ,  $w$  - vetores das variáveis medidas que influenciam a decisão na PR.

$\beta$ ,  $\alpha$  - vetores ou parâmetros desconhecidos (devem ser estimados).

$\xi^{PR}$  - representa a soma dos componentes de utilidade não medidos. Efeitos não observados.

Similarmente, o processo de decisão na PD pode ser medido pela função utilidade:

$$U^{PD} = \beta x^{PD} + \gamma z^{PD} + \xi^{PD} \quad (9)$$

Onde:

$x^{PD}$ ,  $z$ - vetores das variáveis medidas que influenciam a decisão na PD.

$\beta, \gamma$ - vetores ou parâmetros desconhecidos (devem ser estimados).

$\xi^{PD}$  - representa a soma dos componentes de utilidade não medidos. Efeitos não observados.

Diante das duas funções utilidade, para cada tipo de dados respectivamente, pode-se afirmar que:

- observa-se a existência de variáveis que ocorrem apenas na PR ( $w$ ) e apenas na PD ( $z$ );
- a contribuição marginal para a utilidade indicada por  $\beta x$  é a mesma nas duas expressões;
- os modelos de escolha discreta assumem que os componentes aleatórios da função utilidade são Independentemente e Identicamente Distribuídos (IID), entre as alternativas e indivíduos. No modelo combinado PR+PD assume-se que os termos aleatórios têm esta propriedade dentro de cada grupo de dados;
- os termos aleatórios  $\xi^{PR}$  e  $\xi^{PD}$  não são iguais devido a natureza dos experimentos a que estão associadas. Bradley e Daly (1991) afirmam que não se pode dizer qual

dos dois componentes de erro tem maior variância e que não é possível assumir que suas variâncias sejam iguais.

A relação entre as variâncias dos termos de erro,  $\xi^{PR}$  e  $\xi^{PD}$ , é dada pela expressão:

$$\theta^2 = \frac{\text{Var}(\xi^{PR})}{\text{Var}(\xi^{PD})} \quad (10)$$

Onde  $\theta$ :

- é denominado parâmetro escalar e representa a diferença de escala entre os dois modelos, refletindo os respectivos níveis de variação entre os dois tipos de dados (BEN-AKIVA; LERMAN, 1985);
- é um parâmetro escalar positivo que pode assumir valores entre 0 e 1; quanto mais próximo de 1 mais similares são os valores das variâncias de PR e PD (BEN-AKIVA; MORIKAWA, 1990);
- com os dois tipos de dados é possível estimar o valor de  $\theta$  para uma das séries, juntamente com outros parâmetros, como o valor resultante da probabilidade máxima (SWAIT; LOUVIERE, 1993).

Assim, para usar ambos os dados de PR e PD, é necessário multiplicar a equação (9) pela variável aleatória  $\theta$  a fim de igualar as variâncias dos dois métodos.

Este procedimento garante o uso de ambos os métodos na estimação logit, os quais requerem dados com variância igual sobre as observações.

Relacionando as expressões (7) e (10) tem-se:

$$\frac{\text{Var}\xi^{PR}}{\text{Var}\xi^{PD}} = \frac{\pi^2/6\mu_{PR}^2}{\pi^2/6\mu_{PD}^2} = \left(\frac{\mu_{PD}}{\mu_{PR}}\right)^2 = \theta^2 \quad (11)$$

É convencional assumir que o valor do fator escalar dos dados PR é um ( $\mu_{PR}=1$ ).

Assim,  $\theta^2 = \mu_{PD}^2$ , o que significa que o valor do fator escalar ( $\mu$ ) pode ser confundido com o

valor do parâmetro escalar ( $\theta$ ). Estimar o valor do fator escalar de PD representa estimar a escalar relativa com respeito aos dados PR.

No gráfico da Figura 1,  $\tan \alpha = \mu_{PD}/\mu_{PR}$ . Como  $\mu_{PD}/\mu_{PR} = \theta$ , pode-se concluir que a tangente da reta, que passa pela origem e é formada pelas coordenadas dos parâmetros PR e PD é igual ao parâmetro escalar de ligação dos modelos.

Várias formas de estimação combinada dos dados são possíveis (LOUVIERE, HENSHER; SWAIT, 2000). Entre estas formas destacam-se o método seqüencial, descrito por Ben-Akiva e Morikawa (1990) e utilizado neste trabalho, o método simultâneo proposto por Bradley e Daly (1991) e o método que se utiliza da visualização gráfica dos parâmetros proposto por Louviere, Hensher e Swait (2000), também usado aqui.

O método seqüencial estima todos os parâmetros calibrando os modelos PR e PD seqüencialmente. Os passos deste método são apresentados com base em Morikawa (1989) e Ortúzar (2000) e foram implementados no software LMPC (Souza, 1999), adotado no processamento dos dados:

1) Primeiramente, estima-se os coeficientes PD para obter os valores de  $\beta$  e  $\gamma$ ; esses coeficientes, no caso das variáveis  $x$ , são entendidos como a  $\theta\beta$ . Assim substituí-se na equação de PR para formar:

$$V^{PR} = \theta\beta x^{PR} \quad (12)$$

onde:

$V^{PR}$  - é a variável que faz a ligação entre os dois modelos. Ela será chamada de variável composta e é formada pelos atributos comuns aos dois.

2) Com essa nova variável  $V^{PR}$ , junta-se as outras variáveis PR e forma-se a equação:

$$U^{PR} = \lambda V^{PR} + \alpha w^{PR} + \xi^{PR} \quad (13)$$

onde:

$\lambda$  - parâmetro do atributo composto

3) Estima-se os coeficientes de  $U_i^{PR}$  com os dados de PR e obtém-se o valor de  $\lambda$ .

Assim, para obter  $\beta$ , partimos da relação:  $\lambda V^{PR} = \lambda \theta \beta x^{PR}$  e como  $\theta \beta$  é o coeficiente de PD e possui (por teoria) maior variabilidade que PR, então  $\lambda$  é um valor maior do que 1 e inverso ao valor de  $\theta$ ,  $\lambda = \frac{1}{\theta}$ . O produto  $\lambda \theta = 1$  ou seja  $\lambda \theta \beta = \beta$  que é o coeficiente real.

Para corrigir os demais coeficientes basta multiplicar por  $\lambda$  ou dividir por  $\theta$  e assim obter o conjunto de dados secundários para a PD. Então, os coeficientes reais são:

$$U_i = \lambda \theta \beta_i X_i^{PR/PD} + \alpha_j \omega_j^{PR} + \lambda \theta \gamma_k Z_k^{PD} \quad (14)$$

Lembrando que  $\lambda \theta = 1$

A equação estatística final do modelo será composta pelos coeficientes estimados com dados de PR, mais os coeficientes ajustados com dados de PD, corrigidos pelo inverso do coeficiente do atributo composto.

## 5. APLICAÇÃO EMPÍRICA

A pesquisa empírica teve como objeto de estudo o mercado habitacional. O levantamento de dados foi realizado com estudantes da Universidade Regional do Noroeste do Estado do Rio Grande do Sul (UNIJUI). A amostra que participou da pesquisa foi constituída por 450 estudantes.

Quanto às características socioeconômicas dos entrevistados, 58,4% dos estudantes eram mulheres e 41,6% homens; 34,8% tinham até 20 anos, 45,5% entre 21 a 24 anos, 15,3% entre 25 a 29 anos e os restantes 4,4% tinham 30 anos ou mais; o estado civil de 92% era solteiro e apenas 8% eram casados. A grande maioria dos estudantes informou que mantém dependência financeira dos pais, correspondendo a 82% do total. Os demais 18% eram independentes e possuíam renda própria.

Entre os estudantes, 335 haviam saído da casa dos pais para estudar ingressando, assim, no mercado habitacional. Os demais continuavam morando na casa dos pais.

### **5.1 Método de coleta dos dados**

Os dados de PR e PD foram coletados dos mesmos estudantes. No entanto, na pesquisa PR interessaram apenas os estudantes que caracterizavam uma demanda efetiva no mercado habitacional, porque não moravam mais com os pais. Assim, 335 estudantes responderam a pesquisa PR e todos os 450 responderam a pesquisa PD. Na pesquisa PD, os estudantes foram solicitados a considerar duas hipóteses de acordo com sua situação: (a) considere que você vai se mudar de sua atual moradia; (b) considere que você vai sair da casa de seus pais e precisa escolher uma moradia. Em ambos os casos, os estudantes deveriam considerar sua situação financeira no momento e então indicar qual a situação de moradia ele provavelmente escolheria.

Os atributos utilizados para caracterizar a escolha da habitação estão apresentados no Quadro 1. A definição dos mesmos partiu de um estudo exploratório local (BRANDLI; HEINECK, 2003) e de uma revisão bibliográfica sobre demanda habitacional de jovens (BRANDLI, 2004).

O projeto do experimento PD indicou um fatorial completo correspondente a  $2^6 = 64$  combinações dos atributos e seus níveis. Na prática, este número é impossível de avaliar, ou seja, o entrevistado teria dificuldade em escolher uma opção entre as 64 através da compensação entre os atributos (como a técnica PD prescreve). A solução utilizada foi o fatorial fracionário, o qual assume que certos efeitos de interação entre os atributos não são estatisticamente significantes (LOUVIERE; HENSHER; SWAIT, 2000). Assim, a técnica do fracionamento indicou um conjunto de 32 alternativas. Estas alternativas foram então

divididas em 8 conjuntos de 4 alternativas cada e foram submetidos aos entrevistados (FINNEY, 1945 apud SOUZA, 1999, p.151).

Quanto a pesquisa PR, procurou-se reportar os níveis dos atributos os mesmos níveis usados no experimento PD. Assim, na pesquisa PR, os entrevistados relataram sua atual situação de moradia com relação a cada atributo de PD. As entrevistas que apresentaram como resposta níveis diferentes (caso da localização) foram desconsideradas da estimação PR apresentada neste artigo. Assim restaram 281 entrevistas.

Quadro 1: Atributos e níveis do experimento PD

ATRIBUTO	NÍVEIS	DESCRIÇÃO
Qualidade	0	A moradia tem condições menos satisfatórias quanto ao conforto acústico, iluminação e insolação.
	1	A moradia tem boas condições quanto ao conforto acústico, iluminação e insolação.
Arranjo	0	Em grupo (amigos/familiares)
	1	Sozinho
Aluguel	0	150,00 (valores de 0 a 150,00)
	1	300,00 (valores de 150 a 300,00)
Localização/acesso	0	Centro
	1	Próximo da universidade (campus/sede)
Existência de Móvel	0	Sem móvel e equipamentos
	1	Com móvel e equipamentos
Conforto/Uso	0	Espaço por pessoa é restrito, há falta de privacidade e pouca independência no uso do banheiro e no uso da cozinha.
	1	Espaço por pessoa é generoso, você tem a privacidade que deseja, tem total independência no uso do banheiro e no uso da cozinha.

## 5.2 Resultado da estimação dos modelos individuais

As Tabelas 1 e 2 mostram, para cada atributo, o valor dos coeficientes, o erro, o teste t e o intervalo de confiança, para os modelos estimados apenas com dados PR e apenas com dados PD, respectivamente.

Os resultados do ajuste estatístico do modelo PR indicam o arranjo aparece como o atributo de maior importância na função utilidade, seguido do valor do aluguel, conforto, qualidade, localização e existência de móvel. Considerando a condição de  $t > 1,96$  para (1 -

$\alpha) = 95\%$ , observa-se que todos os parâmetros da função utilidade obtiveram coeficientes significativos a 0,05.

Quanto ao arranjo, a análise das utilidades das alternativas mostra que a maioria das escolhas tiveram o arranjo em grupo entre os atributos. A técnica da máxima verossimilhança atribui um valor para arranjo em grupo que indica que os cartões que o têm sejam os mais escolhidos, como de fato o foram. No entanto, existiram alguns poucos casos com arranjo sozinho e aluguel baixo que deveriam ser ainda mais escolhidos. Isto na realidade não ocorreu. A lógica diz que isto é melhor. Da mesma forma existiam alguns casos de moradia em grupo e aluguel alto que deveriam ser rejeitados, mas não o foram. Isto mostra que apenas o fato de ter moradia em grupo já resulta numa utilidade boa, que não pode ser rebaixada pelo aluguel.

A leitura da função utilidade mostra que o sinal negativo do atributo valor de aluguel indica uma maior utilidade na faixa de 150,00 reais. O parâmetro obtido no atributo conforto confirmou a maior utilidade na condição satisfatória, do que na condição contrária. O mesmo acontece com a qualidade. Estes dois atributos juntos na condição mais favorável tendem a aumentar a utilidade. A localização mais escolhida é a próxima da universidade. Os estudantes escolheram mais habitações com mobília, sendo a utilidade de habitações que incorporam mobília maior do que as que não possuem.

Os resultados do ajuste estatístico do modelo PD indicam que o valor do aluguel como o atributo mais significativo na escolha da habitação pelo estudante, seguido do conforto, arranjo de moradia e existência de mobília. Qualidade e localização não apareceram significativos na função utilidade.

O processo de decisão foi similar ao encontrado no modelo PR no que se refere ao comportamento dos atributos. A diferença obtida aparece na utilidade do arranjo que é maior para sozinho. O modelo PD indicou que a maioria das escolhas tiveram o arranjo sozinho

entre os atributos. No entanto, para uma situação onde o arranjo é sozinho, mas o aluguel alto, este último atributo baixa a utilidade.

Tabela 1: Resultados do ajuste estatístico do modelo PR

Atributo	Coefficiente	Erro	Teste t	IC (t=2.5%)
Arranjo	-2.1412	0.2544	-8.4180	[-2.64 ; -1.64]
Aluguel	-1.6141	0.1663	-9.7040	[-1.94 ; -1.29]
Localização	0.6176	0.1305	4.7339	[0.36 ; 0.87]
Existência mobília	0.3775	0.1212	3.1146	[0.14 ; 0.61]
Qualidade	1.2146	0.1543	7.8711	[0.91 ; 1.52]
Conforto	1.3080	0.1488	8.7882	[1.02 ; 1.60]
Número de entrevistas: 281		Número de casos: 281		
F (0) = -1014.6679		F (B) = -806.8433		
LR (-2[F(0)-F(B)]) = 415.6494		Eficiência = 1.000		
$\rho^2 = 0.2048$		QMR = 0.21429		

Tabela 2: Resultados do ajuste estatístico do modelo PD

Atributo	Coefficiente	Erro	Teste t	IC (t=2.5%)
Qualidade	- 0.0711	0.1072	-0.6629 n/s	[-0.286; 0.143]
Arranjo	0.3588	0.1089	3.2952	[0.141; 0.577]
Aluguel	-0.7686	0.1151	-6.6743	[-0.999; -0.538]
Localização	0.0265	0.1072	0.2471 n/s	[-0.188; 0.241]
Mobília	0.3133	0.1085	2.8880	[0.096; 0.530]
Conforto	0.5730	0.1116	5.1357	[0.350; 0.796]
Número de entrevistas: 450		Número de casos: 450		
F (0) = -623.8325		F (B) = -563.3687		
LR (-2[F(0)-F(B)]) = 120.9275		Eficiência = 0.7742		
$\rho^2 = 0.0969$		QMR = 0.0723		

### 5.3 Resultado da estimação com dados combinados

O primeiro passo para a estimação do modelo conjunto foi a realização do teste visual para igualdade dos vetores parâmetros. O gráfico da Figura 2 apresenta o teste visual para a identificação desta igualdade dos vetores parâmetros sobre os dois tipos de dados de preferências, onde a inclinação da reta determina o valor do fator escalar. O gráfico sugere que entre os atributos comuns aos dois modelos, apenas o conforto, a mobília e o aluguel

comportam-se de maneira que os parâmetros sejam tenham o mesmo comportamento em ambos os conjuntos de dados, embora os fatores escalares difiram. A tangente do ângulo de inclinação da reta, traçada a partir destes três atributos de ligação, indica que o parâmetro escalar, neste caso é 0,5558.

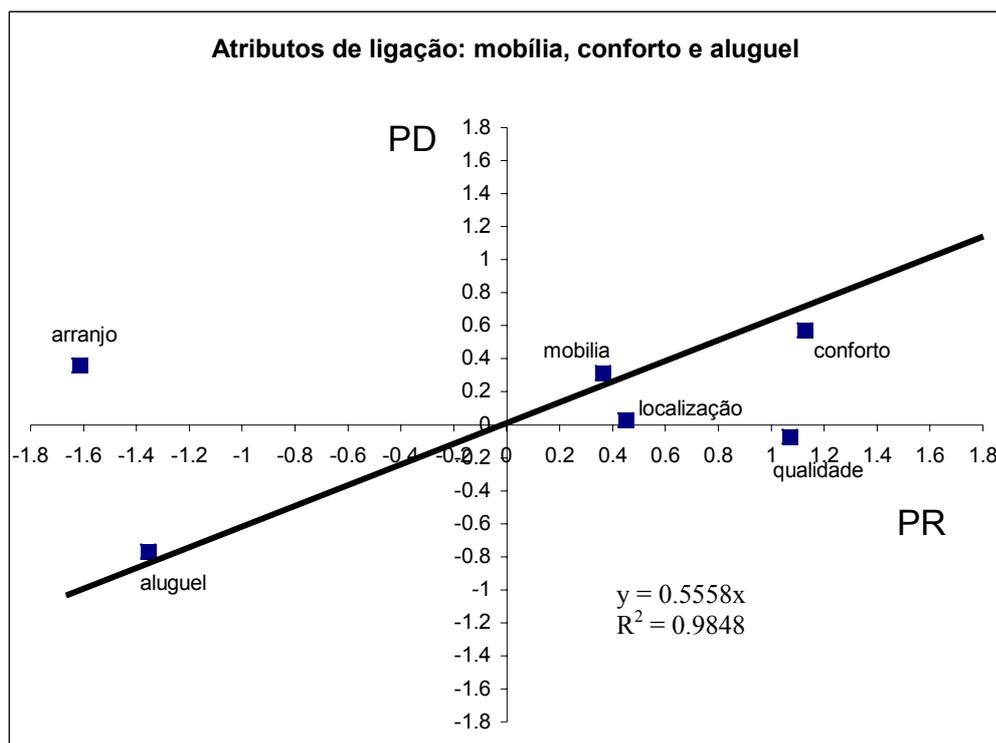


Figura 2 – Teste visual para igualdade dos vetores parâmetros

Para a estimação conjunta pelo método seqüencial adotou-se os seguintes procedimentos, cujos resultados estão resultados estão na Tabela 3.

- (1) **Estimação dos coeficientes PD:** modelo gerado anteriormente;
- (2) **Definição da variável composta:** neste caso todos os atributos são comuns a ambos os modelos. Para a determinação da variável composta foram considerados apenas os atributos que obtiveram comportamentos similares e significância a 0,05. Assim, foram desconsiderados: os atributos localização e qualidade, pois não apresentaram significância a

0,05 e o atributo arranjo pois apresentou um comportamento oposto nos dois modelos. Finalmente o atributo composto incluiu os atributos aluguel, conforto e mobília.

(3) **Estimação do valor do parâmetro escalar com os dados de PR:** Neste caso, o coeficiente do atributo composto obteve um valor de 2,0284 o que indica um parâmetro escalar igual a  $\theta = 0,4929$ . Comparando-se ao valor de  $\theta = 0,4727$  obtido pela tangente observa-se que os valores são muito próximos.

(4) **Correção dos dados e obtenção do modelo final**

O modelo ajustado com ambos os tipos de dados é mostrado na coluna (4) da Tabela 3.

Tabela 3: Resultados do ajuste estatístico do modelo PR+PD

	(1)	(2) e (3)	(4)
Atributo	Coeficiente (teste t)	Integração Seqüencial	Modelo corrigido PR+PD
Arranjo X	0.3588 (3.29)	<b>-2.1038 (-8.33)</b>	0.1768
Localização	0.0265 (0.24) n/s	<b>0.6096 (4.67)</b>	<b>0.6096</b>
Qualidade	-0.0711 (-0.66) n/s	<b>1.1917 (7.76)</b>	<b>1.1917</b>
Aluguel	- 0.7686 (-6.67)	atributo composto	-0.3788
Existência mobília	0.3133 (2.88)	atributo composto	0.1544
Conforto	0.5730 (5.13)	atributo composto	0.2824
Coeficiente do atributo composto (1/ $\theta$ )		1/ $\theta = 2.0284$ (13.51) $\theta = 0.4929$	
Número de observações	450	281	
F(0)	-623.8325	-1014.6679	
F(B)	-563.3687	-809.6685	
LR (-2[F(0)-F(B)])	120.9275	409.9999	
$\rho^2$	0.0969	0.2020	

Neste modelo, a qualidade passou a ser o atributo mais importante da função utilidade, seguido da localização. O arranjo de moradia consolidou sua maior utilidade na condição de moradia sozinho e a localização sua maior utilidade na condição próxima à universidade

A análise das escolhas das alternativas mostra que os atributos arranjo e existência de mobília afetam de maneira similar as utilidades. Uma situação com arranjo de moradia sozinho e sem mobília tem uma utilidade parecida à situação com arranjo de moradia em grupo e com mobília, mantendo-se os demais atributos nos mesmos níveis nas duas situações.

As alternativas mais escolhidas são aquelas que possuem boa qualidade ou combinam o aluguel baixo e a localização próxima à universidade. A situação contrária, aluguel alto e localização no centro apresenta uma utilidade ruim, mesmo que todos os demais atributos estejam na sua condição mais favorável.

A situação que possui a localização próxima à universidade, com piores condições de qualidade e conforto e sem mobília é tão escolhida quanto as situações que são no centro, com melhores condições de qualidade e conforto e com mobília.

A análise das utilidades gerada a partir da combinação dos níveis dos atributos indica a utilidade daquela alternativa. A Figura 3 mostra as situações de moradia de maior utilidade resultantes dos três modelos estimados.

Em todos os modelos apareceram diferenças na importância relativa de cada atributo na contribuição da utilidade total.

As diferenças dos modelos individuais foram que em PD a melhor situação indicou qualidade menos satisfatória, mas sem significância estatística, enquanto que nos outros dois modelos a qualidade preferida tem condições satisfatórias. No modelo PR a melhor situação indicou a moradia em grupo, no modelo PD a melhor situação indicou a moradia sozinho. O modelo conjunto indicou uma preferência por moradia sozinho. As demais características, valor do aluguel, localização, existência de mobília e conforto, confirmaram o mesmo comportamento nos três modelos.

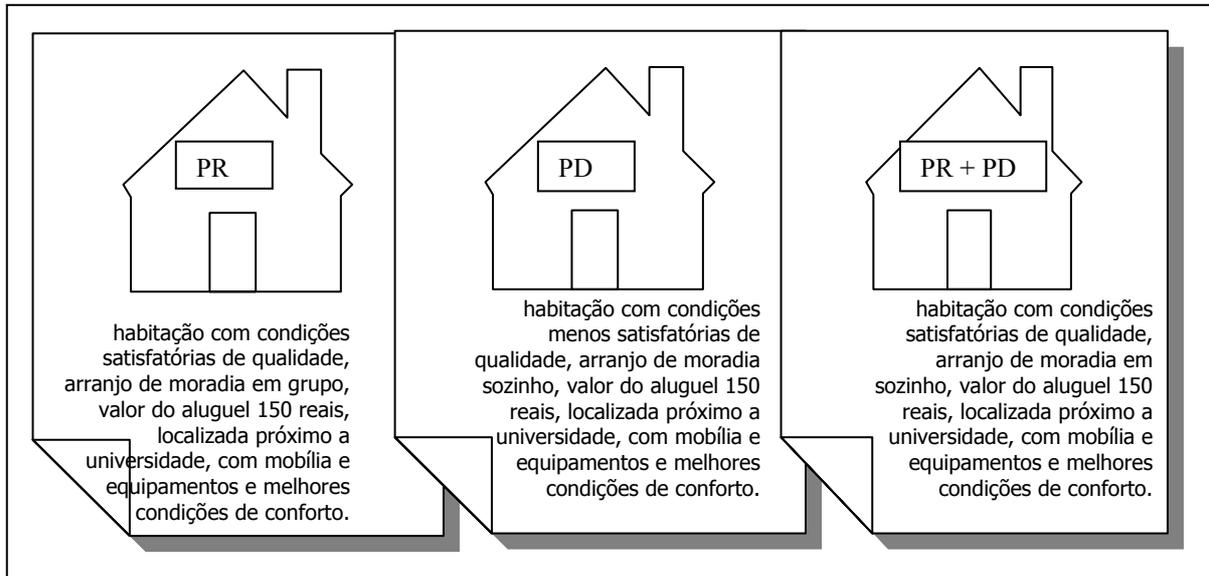


Figura 3: Situação de moradia de maior utilidade nos três modelos.

## 6. DISCUSSÃO E CONCLUSÕES

Este artigo combinou dados de preferência declarada e revelada com o objetivo de obter as vantagens da complementaridade das forças e franquezas de ambas os tipos de dados. A abordagem envolveu (a) a estimação de um modelo apenas com dados de PR; (b) a estimação de um modelo apenas com dados de PD; (c) a estimação sequencial com dados de PR e PD.

As funções utilidades obtidas em cada modelo mostraram a importância relativa do atributo, a tendência de comportamento através dos sinais e o seu nível de significância. A análise dos resultados dos modelos separadamente, no entanto, indicou diferenças e similaridades entre o comportamento dos atributos.

Os resultados confirmam o posicionamento de Earnhart (2002), de que os modelos podem mostrar as diferenças entre a preferência revelada e a declarada, permitindo que sejam

indicados quais os parâmetros similares no processo e quais os que fazem com que o comportamento atual difira das intenções declaradas.

O modelo estimado com os dados PR+PD confirmou os resultados dos modelos individuais na maioria dos parâmetros. Este modelo, entretanto, não foi obtido pela estimação de uma nova matriz com todos os dados PR e PD e sim corrigindo os coeficientes de PR a partir dos dados de PD.

Apesar dos resultados, não se pode afirmar que a combinação dos dados melhorou o poder explanatório do modelo, uma vez que não se pôde obter os testes estatísticos dos modelos conjuntos, por questões relacionadas ao software utilizado na estimação.

## REFERÊNCIAS

ADAMOWICZ, W.; LOUVIERE, J.; WILLIAMS, M. Combining revealed and stated preference methods for valuing environmental amenities. **Journal of Environmental Economics and Management.**, v. 26, p. 271-292, 1994.

BEN-AKIVA; LERMAN. **Discrete choice analysis: theory and application to travel demand.** Cambridge, Massachusetts: The MIT Press, 1985. 390 p. (MIT Press Series in Transportation Studies.)

BEN-AKIVA, M.; MORIKAWA, T. Estimation of switching models from revealed preferences and stated intentions. **Transportation Research part A - Policy and Practice**, v. 24, n. 6, p. 485-495, 1990.

BRADLEY, M. Use of the logit scaling approach to test for rank-order and fatigue effects in stated preference data. **Transportation**, v. 21, n. 2, p. 167-184, 1994.

BRADLEY, M.A.; DALY, A. J. Estimation of logit choice models using mixed stated preference and revealed preference information. In: INTERNATIONAL CONFERENCE ON TRAVEL BEHAVIOUR, 6 th., 1992, Quebec. **Proceedings...** p. 116-133.

BRADLEY, M.A.; KROES, E.P. Simultaneous analysis of stated preference and revealed preference information. In: PTRC 18TH SUMMER ANNUAL MEETING SEMINAR ON TRANSPORTATION PLANNING METHODS, 18 th., September, 1990, University of Sussex. 1990, England, **Proceedings...** p. 1-16.

BRANDLI, L.L. Modelo de demanda habitacional de estudantes numa perspectiva de desenvolvimento local. 2004. 308 p. Tese (Doutorado em Engenharia de Produção) - Programa de Pós-Graduação em Engenharia de Produção, UFSC, Florianópolis, 2004.

BRANDLI, L.L.; HEINECK, L.F.M. A iniciação no mercado habitacional de estudantes universitários e a escolha da habitação. In: XXIII Encontro Nacional de Engenharia de Produção. 2003, Ouro Preto, MG, **Anais em CD...**, 21 a 24 de out de 2003, Ouro Preto, 2003.8p.

BROWNSTONE, D.; BUNCH, D.S.; TRAIN, K. Joint mixed logit models of stated and revealed preferences for alternative-fuel vehicle. **Transportation Research part B - Methodological**, v. 34, p. 315-338, 2000.

EARNHART, D. Combining revealed and stated preference methods to value environmental amenities at residential locations. **Land Economics**, v. 77, n. 1, p. 12-29, Feb. 2001.

\_\_\_\_\_. Combining revealed and stated data to examine housing decisions using discrete choice analysis. **Journal of Urban Economics**, v. 51, n. 1, p. 143-169, Jan. 2002.

HENSHER, D.A. Integrating revealed preference and stated response data into a jointly estimated hierarchical mode choice model. In: AUSTRALASIAN TRANSPORT RESEARCH FORUM CONFERENCE, Seventeenth., 1992, **Working Paper ITS - WP - 92-9**. Canberra, 1992. p. 1-9.

\_\_\_\_\_. Stated preference analysis of travel choices: the state of practice. **Transportation**, v. 21, n. 2, p. 107-133, 1994.

INTRILIGATOR, M; BODKIN, R.; HSIAO, C. **Econometric models, techniques, and applications**. 2. ed. United States of America: Prentice Hall, 1996. 653 p.

KROES, E.P.; SHELDON, R. J. Stated preference methods. An introduction. **Journal of Transport Economics and Policy**, v. 22, n. 1, p. 11-25, 1988.

MORIKAWA, T. **Incorporating stated preference data in travel demand analysis**. 1989. 203 f. (Doctor of Philosophy) - Department of Civil Engineering, Massachusetts Institute of Technology. Cambridge, 1989.

ORTÚZAR, J. D. **Modelos econométricos de elección discreta**. Chile: Ediciones Universidad Católica de Chile, 2000. 249 p.

SCHMITZ, R. **Uma contribuição metodológica para avaliação da tarifa de pedágio em rodovias**. Florianópolis, 2001. 189 f. Tese (Doutorado em Engenharia de Produção) - Programa de Pós Graduação em Engenharia de Produção, UFSC.

SENNA, L.A. ; TONI, J.; LINDAU, L.A. O valor monetário atribuído pelos usuários ao conforto no transporte público. In: ENCONTRO DA ASSOCIAÇÃO NACIONAL DE

PESQUISA E ENSINO EM TRANSPORTES, VIII., 1994, Recife. **Anais...** Recife, 1994. p. 95-105.

SOUZA, O. A. **Delineamento experimental em ensaios fatoriais utilizados em preferência declarada.** Florianópolis, 1999. 179 f. Tese (Doutorado em Engenharia de Produção) - Programa de Pós-graduação em Engenharia de Produção, UFSC, 1999.

SOUZA, O. A. et al. Modelagem da decisão pela escolha portuária para exportação de produtos em contêiner utilizando-se das técnicas de preferência declarada. In: ANPET, 2000, **Engenharia de Tráfego e Transportes: Avanços para uma Era de Mudanças.** Gramado, 2000. p. 537-551.

SWAIT, J.; LOUVIERE, J. The role of the scale parameter in estimation and comparison of multinomial logit models. **Journal of Marketing Research**, v. 30, p. 305-314, 1993.

SWAIT, J.; LOUVIERE, J.J; WILLIAMS, M. A sequential approach to exploiting the combined strengths of SP and RP data: application to freight shipper choice. **Transportation**, v. 21, n. 2, p. 135-152, 1994.

TIMMERMANS, H.; MOLIN, E.; VAN NOORTWIJK, L. Housing choice processes: stated versus revealed modeling approaches. **Neth. J. of Housing and the Built Environment**, v. 9, n. 3, p. 215-227, 1994.

WARDMAN, M. A comparison of revealed preference and stated preference models of travel behaviour. **Journal of Transport Economics and Policy**, v. 22, n. 1, p. 71-91, 1988.