A Methodology to Compute Regional Housing Index Price using Matching Estimator Methods

Paredes - Araya, Dusan
University of Illinois

04. May 2009

A Methodology to Compute Regional Housing Index Price using Matching Estimator Methods

Dusan J.C. Paredes

REAL, University of Illinois, South Mathews Av. 607, Urbana 61801-3671, IL, USA

1-217-244-7226

paredes1@illinois.edu

Abstract. This paper proposes a methodology for a spatial cost index of housing that considers spatial heterogeneity in properties across regions. The index is built by combining three different techniques to reduce the spatial heterogeneity in housing: Quasi-experimental methods, hedonic prices and Fisher spatial price index. Using microdata from the Chilean survey CASEN 2006, it is shown that the quasi-experimental method called Mahalanobis metric within propensity score calipers (MMWPS) leads to a significant reduction in the potential bias. The technique matches dwellings of a particular region with other properties of similar characteristics in the benchmark region (Metropolitan region). Once the houses are matched, a hedonic price model is computed, and a regional housing price matrix is created using Fisher spatial price indices. The paper concludes the existence of price differentials for homogeneous houses across regions in Chile.

Keywords: Housing Cost, Index Hedonic Prices Index, Matching Estimator, Spatial Fisher Index.

INTRODUCTION

A regional housing price index could provide important information in the investigation of the housing market and for the development of regional public policy. In the housing market, the price index provides a measure for regional housing demand, regional price trends, and residential real estate investment decisions. Concerning public policy, the price index would be helpful in the formulation and design of housing policies, social housing programs or any public policy focused on regional housing markets. Therefore, to design adequate regional housing policies and to understand the dynamics of the housing market, the regional scientists must be able estimate precisely housing prices index across different regions or spatial units.

The contribution of this paper is to take account of the heterogeneity in the comparison of regional dwellings using a quasiexperimental control group method (Rosenbaum & Rubin, 1983). Using this method, we match dwellings between different regions with similar characteristics and quality. The output will be two samples (one for each region) of houses with homogeneous characteristics; which will allow comparing house prices through a regional housing price index.

In this paper, we use three kinds of quasiexperimental control group methods: 1) nearest matching on the propensity score, 2) Mahalanobis matching including the propensity score and 3) Mahalanobis matching within score calipers. This paper evaluates the three methods in the context of the housing price index and chooses the best method based on reduction of the average regional bias measured trough standardized differences of the houses characteristics between the spatial units (Tritchler, 1995). Indeed, we measure the difference in mean as a percentage of the variance for each region with respect to metropolitan region (base region). Large standardized differences indicate high geographic heterogeneity while small implies low ones.

The matching method will allow identifying one "control house" in the metropolitan region for each "treatment house" in any region[1], both having statistically similar characteristics. Therefore, one generates two samples with the same number of observations. Hedonic regressions are running on these samples, estimating the hedonic coefficients for the characteristics in each matched sample. Using Spatial Fisher Housing Price Index and his superlative property, a regional housing index price in Chile for 2006 is calculated.

The results show that the Mahalanobis matching within score calipers was the best method to reduce the geographic bias for each covariate among regions and the regional differential in propensity score in the Chilean case. In addition, the Regional Fisher Price Index shows that the region II (Antofagasta) is the most expensive of the country.

This paper contains five sections. Section two surveys the relevant regional housing price index literature. Section three discusses the quasiexperimental methods used and

[1] There exist twelve regions in Chile and one Metropolitan Region. Given that the Metropolitan region is the largest one in the country, each house of the regions was matched to a "control house" from the Metropolitan Region.

presents the hedonic functions. Section four reports the estimation of quasiexperimental methods and parameters for hedonic regression. Finally, section five presents the conclusions.

REVIEWING THE REGIONAL INDEX PRICE LITERATURE

Rosen's (1974) work provided the basis for the use of hedonic regression as the principal tool to investigate housing prices. Although there is a broad consensus for the appropriateness of hedonic regression to build a house index price, most of the applications have not considered the geographic heterogeneity in houses. However the recent literature offers some advances in this area.

Edwin and Simenauer (1996) proposed a regional house price index for the United States using national data, and incorporated geographic heterogeneity through regional fixed effects. The paper revealed significant differences between the index published by the National Association of Realtors (NAR) and their proposed regional index. Despite efforts to construct a regional index, the authors did not calculate a hedonic regression for each region, thereby assuming homogeneous behavior of the parameters over space.

A different approach is taken by Forret (1991) who partially incorporated the geographic heterogeneity in his construction of the regional house price index in 1985 for regions in England. The author recognized the regional heterogeneity in housing, arguing that regional differences in price may stem from the neighborhood and physical characteristics of the respective regional housing stocks. To extract the heterogeneity in the estimation process, the author discomposed the regional differentials between those attributed to "housing characteristics" and "unambiguously attributable to differences in regional location". In summary, the author discomposed the regional differential price, establishing a "bias" generated by heterogeneity, but did not use hedonic regression to solve the econometric problem associated with from geographic heterogeneity.

The importance of regional heterogeneity can appear at the intra-regional spatial levels, for example, within metropolitan areas. Thibodeau (1989) computed a tenure specific hedonic housing price index for sixty metropolitan areas in the US. Calculating the hedonic coefficients for each MSA, the houses are "priced" at market value. Finally, the average ratio of the housing value was estimated for each MSA, in time and space, reflecting the housing index price. His paper computed the hedonic regression for each MSA, but the interpretation of the regional index is debatable because the coefficients used to calculate the value of the house came from heterogeneous housing data. In other words, the coefficient hedonics were calculated with regional housing data, but without any attempt reducing geographic bias.

In spite of great strides made in this area, the spatial aspect still requires more consideration in constructing a regional house price index. Costello and Watkins (2002) highlighted the importance of local house price indices. According to them, the index must consider the minimum geographic scale available, and respect the differences among urban markets. In this sense, heterogeneity is a vital element in building the index, especially when the data show high levels of heterogeneity. According to Paredes and Aroca (2008), geographic heterogeneity could be reduced using the quasi-experimental control group method. Before building the regional price index, the authors calculated the bias between any pair of regions in Chile. The bias is the average difference in the independent variables to construct the index price. Using nearest neighbor matching estimator, the authors matched similar houses in two regions. With this methodology, they considerably reduced the bias among regional houses. Finally, they calculated a Fisher regional index price using hedonic regression and incorporate a methodology to reduce the regional heterogeneity. However, they do not explore with quasiexperimental methods fit better for this purpose. A better quasiexperimental method could lead to more statistically reliable regional house index price.

Summarizing, only few articles focus on building a regional housing price index that considers spatial heterogeneity. Mills and Simenauer (1996) and Thibodeau (1989) marginally incorporate the regional aspect, but do not make the effort to embrace the spatial dimension. Forret (1991) identifies the heterogeneity problem and developed a measure of the bias, but does not provided a solution for it. Paredes and Aroca (2008) demonstrated an alternative method, but did not test with alternative quasi-experimental control groups. Quasiexperimental control groups may provide the methodology to construct the regional price index; however the present article will propose different quasi-experimetnal methods to improve the “reduction bias” proposed by Paredes and Aroca.

METHODOLOGY

In this paper two separate methodologies are used to calculate the regional housing price index. The first part of this section describes three different matching methods: 1) nearest available matching on the estimated propensity score, 2) Mahalanobis metric matching including the propensity score and 3) nearest available Mahalanobis metric matching within calipers defined by the propensity score. Each one must be evaluated in order to select the best method. The second part of this section explains hedonic regression and introduces the spatial Fisher price index.

MATCHING ESTIMATOR METHODS

The quasi-experimental methods estimate the effect that treatment would have had on a unit that, in fact, did not receive the treatment (Rosenbaum and Rubin, 1983). Formally, the literature calls the treated group as “treated group” and the group without treatment as “potential control group” (Rubin (1976); Rosenbaum and Rubin (1983)).

The concept of treatment applies to all situations where two groups exist, a treated and control group. In this sense, the origin of the treatment could be diverse and the consequences analyzed in different contexts. For example, Hii and Frei (2002) studied the effect of electronic banking on the profitability of the bank clients. D'Agostino (1998) used matching method to study the risk for women during pregnancy. Glasmeier & Farrigan (2007) use quasiexperimental methods to examine the effect of state-run prisons constructed in rural counties. Whatever the application, they all have a common denominator: a small treated group and a large group with no treatment.

In the regional housing market, the treatment can be considered as the presence of a house in a specific region. To know the regional differential price of the house, the ideal situation would be to have exactly the same house in two different regions at the same time, but this comparison is impossible. The alternative is that each regional house is compared to a similar house in the benchmarking region. The comparison is realized between each region $i=\{1,2,..,n\}$ and the benchmarking region, taking in consideration that each house belonging to $i=\{1,2,..,n\}$ could have different characteristics than the one in the benchmarking region, which should be the largest region of the country in order to have more degree of freedom to build the potential control group and it should be like the non-treated region.

Summarizing, the matching estimator produces a control group from the benchmarking region for each region $i = \{1, 2, .., n\}$ with statistically similar covariates. This process reduces the geographic heterogeneity and f to construction of a regional housing price index that can be compared among regions. To understand the method this paper present a short introduction of the propensity score and matching estimator using the notation of Rosenbaum and Rubin (1985) adapted to the housing market.

Let $\mathbf{x}$ characterize the vector for a specific house, and the z variable indicates whether the house is located in a specific region $(z=1)$ or in the benchmarking region (metropolitan area)$(z=0)$. The propensity score $e(\mathbf{x})$ is defined as conditional probability to be located in the region i given the covariates, that is to say $e(\mathbf{x}) = \Pr(z=1|\mathbf{x})$. The matching using $e(\mathbf{x})$ will balance the distributions of $\mathbf{x}$ between the region i and the benchmarking one (Rosenbaum and Rubin, 1983).

At this point, the matching estimator presents two issues. First, the functional form for $\Pr$ is unknown, therefore it must be estimated from the available data. Secondly, $e(\mathbf{x})$ has a continuous metric and as it is impossible for two $e(\mathbf{x})$ to match exactly, it is therefore necessary to choose an objective criterion to match similar $e(\mathbf{x})$.

The literature highlights different functional forms for the probability. For the binary case, Smith (1997) establishes that there is no critical differences for the popular logit and probit densities, excepting for the constraint on the data generation process where the probabilities are independent of irrelevant alternative impose by the logit. The present paper works with a probit model to estimate $e(\mathbf{x})$.

Regarding the second issue, this paper compare three methods to match the propensity score: nearest-neighbor matching (NNM), Mahalanobis metric matching with propensity score (MMPS) and Mahalanobis metric within propensity score calipers (MMWPS). The adjustment for each matching method is evaluated using standardized differences. Specifically, the method with the greatest bias reduction measured through standardized differences will be chosen as the best fit model.

NEAREST-NEIGBOUR-MATCHING (NNM)

The NNM matches each house belonging to region i with a house in the benchmarking region with a similar $e(\mathbf{x})$. To formalize the discussion, assume that $e_i^k\left(\mathbf{x}_i^k\right)$ represents the propensity score of house k in region i considering the covariates $\mathbf{x}_i^k$. Let $e_{mr}^n\left(\mathbf{x}_{mr}^n\right)$ represents the propensity score of the house n in the benchmarking region according the covariates $\mathbf{x}_{rm}^n$. N_i and N_{rm} are the numbers of observations of regions $i=\{1,2,..,12\}$ and benchmarking region rm, respectively. Putting all this together, houses are matched using the equation

$$C^{NNM}\left(e_i^k\right)=\min_n\left\|e_i^k-e_{rm}^n\right\|, n\in N_{rm} \quad (1)$$

In this case, $\|(.)\|$ is either based on comparing the index function or it is obtenained through a distance metric. This matching selects a control observation just once for each region i, therefore control houses are drawn without replacement. In addition, control group are built independently for each region, so a house in the benchmarking region can belong to more than one control group.

MAHALANOBIS METRIC MATCHING INCLUDING THE PROPENSITY SCORE AS COVARIATE (MMPS).

The Mahalanobis metric matches two observations using the Mahalanobis distance of the covariates. This method, instead of minimizing the difference in the propensity scores between treated and control observations, finds for each treated observation a control individual with the closest characteristics estimated through the Mahalanobis distance. The MMPS uses the Mahalanobis distance, but includes the propensity score as a covariate. Houses are matched using the equation:

$$DM=\left(\mathbf{x}_i^k-\mathbf{x}_{rm}^n\right)^T\mathbf{C}^{-1}\left(\mathbf{x}_i^k-\mathbf{x}_{rm}^n\right) (2)$$

where $\mathbf{x}_i^k$ and $\mathbf{x}_{rm}^n$ include $e_i^k\left(\mathbf{x}_i^k\right)$ and $e_{mr}^n\left(\mathbf{x}_{mr}^n\right)$. Rubin and Thomas (2000) recognize the advantages of including $e(\mathbf{x})$ in NNM, especially in handling possible problems arising from selection bias.

MAHALANOBIS METRIC WHITIN PROPENSITY SCORE CALIPERS (MMWPS)

The MMWPS is an hybrid method that defines a subset control candidates using the propensity score as a caliper, and selects the control using the Mahalanobis metric on the covariates, including the propensity scores (as MMPS). Sequentially, the first step for MMWPS is to find the closest controls using the propensity score according to the caliper defined. With this subset defined, the possibilities of finding a control with the Mahabalobis method are higher than with the last two proposed matching techniques. The second step for MMWPS is similar to MMPS namely, to match houses with similar Mahalanobis metrics using the covariates and propensity score to compute the metric. The key element of this method is the definition of the caliper scalar. Rosenbaum and Rubin (1985) suggest choosing the caliper based on the variance of the propensity score for each group. If σ_i^2 is the variance for the propensity score in each region $i=\{1,2,..,n\}$ and σ_{rm}^2 is the variance for the benchmarking region, the caliper should be a function of $\sigma=\left[\left(\sigma_i^2+\sigma_{rm}^2\right)/2\right]^{1/2}$. This paper follows Crochran and Rubin (1973), who recommend a caliper width of $c=0.2\sigma$.

HEDONIC PRICE AND SPATIAL FISHER PRICE INDEX

The hedonic price is the standard methodology for studying heterogeneous goods. Rosen (1974) provided the theoretical background for the interrelationship among supply, bid price function of consumers, and hedonic prices. Particularly, Chile has thirteen regions, therefore this article estimates a regression hedonic for each region $i=\{1,2,..,12\}$ and the capital of the country (metropolitan region) will be the benchmarking.

$$\ln P_i^k=\beta_{0i}+\sum_{j=1}^{J}\beta_{ji}\mathbf{x}_i^k+\varepsilon_i^k \quad (3)$$

$$\ln P_{rm}^{nik} = \beta_{0rm}^{i} + \sum_{j=1}^{J} \beta_{jrm}^{i} \mathbf{x}_{rm}^{nik} + \varepsilon_{rm}^{nik} \quad (4)$$

In equation (3), P_i^k is the price of the house k in the region i and $\mathbf{x}_i^k$ are the characteristics for the same house. The coefficients β_{0i} and β_{ji} vary for each region i. Equation (4) P_{rm}^{nik} represents the price for the housing n of the benchmarking region rm matched with the housing k in the region i. Given that the three matching methods are nearest neighbor, and then k must be equal to n.

The key of the regional comparison is the estimate of β. In this case, the $\mathbf{x}$ covariates are statistically similar for each region $i = \{1, 2, .., 12\}$ and their control groups samples, therefore the difference in β can be attributed to price differences and no quality differences. Otherwise, without the reduction in geographic bias, different β could be due to different distributions of X. If this were the case, for example if β_{ji} were greater than β_{jrm}^{i}, it could be that region i has the characteristic j with higher quality than rm and it generates as a consequence a higher shadow price, biasing the price index.

Having estimated β coefficients for each region and control groups, the next step is to estimate the value of the housing according to $\mathbf{x}$. Previous research has shown that the advantage of the hedonic regression is that it is able to maintain the quality constants. However, the spatial comparison demands attention to the influence of the basket of characteristics on the value of the house in both regions. For example, suppose that region i has a "housing basket" completely different than the control group. In this case, the selection of one basket (region i or non-treated group) could incorporate bias in the estimation of the regional housing price index. This problem has been broadly documented in temporal price indexes, such as the Paasche and Laspeyres price index. This paper proposes a geometric mean of these two indexes to estimate the regional housing price index, as follows:

$$I_{i/in} = 0.5\left(\ln\left(\overline{P}_{rm}^{i}\right) - \left(\beta_{0i} + \sum_{j=1}^{J} \beta_{ji} \overline{\mathbf{x}}_{rm}^{i} \right) \right) + 0.5\left(\left(\beta_{0rm}^{i} + \sum_{j=1}^{J} \beta_{jrm}^{i} \overline{\mathbf{x}}_{i} \right) - \ln\left(\overline{P}_{i}\right) \right) (5)$$

This price index is called Fisher Regional Housing Price Index. This measure removes the bias due to different housing baskets given the previous procedure to build the samples. It allows the “price” to be the average house characteristics of the region i at prices β^{i}_{jrm} in the benchmarking region and the average benchmarking house characteristics priced at regional prices β_{ji}. Finally, the geometric mean represent the index price between the region i and benchmarking region rm. Nonetheless, the regional housing price index must be constructed between all regions, and not only between region $i = \{1, 2, .., 12\}$ and rm.

The Fisher Price Index allows the direct comparison among regions using the superlative propriety (Diewert, 1978). The Spatial Fisher Housing Price Index between the region i and the region $j \in i = \{1, 2, .., 12\} \ \forall i \neq j$ is given by the quotient between the index price of the region i and j respect to rm. This property can be established as

$$IF_{i/j} = \frac{IF_{i/rmi}}{IF_{j/rmj}} \quad (6)$$

With this information, this paper reports a matrix of regional housing index price for the twelve regions plus the benchmarking one.

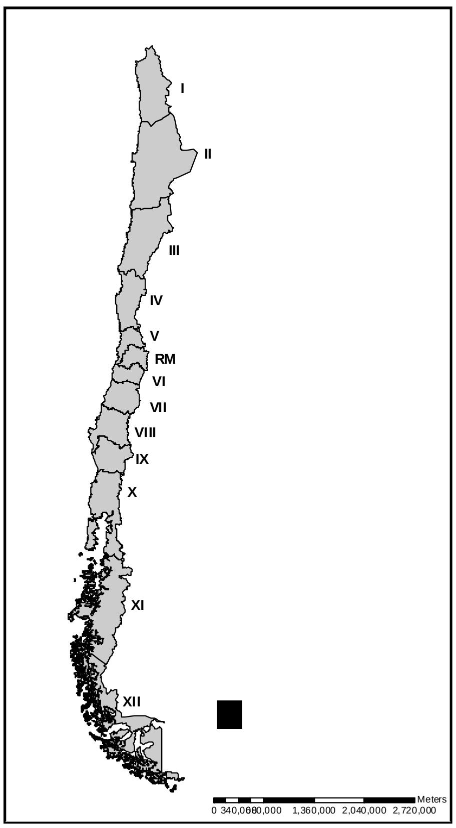
$$MIF = \begin{bmatrix} 1 & IF_{2/1} & . & . & IF_{13/1} \\ & . & & & \\ & & . & & \\ & & & . & \\ F_{1/13} & . & . & . & 1 \end{bmatrix} \quad (7)$$

The cell identified in the first row and second column represents the index between the region 1 and 2. If this number is greater than 1, it means that an individual that moves from 1 to 2 should spend, on average additional percentage of rent equal to $(IF_{2/1} - 1)$.

THE DATA

The data are extracted from the Chilean Household Survey 2006. This sample consists of 268,873 individuals covering thirteen regions: twelve regions plus the Metropolitan area (benchmark) which will be used as the benchmarking region because is the largest in the country. Figure 1 describes the geographic distribution of the twelve regions (I to XII) and Metropolitan region (RM). Since hedonic regressions require the house' price and characteristics, only 7,184 renter heads of household were selected from the CASEN 2006. Other filters were added in order to eliminate outliers and missing data, leaving 7,094 observations. The data include a weight variable therefore the hedonic regressions are estimated with weighted sample equal to 649.328 observations.

Fig. 1 Chile's Map



The CASEN survey provides information about housing characteristics, but its quality is limited. There is no information about important physical characteristics, such as the size of the house in squared meters, its age, and whether there is a garage. Also there is not data for spatial location, for example, its distance to the Central Business District, or distance from job, or whether it is close to services centers. These limitations restrict the available models used to estimate regional housing price index.

On the other hand, CASEN gives detailed information about quality characteristics of the housing, but the categorical metric must be changed into a continuous variable. For example the variables for the quality of the floor are categorical (bad, normal, and good). Therefore, they cannot be included directly in the hedonic regressions. Taking these categorical variables, we estimated a continuous quality housing variable using a methodology proposed by Economic Commission for Latin America and the Caribbean (ECLAC).

ECLAC (1996) provides a methodology for constructing a dummy variable that indicates acceptable or unacceptable quality housing using categorical variables. The constructed dummy variable is used as a dependent variable to estimate a logit model using different covariates to estimate the probability for having acceptable quality housing (Paredes and Aroca, 2008). Finally, the probability is multiplied by 100 in order to compute a housing quality index.

Table 1: Covariates

Variables*	**Description Original Covariate**	**As used for estimating the Propensity score**	**Mean**	**Standard Deviation**	**Min**	**Max**
Bedroom	Number	Integers	2.40	1.04	0	12
Room Alternative use	Number	Integers	0.86	0.47	0	3
Restroom	Number	Integers	1.08	0.49	0	5
Kitchen	Number	Integers	0.81	0.41	0	3
Quality Housing	Index Quality	Index between 0 and 1	88.55	9.30	28.53	99.97
Water Heater	Dummy	1 = yes, 0 = no	0.61	0.49	0	1
Phone	Dummy	1 = yes, 0 = no	0.28	0.45	0	1
Cable TV	Dummy	1 = yes, 0 = no	0.28	0.45	0	1
Crowding	Number	Persons per bedroom	1.59	0.82	0.09	8.00
Education	Household Head's Education	Integers, education's year	11.08	3.95	0	20
Age	Years	Years	41.29	12.82	16	98
Civil Status	Dummy	1 = married, 0 = no married	0.45	0.50	0.00	1.00
Sex	Dummy	1 = man, 0 = woman	0.75	0.43	0.00	1.00

*** Descriptive statistics for 7094 observations**

Table 1 provides country descriptive statistics for selected variables to estimate matching methods and hedonic regressions. The first group represents characteristics of the house, such as bedrooms and restrooms. A second group represents the head of the household's characteristics, such as education or age.

TESTING RESULTS FOR MATCHING ESTIMATOR METHODS

This section tests the three matching estimator methods: NNM, MMPS and MMWPS. The quality of each matching method is defined by reducing the difference between regional and metropolitan average housing characteristics of the sample that will be used to build the index price. The regional housing difference can be measured in two dimensions: covariates and propensity score. In the first case two houses could be different if their characteristics are different. Some comparability problems could appear because there are several covariates. Therefore, the differences could exist in some variables and not the others. This problem can be overcome using the single-measure propensity score. This number represents the probability of belonging to a specific region for treated and non treated house, therefore a similar propensity score implies similar houses under the assumptions detailed by Rosenbaum and Rubin (1985). This paper uses both measures to prove the adjustment of matching methods.

Table 2 provides the baseline to analyze the regional heterogeneity of housing, using means and standard deviations. The first twelve rows represent each region and row thirteen, the Metropolitan area (benchmark). The first column has the designated name for each region, which is used in this paper to refer to each region. The second and third column is the sample and weighted sample by region, respectively. The Housing Rent column is the regional average of the natural logarithm housing price. Finally, the rest of the columns are variables used to estimate the three matching estimator methods.

Table 2 shows different average values for each one of the thirteen regions. For example, region XI has a mean of 2.61 bedrooms per house, but it is different than the 2.11 of region II. These values show that the average of house characteristics is different among regions. These differences are evident for almost all variables, specifically in Housing Quality; showing the presence of the geographic heterogeneity through the regions.

Table 2: Sample means of Covariate for CASEN 2006 by Region.

Region	Sample Weighted	No. of Obs. Sample	Housing Rent	Bedroom	Room Another use	Restroom	Kitchen	Quality Housing	Crowding	Education	Age
I	22858	214	11.10	2.29	0.75	0.99	0.74	79.44	1.74	11.06	41.28
			(0.52)	(1.25)	(0.55)	(0.44)	(0.46)	(14.8)	(0.91)	(3.75)	(12.8)
II	22838	293	11.26	2.11	0.79	1.06	0.77	87.71	1.91	11.53	39.18
			(0.59)	(1.00)	(0.49)	(0.41)	(0.47)	(9.52)	(1.11)	(3.64)	(12.3)
III	9880	195	10.95	2.39	0.87	1.01	0.84	90.26	1.61	11.26	39.52
			(0.52)	(0.89)	(0.45)	(0.33)	(0.37)	(7.38)	(0.84)	(3.62)	(11.1)
IV	15642	177	10.93	2.47	0.78	1.07	0.76	87.57	1.53	10.92	41.40
			(0.58)	(1.16)	(0.44)	(0.51)	(0.42)	(9.81)	(0.69)	(4.12)	(13.7)
V	69050	854	11.08	2.46	0.94	1.11	0.91	92.39	1.56	11.26	42.81
			(0.51)	(0.91)	(0.39)	(0.42)	(0.30)	(5.46)	(0.75)	(3.72)	(13.4)
VI	27333	522	10.82	2.43	0.93	0.98	0.88	82.95	1.59	10.11	41.46
			(0.49)	(0.97)	(0.41)	(0.38)	(0.33)	(12.4)	(0.78)	(3.87)	(12.6)
VII	24592	357	10.75	2.41	0.87	1.03	0.79	84.59	1.60	10.11	42.48
			(0.52)	(1.00)	(0.48)	(0.42)	(0.40)	(12.2)	(0.88)	(4.21)	(13.0)
VIII	60946	904	10.87	2.34	0.86	1.02	0.78	85.95	1.61	11.03	40.78
			(0.55)	(0.87)	(0.47)	(0.42)	(0.43)	(9.32)	(0.80)	(3.83)	(12.9)
IX	24353	415	10.78	2.37	0.84	1.03	0.72	87.47	1.52	10.73	40.85
			(0.49)	(0.86)	(0.50)	(0.35)	(0.45)	(7.66)	(0.77)	(4.13)	(13.2)
X	36297	693	10.98	2.48	0.71	1.00	0.64	89.90	1.55	10.81	39.07
			(0.50)	(1.12)	(0.50)	(0.37)	(0.48)	(8.17)	(0.78)	(3.95)	(11.5)
XI	4706	174	11.30	2.61	0.70	1.01	0.63	93.68	1.41	11.42	38.99
			(0.52)	(0.90)	(0.52)	(0.31)	(0.48)	(6.05)	(0.67)	(3.86)	(10.7)
XII	5504	127	11.30	2.43	0.77	1.03	0.66	79.53	1.57	10.60	41.50
			(0.49)	(0.96)	(0.47)	(0.56)	(0.47)	(15.5)	(0.88)	(4.41)	(11.5)
RM	325329	2169	11.42	2.41	0.91	1.20	0.89	90.96	1.59	11.52	42.08
			(0.58)	(1.14)	(0.45)	(0.60)	(0.35)	(5.99)	(0.82)	(3.99)	(13.0)
Total	649328	7094									

These differences are also found in the standard deviations. For example, the variable "bedrooms" variable in region IX has a standard deviation of 0.86 while it is 1.25 in region I. This situation implies that not only that there are differences in the averages, but also in the characteristics' variability within regions.

Two technical details are needed to correctly interpret the regional heterogeneity in the Table 2resulting from the different metrics between variables and the difficulty of making a comparison between regions, taking into account the mean and variance at the same time. The first detail reflects the impossibility of making comparison among means because they are estimated using different metrics.

To deal with these difficulties, this paper uses the standardized difference. Table 5 exhibits the standardized difference for the regional means sample in Table 2. This statistical measure is comparable without taking account the metrics of each variable and represents the difference between means as a percentage of the variance. The covariates for matching estimator are between the second and eighth column. The last column is the average propensity score between each region and the metropolitan region.

The largest differences are between extreme regions (i.e. regions I and XII) and the metropolitan (central) region. According to Rosenbaum and Rubin (1985), standardized differences under 10% are acceptable to make comparison among the regions. In this case, the 37.5 percent of the difference would be acceptable to make comparison among regions. Using the idea Rosenbaum and Rubin (1985), the quality for each matching method is evaluated by looking at the standardized difference measured against situation without the matching estimator method provided by the table 2. The situation described by this table is called "base line".

The last column contains the differences of the propensity score between each region and the metropolitan region. According to the methodology, the matching estimator methods are constructed using a probit model. Specifically, for each region the probability was estimated for each house (region and metropolitan region) belonging to the analyzed region. Consequently, if the average characteristics are different between the metropolitan region and the other regions, then the average propensity will be different as well.

The first test is estimated for NNM matching method and the regional means are displayed in Table 6. In regard to covariates, they exhibit a reduction in the heterogeneity. Table 6 shows a significant reduction of the differences with respect to the baseline. Indeed, 64 percent of the averages are under 10 percent. Some variables, such as restroom and kitchen, show systematic reductions in the heterogeneity for all regions. Nevertheless, there are some variables with persistent differences; for example housing quality variables. The bias measured as difference in propensity score exhibits a significant narrowing in relation to the base line. The difference is close to zero in four regions, indicating the average reduction in the heterogeneity.

In spite of the NNM's estimation improvements, a problem still is in place: the propensity score could match two houses with similar propensity, but different characteristics. To solve this problem, Rosenbaum and Rubin (1983) proposed the use of the Mahalanobis metric as a propensity score and include the propensity score (probit probability) as covariate. The propensity score contains rich information about differences, and its incorporation could reduce problems associated with bias selection. The results for this method are displayed in Table 7.

The MMPS fits better than NNM. The MMPS' heterogeneity reduction is larger than NNM, reaching to 89.58 percent. Some variable biases have been reduced to fewer than 10 percent, such as bedroom, room alternative use, kitchen or crowding. Using the one-dimensional variable (propensity score), it shows a marginal improvement regarding NNM model. Nevertheless, the MMPS could not reduce the bias appropriately for region I, where the difference propensity score is 0.09. In any case, this model considerably reduces the geographic bias with respect to the baseline compared to NNM

Both of the previous methods have reduced the geographic bias in some variables and regions, but there is some uncertainty about the quality of the matching. Specifically, the nearest neighbor methods could be problematic if they match "nearest neigbors" with large differences in propensity scores. The MMPS reduces the probability of this event, but not completely. The theory recommends using a caliper to reduce the subset of potential matches before matching houses.

The two previous methods have not been strict with regard to the difference between treatment and control's propensity score. For example, if there is no overlapping between both distribution (treatment and non-treatment), then the method can match observations with different propensity scores. This phenomenon is highly correlated with characteristics of the distribution, specifically the variance. It means larger variance implies a greater probability to find matched observations with different propensity scores. For this reason the literature suggests using a caliper as a radius to avoid matching observations with extremely different propensity score (Rosenbaum and Rubin, 1983). Rosenbaum and Rubin (1985) suggest that the caliper must be calculated using the variance of treatment and non-treatment group to avoid distributional effects on the selection of the caliper, such as the methodology described previously. The caliper makes it possible to reduce the potential candidates in the non-treatment group to those that stay within the caliper-defined radius. Finally, the Mahalanobis metric is used to find the control observations inside this subset.

Thus, the Table 8 shows the output for the MMWPS method. This method reduces the heterogeneity more than last matching methods. The reduction is 90.63 percent compared to the baseline, removing completely the bias in room alternative use, restroom and kitchen. On the other hand MMWPS reduce the bias in housing quality, although not to less than 10 percent in some regions, but the average improvement is significant. This method proves the best fit, reaching to 90.63 percent, better than 64 percent and 89.58 percent of the previous methods.

Table 3: Summary Matching Methods

Region/Method	PS Base Line	PS NNM	PS MMPS	PS MMWPS
I	0.24	0.05	0.09	0.06
II	0.06	0.00	0.01	0.01
III	0.01	0.00	0.00	0.00
IV	0.09	0.00	0.02	0.02
IV	0.04	0.00	0.00	0.00
V	0.05	0.01	0.01	0.01
VI	0.18	0.07	0.04	0.02
VII	0.11	0.02	0.01	0.01
VIII	0.16	0.07	0.05	0.03
X	0.12	0.01	0.01	0.01
XI	0.15	0.03	0.01	0.01
XII	0.22	0.05	0.02	0.02
Differences Variables	37.50%	64.00%	89.58%	90.63%

The summary for these models is presented in the table 3. The best method is MMWPS. This method reduces the difference in propensity score considerably. The comparison between the first and fourth column indicates the improvement achieved using different matching estimator methods.

RESULTS FOR HEDONIC REGRESSION AND PRICE INDEX MATRIX

Each region has two groups of hedonic prices. The first group was computed using the housing weight data available to each region and the second one represents the same methodology applied to each regional control group belonging to the Metropolitan region. The results are provided in Table 9.

The variables used represent characteristics of the head of household and housing attributes. The upper level of the table contains the estimation with regional data. Given the log-linear specification, the beta coefficient could be approximated as a semi-price elasticity. Most of the variables show the expected sign. For example, the education shows a positive coefficient and it is significant for most of the regions. This supports the hypothesis that people get better houses if they achieved higher education levels. Civil status has an irregular behavior through the regions, showing that there is little evidence supporting significant differences in housing price for different civil status of the head of household. The sex of the head of household has a similar behavior, but there are some regions (IV and XII) with high positive housing price elasticity between male households and prices. On the other hand, the "housing quality" variable has the expected positive sign for the all the regions, except the first region. Finally, the adjustment of the model lies between 0.49 and 0.85, indicating high variability in the adjustment levels, caused likely by the variability of the sample sizes.

The lower part of the table reflects the same hedonic regression methodology, but applied to the control group. One of the advantages of this methodology is that it allows comparison in the hedonic prices estimation for "comparable houses". In this sense, the hedonic prices are comparable because they come from comparable samples, reducing quality bias among regions. In general, the adjustment for the second group is similar to the first.

From an economic point of view, there are "regional prices" and "quantities" for each attributes of the houses at the location region and their similar at the Metropolitan one. The next step is to build a housing regional price index that defines the bundle to value the differences. The first alternative should be to take the regional bundle (regional average house) and compute his value with regional price and control group price (metropolitan average). The second alternative consists of comparing the metropolitan bundle to metropolitan and regional prices. Both these approaches have been to a great extent documented as Paasche and Laspeyres Spatial Price Indices. These approaches present several restrictions. In the first place, each index would bias the information because it would take into account the structure of one spatial unit of the data (treatment or control group). In second place, these indexes only allow for computing the index between one region and the Metropolitan one and it does not allow for comparison among all regions.

To face this problem, the paper computes the Fisher Spatial Price Index (Diewert, 1976). This index considers both bundles, for the region and metropolitan region, reducing the geographical bias and allowing comparability across regions. The table 4 contains the Regional Housing Price Index for Chilean regions. Each cells on the rows shows the price index between the region in the column j and the one on row i which has been set at the numeraire equal one. For example, the housing cost between the region I and the region II is 1.22. This indicates that the housing prices are higher at region II in 22 percent. The matrix shows different results compared to the fourth column in the table 2 called "Housing Rent". Considering only the average house price in the region, the metropolitan region has the most expensive ones in the country.

According to table 4, the differential regional prices have a different behavior when the heterogeneity is considered in the estimation process. Particularly, the region most expensive region in housing is region II, where the maximum difference reaches 73 percent and it is between regions II and VII.

Table N4: Matrix Regional Housing Price Index

Region	**I**	**II**	**III**	**IV**	**V**	**VI**	**VII**	**VIII**	**IV**	**X**	**XI**	**XII**	**RM**

I	1.00	**1.22**	0.81	0.76	0.80	0.74	0.71	0.84	0.75	0.84	0.94	0.87	**1.03**
II	0.82	1.00	0.66	0.63	0.65	0.61	0.58	0.69	0.62	0.69	0.77	0.72	0.85
III	**1.24**	**1.51**	1.00	0.95	0.99	0.92	0.87	**1.04**	0.93	**1.04**	**1.16**	**1.08**	**1.28**
IV	**1.31**	**1.60**	**1.06**	1.00	**1.04**	0.97	0.92	**1.09**	0.98	**1.10**	**1.23**	**1.14**	**1.35**
V	**1.26**	**1.53**	**1.01**	0.96	1.00	0.93	0.89	**1.05**	0.94	**1.05**	**1.18**	**1.10**	**1.30**
VI	**1.35**	**1.64**	**1.09**	**1.03**	**1.07**	1.00	0.95	**1.13**	**1.01**	**1.13**	**1.27**	**1.18**	**1.39**
VII	**1.42**	**1.73**	**1.14**	**1.08**	**1.13**	**1.05**	1.00	**1.18**	**1.06**	**1.19**	**1.33**	**1.24**	**1.47**
VIII	**1.20**	**1.46**	0.97	0.91	0.95	0.89	0.84	1.00	0.90	**1.00**	**1.12**	**1.05**	**1.24**
IV	**1.33**	**1.62**	**1.07**	**1.02**	**1.06**	0.99	0.94	**1.11**	1.00	**1.11**	**1.25**	**1.16**	**1.38**
X	**1.19**	**1.45**	0.96	0.91	0.95	0.89	0.84	1.00	0.90	1.00	**1.12**	**1.04**	**1.23**
XII	**1.07**	**1.30**	0.86	0.81	0.85	0.79	0.75	0.89	0.80	0.89	1.00	0.93	**1.10**
XII	**1.14**	**1.39**	0.92	0.87	0.91	0.85	0.81	0.96	0.86	0.96	**1.07**	1.00	**1.18**
RM	0.97	**1.18**	0.78	0.74	0.77	0.72	0.68	0.81	0.73	0.81	0.91	0.85	1.00

A spatial interesting pattern in housing price emerges and it seems that prices are higher in the extreme North and South of the country and lower in the central regions. For example, I, II, XI and XII region show higher price, indicating a no-random distribution of the regional house price index in the space. The value of this matrix is the clearness to show the spatial distribution of the housing price in contrast to the average showed in Table 2.

CONCLUSION

This paper computes a regional housing price index using quasi-experimental control group methods, hedonic regression and a spatial index price. The adoption of quasi-experimental control group methods reduce the geographical heterogeneity in the regional comparison of housing. Subsequently a matrix of regional housing price index is reported using hedonic price and fisher spatial price index. This methodology improves the results obtained from just averaging regional housing price. The application of the methodology to Chilean data shows dramatic changes in the results. For example, Region II (Antofagasta) was 2 percent cheaper than the Metropolitan Region using simply average price, while region II was 18 percent more expensive than Metropolitan region according to the propose methodology. The different results are maintained for any pair of regions in the matrix indicating the contradictory results between simple average price and a regional housing price index that considers the geographical heterogeneity

Quasi-experimental control group methods applied to the regional housing price index allows reducing the bias caused by geographical heterogeneity and permits price comparisons among houses belonging to different regions and since a hedonic price set is constructed using comparable samples.

Particularly, this paper tested three different control group methods: 1) nearest matching on the propensity score, 2) Mahalanobis matching including the propensity score and 3) Mahalanobis matching within score calipers. The evaluation of these methods applied on Chilean data shows that the third one has the largest reduction of the average regional bias measured trough standardized differences. In addition, using Rosenbaum & Rubin (1985) the standardized differences under 10 percent can be considered not significant; our results comply with this criterion, because 90.63 percent of the covariates differences among regions were reduced. On the other hand, the average difference in propensity score changed from 0.12 to 0.2, indicating a good performance for this matching method.

The matrix of regional housing price indices obtained through the computation of Fisher spatial price index and it provides a better estimate than the simple average price (table 2).. Analyzing the average in table 2, the region with the more expensive housing is the metropolitan area. However, this comparison does not take in to account the geographical heterogeneity broadly discussed in this paper. Addressing this restriction through quasi-experimental control group methods, the regional housing price index shows that the metropolitan area does not have the most expensive housing. In this matrix, the most expensive housing is Region II (Antofagasta), some 18% above of the value in the metropolitan region. Moreover, using the superlative index property, it is possible to find the price index for any region. The difference between regions can be as much as 73 percent as in the case of Region II (Antofagasta) and Region VII.

Finally, this paper shows that the Regional Housing Price Index must include account for geographical heterogeneity. Otherwise, as in the case of the average price methodology, it can generate an inaccurate picture of the regional cost of housing, and, in turn, provide incorrect signals to national and regional government agencies whose policy initiatives center on reduction of regional disparities.

Table 5: Standardized Difference by Region Baseline

Region	Bedroom	Room Alternative use	Restroom	Kitchen	Quality Housing	Crowding	Education	Age	Difference Pscore
I	**-9.84**	-31.91	-39.53	-34.64	-101.58	16.25	-11.85	**-6.21**	0.24
II	-28.05	-25.91	-26.75	-26.86	-40.84	31.97	**0.32**	-22.91	0.06
III	**-1.74**	**-9.37**	-38.06	-14.02	-10.52	**1.64**	**-6.92**	-21.08	0.01
IV	**5.82**	-30.05	-22.00	-31.69	-41.73	**-8.39**	-14.76	**-5.15**	0.04
V	**4.63**	**5.32**	-17.46	**6.01**	24.84	**-4.42**	**-6.65**	**5.48**	0.05
VI	**2.57**	**3.86**	-42.65	**-3.23**	-81.98	**-0.34**	-35.92	**-4.86**	0.18
VII	**0.39**	**-9.80**	-31.69	-25.38	-66.15	**0.34**	-34.28	**3.04**	0.11
VIII	**-6.57**	-11.11	-33.50	-28.27	-63.96	**2.52**	-12.55	-10.08	0.16
IV	**-3.61**	-15.27	-32.73	-42.23	-50.78	**-9.68**	-19.43	**-9.41**	0.09
X	**6.07**	-42.00	-39.02	-58.57	-14.85	**-5.93**	-17.88	-24.51	0.12
XI	19.59	-44.83	-39.60	-61.40	45.06	-23.76	**-2.56**	-25.86	0.15
XII	**2.42**	-30.72	-28.28	-53.87	-96.97	**-2.43**	-21.89	**-4.78**	0.22

Bold Numbers are less than 10% in absolute value except for Difference Pscore

Table 6: Standardized Difference by Region Nearest Neighbor

Region	Bedroom	Room Alternative use	Restroom	Kitchen	Quality Housing	Crowding	Education	Age	Difference Pscore
I	-15.54	25.74	-13.48	41.48	-49.56	36.25	-46.14	15.09	0.05
II	**-3.24**	**3.37**	**3.54**	**-0.75**	**1.43**	**-3.21**	**5.47**	**-1.22**	0.00
III	**-2.96**	**4.62**	**-7.55**	**4.07**	**-1.45**	**-4.31**	**-2.93**	**8.44**	0.00
IV	**-4.93**	**4.66**	**-8.59**	**1.28**	**0.59**	10.83	**7.69**	-12.16	0.00
V	-27.51	-11.78	**0.00**	**-7.28**	14.40	17.77	**7.22**	-12.96	0.01
VI	-25.05	-11.15	**8.58**	**-3.88**	-40.69	25.20	**0.64**	18.95	0.07
VII	**-7.52**	**6.54**	**1.64**	11.80	-16.49	11.84	**2.33**	**5.59**	0.02
VIII	**-1.73**	**1.72**	**1.55**	**9.35**	-45.68	11.65	-25.01	20.71	0.07
IV	**2.36**	**-0.44**	**3.24**	**5.27**	**-8.37**	**5.12**	**-4.40**	14.09	0.00
X	**7.32**	**0.55**	**2.14**	**-4.48**	**-0.33**	**-5.07**	**-0.48**	**-0.39**	0.01
XI	**-2.55**	21.46	**2.21**	22.77	34.39	**-1.40**	13.83	**-1.66**	0.03
XII	**-4.37**	-11.63	**5.59**	**-5.02**	-18.10	**1.78**	-17.38	**8.79**	0.05

Bold Numbers are less than 10% in absolute value except for Difference Pscore

Table 7: Standardized Difference by Region Mahalanobis - Propensity Score covariate

Region	Bedroom	Room Alternative use	Restroom	Kitchen	Quality Housing	Crowding	Education	Age	Difference Pscore
I	**0.73**	**0.00**	10.48	**-2.28**	-43.82	**-1.48**	**3.88**	12.30	0.09
II	**0.91**	**-0.92**	**2.11**	**0.97**	**-4.57**	**1.41**	**0.12**	**0.04**	0.01
III	**0.74**	**0.00**	**0.00**	**0.00**	**6.36**	**6.52**	**-2.49**	**3.04**	0.00
IV	**4.37**	**1.53**	**2.80**	**0.00**	**1.48**	**-0.87**	**-2.96**	**2.92**	0.00
V	**-0.32**	**0.00**	**2.68**	**0.00**	**2.69**	**7.00**	**-5.55**	**0.03**	0.01
VI	**4.61**	**1.85**	**1.10**	**-1.13**	-28.52	**-1.18**	**-6.82**	10.34	0.04
VII	**3.51**	**-0.97**	**1.08**	**-2.54**	**-6.05**	**-3.31**	**1.85**	**2.92**	0.01
VIII	**3.46**	**0.67**	**1.48**	**-0.91**	-24.80	**1.04**	**-5.58**	10.69	0.05
IV	**7.82**	**0.97**	**4.84**	**-2.14**	-14.16	**-5.20**	**-0.42**	**3.49**	0.02
X	**4.77**	**-1.34**	**0.00**	**-0.79**	13.29	**-2.56**	**-0.92**	**-1.08**	0.01
XI	**1.34**	**0.00**	**-9.83**	**0.00**	**-5.57**	**-0.83**	**-7.36**	**1.34**	0.01
XII	**4.29**	**0.00**	**0.00**	**-3.19**	**-6.29**	**-7.69**	12.81	**-2.94**	0.02

Bold Numbers are less than 10% in absolute value except for Difference Pscore

Table 8: Standardized Difference by Region Mahalanobis - Propensity Score covariate. Caliper 0.2 * Variance Pscore

Region	Bedroom	Room Alternative use	Restroom	Kitchen	Quality Housing	Crowding	Education	Age	Difference Pscore
I	**-2.80**	**0.00**	**3.52**	**-3.48**	-42.77	**0.72**	12.29	11.50	0.06
II	**0.91**	**-0.92**	**2.11**	**0.97**	**-4.57**	**1.41**	**0.12**	**0.04**	0.01
III	**0.74**	**0.00**	**0.00**	**0.00**	**6.36**	**6.52**	**-2.49**	**3.04**	0.00
IV	**4.37**	**1.53**	**2.80**	**0.00**	**1.48**	**-0.87**	**-2.96**	**2.92**	0.00
V	**-0.31**	**0.00**	**2.66**	**0.00**	**1.44**	**7.78**	**-5.58**	**0.14**	0.01
VI	**7.24**	**2.25**	**0.00**	**0.00**	-15.33	**-8.59**	**2.70**	**5.52**	0.02
VII	**2.66**	**0.00**	**1.12**	**-1.28**	**-5.31**	**-3.06**	**3.18**	**4.73**	0.01
VIII	11.28	**0.87**	**4.58**	**-1.15**	-17.29	**-1.14**	**0.82**	**5.90**	0.03
IV	**7.82**	**0.97**	**4.84**	**-2.14**	-14.16	**-5.20**	**-0.42**	**3.49**	0.02
X	**3.77**	**-1.45**	**-0.84**	**0.00**	10.95	**-1.53**	**-1.92**	**-1.10**	0.01
XII	**1.34**	**0.00**	**-9.83**	**0.00**	**-5.57**	**-0.83**	**-7.36**	**1.34**	0.01
XII	**4.29**	**0.00**	**0.00**	**-3.19**	**-6.29**	**-7.69**	12.81	**-2.94**	0.02

Bold Numbers are less than 10% in absolute value except for Difference Pscore

Table 9: Hedonic Regression for Treatment and Control Housing

Variable/Region		I		II		III		IV		V		VI		VII		VIII		IX		X		XI		XII	
		Beta	t-ratio	Beta	t-ratio	Beta	t-ratio	Beta	t-ratio	Beta	t-ratio	Beta	t-ratio	Beta	t-ratio	Beta	t-ratio	Beta	t-ratio	Beta	t-ratio	Beta	t-ratio	Beta	t-ratio
TREATMENT	**Constant**	11.64	(92.38)	7.75	(72.42)	9.42	(73.34)	7.88	(87.72)	9.06	(118.7)	7.81	(90.32)	9.05	(188.6)	8.57	(111.8)	8.40	(153.2)	9.38	(103.4)	10.36	(50.07)	10.03	(127.3)
	Education	0.03	(12.19)	0.03	(19.60)	0.04	(26.30)	0.03	(19.13)	0.02	(16.12)	0.02	(13.46)	0.03	(28.03)	0.02	(15.36)	0.00	(-2.69)	0.03	(22.75)	0.02	(8.966)	0.05	(28.38)
	Civil Status	0.29	(26.07)	-0.04	(-6.33)	0.18	(18.86)	-0.16	(-18.7)	-0.01	(-2.62)	0.08	(9.905)	-0.03	(-5.78)	0.00	(0.626)	-0.11	(-16.9)	-0.04	(-6.10)	-0.02	(-1.57)	0.05	(6.525)
	Sex	0.05	(4.831)	-0.02	(-2.08)	0.02	(1.294)	0.12	(13.56)	0.00	(0.236)	-0.02	(-2.47)	-0.06	(-10.4)	-0.04	(-5.29)	0.05	(6.707)	0.01	(1.527)	0.03	(1.387)	0.17	(19.27)
	Age	0.00	(3.254)	0.00	(-6.54)	0.01	(15.28)	0.00	(9.289)	0.00	(7.573)	0.01	(26.20)	0.00	(8.380)	0.00	(-2.39)	0.00	(-3.94)	0.00	(-8.71)	0.01	(15.31)	0.00	(1.247)
	Bedroom	0.06	(14.62)	0.06	(16.14)	0.06	(10.48)	-0.06	(-13.7)	0.10	(38.09)	0.10	(24.74)	0.15	(51.13)	0.05	(13.10)	0.09	(25.12)	0.09	(38.02)	0.00	(0.063)	0.24	(48.68)
	Room Alt. Use	-0.01	(-0.68)	0.12	(9.648)	-0.06	(-3.39)	-0.10	(-6.76)	-0.16	(-18.0)	-0.02	(-1.49)	0.04	(6.078)	0.16	(21.59)	0.03	(4.393)	0.13	(15.72)	0.22	(11.26)	-0.18	(-16.6)
	Restroom	-0.13	(-10.1)	0.17	(25.44)	0.46	(34.93)	0.37	(37.50)	0.43	(93.62)	0.44	(47.92)	0.32	(48.92)	0.32	(47.09)	0.35	(45.80)	0.26	(34.75)	0.64	(27.69)	0.10	(13.80)
	Kitchen	0.45	(30.00)	-0.12	(-9.10)	0.16	(7.543)	0.17	(11.65)	0.12	(12.28)	0.23	(16.72)	0.12	(13.90)	0.01	(0.936)	0.08	(8.460)	-0.08	(-8.44)	-0.28	(-9.86)	-0.07	(-6.31)
	Quality	-0.02	(-8.93)	0.03	(24.72)	0.00	(-0.62)	0.02	(21.81)	0.01	(10.67)	0.02	(15.70)	0.00	(7.767)	0.02	(16.67)	0.02	(29.29)	0.01	(8.549)	-0.01	(-2.87)	0.00	(-2.07)
	Water Heater	0.44	(43.92)	0.27	(39.52)	0.25	(23.50)	0.47	(51.72)	0.27	(36.89)	0.20	(21.26)	0.33	(49.81)	0.25	(33.43)	0.28	(30.66)	0.37	(51.73)	0.36	(16.33)	0.13	(9.925)
	Phone	0.13	(11.07)	0.13	(17.58)	0.11	(10.44)	0.26	(24.66)	0.04	(7.117)	-0.02	(-1.69)	0.05	(6.665)	0.16	(21.74)	0.18	(24.57)	0.21	(25.60)	-0.12	(-7.27)	0.24	(25.97)
	Cable TV	-0.04	(-4.26)	0.01	(0.954)	0.11	(10.16)	-0.07	(-6.30)	0.20	(40.53)	0.10	(12.83)	0.18	(24.92)	0.19	(25.63)	0.18	(24.75)	0.03	(4.625)	0.30	(19.92)	0.15	(16.57)
	R2	0.53		0.49		0.64		0.53		0.58		0.63		0.71		0.64		0.65		0.63		0.77		0.85	
CONTROL	**Constant**	10.24	(60.64)	8.40	(75.86)	8.98	(72.58)	8.18	(112.6)	8.28	(87.88)	9.00	(103.9)	9.27	(189.6)	9.76	(123.6)	9.06	(129.7)	10.58	(113.4)	7.35	(26.69)	9.62	(72.64)
	Education	0.02	(8.325)	0.05	(33.01)	0.02	(12.58)	0.02	(18.77)	0.01	(3.880)	0.02	(12.48)	0.02	(19.28)	0.04	(34.16)	0.03	(17.65)	0.08	(58.62)	0.00	(0.742)	0.08	(27.29)
	Civil Status	0.56	(55.36)	-0.09	(-14.7)	-0.22	(-19.5)	-0.01	(-1.34)	-0.05	(-8.92)	-0.05	(-5.89)	-0.02	(-2.91)	-0.09	(-13.2)	0.12	(13.93)	-0.13	(-19.2)	-0.08	(-6.18)	0.33	(20.22)
	Sex	-0.35	(-30.7)	-0.08	(-11.9)	-0.06	(-5.39)	-0.02	(-2.51)	0.06	(9.241)	-0.17	(-17.6)	0.00	(-0.22)	-0.05	(-7.68)	0.09	(8.099)	0.06	(8.269)	0.09	(6.774)	-0.09	(-7.74)
	Age	0.00	(8.574)	0.01	(32.57)	0.00	(1.257)	0.01	(45.29)	0.00	(19.80)	0.00	(13.96)	0.00	(9.170)	0.01	(33.29)	0.00	(5.025)	0.01	(22.18)	0.00	(-5.57)	0.03	(42.79)
	Bedroom	-0.06	(-14.0)	-0.02	(-5.28)	0.13	(20.63)	-0.02	(-5.26)	-0.02	(-4.57)	0.03	(7.626)	0.08	(25.99)	0.06	(15.31)	0.07	(14.07)	-0.04	(-13.7)	0.01	(1.547)	-0.04	(-4.32)
	Room Alt. Use	0.15	(22.63)	0.37	(30.51)	0.15	(8.341)	-0.04	(-3.60)	0.46	(45.44)	0.14	(13.05)	0.21	(29.69)	0.03	(3.754)	0.02	(1.545)	-0.02	(-1.86)	0.20	(12.35)	0.39	(25.94)
	Restroom	0.49	(32.36)	0.40	(61.46)	0.07	(4.829)	0.45	(53.17)	0.43	(75.52)	0.56	(61.60)	0.28	(43.68)	0.30	(41.10)	0.21	(20.79)	0.34	(43.24)	0.61	(30.40)	0.35	(31.22)
	Kitchen	-0.20	(-13.6)	0.03	(2.115)	-0.12	(-5.48)	0.25	(20.56)	-0.15	(-13.3)	0.11	(8.049)	-0.07	(-7.81)	-0.07	(-6.18)	0.08	(5.844)	0.11	(11.00)	-0.19	(-7.93)	-0.16	(-11.9)
	Quality	0.00	(1.644)	0.01	(7.652)	0.02	(10.64)	0.02	(19.81)	0.02	(18.13)	0.01	(10.47)	0.01	(15.47)	0.00	(1.819)	0.01	(13.78)	-0.01	(-8.27)	0.04	(11.30)	-0.01	(-7.04)
	Water Heater	0.11	(8.996)	0.30	(42.00)	0.28	(21.36)	0.17	(21.70)	0.19	(24.34)	0.21	(25.16)	0.12	(16.41)	0.19	(24.37)	0.38	(34.55)	0.13	(16.21)	-0.07	(-4.04)	0.10	(6.691)
	Phone	0.11	(10.03)	0.10	(14.38)	0.20	(18.19)	0.17	(24.09)	0.24	(36.32)	0.16	(18.26)	0.16	(24.75)	0.14	(20.77)	-0.11	(-11.7)	0.22	(30.72)	0.29	(20.40)	0.48	(30.17)
	Cable TV	0.27	(18.21)	0.16	(22.65)	0.02	(1.705)	0.11	(13.88)	-0.06	(-9.88)	-0.07	(-7.75)	0.39	(50.30)	0.19	(26.61)	0.29	(29.29)	0.04	(4.441)	0.20	(14.55)	0.00	(0.131)
	R2	0.61		0.66		0.48		0.61		0.56		0.63		0.69		0.56		0.53		0.48		0.77		0.87	

REFERENCES

Barros, J. A., and Hernandez, J. (2007). Housing and Urban Location Decisions in Spain: An Econometric Analysis with Unobserved Heterogeneity. *Urban Studies , 44* (9), 1657-1676.

Basu, S., and Thibodeau, T. G. (1998). Analysis of Spatial Autocorrelation in House Prices. *Journal of Real Estate Finance and Economics , 17* (1), 61-85.

Can, A. (1990). The Measurement of Neighborhood Dynamics in Urban House Prices. *Economics Geography , 66* (3), 254-272.

Can, A., and Megbolugbe, I. (1997). Spatial Dependence and House Price Index Construction. *Journal of Real Estate Finance and Economics , 14*, 203-222.

CELADE. (1996). *Déficit Habitacional y Datos Censales Sociodemográficos: una Metodologia.* Santiago: Documento LC/DEM.R.267.

Cochran, W., and Rubin, D. (1973). Controlling Bias in Observational Studies: A review. *Sankhya , 35*, 417-446.

Costello, G., and Watkins, C. (2002). Towards a System of Local House Price Indices. *Housing Studies , 17* (6), 857-873.

D'Agostino, R. (1998). Propensity Score Methods for Bias Reduction in the Comparison of a Treatment to a Non-Randomized Control Group. *Statistics in Medicine , 17*, 2265-2281.

Diewert, W. E. (1976). Superlative and Exact Price Index. *Journal of Econometrics , 4*, 115-145.

Diewert, W. E. (1978). Superlative Index Numbers and Consistency in Aggregation. *Econometrica , 46* (4), 883-900.

Forret, D. (1991). An Analysis of House Price Differentials between English Regions. *Regional Studies , 25* (3), 231-238.

Gatzlaff, D., and Ling, D. (1994). Measuring Change in Local House Prices: An Empirical Investigation of Alternative Methodologies. *Journal of Urban Economics , 35*, 221-244.

Glasmeier, A., and Farrigan, T. (2007). The Economic Impacts of the Prison Development Boom on Persistently Poor Rural Places. *International Regional Science Review , 30* (3), 274-299.

Goodman, A. C. (1978). Hedonic Prices, Price Indices and Housing Markets. *Journal of Urban Economics , 5*, 471-484.

Griliches, Z. (1971). *Hedonic Price Indices for automobiles: an Econometric Analysis of Quality Change.* Massachusetts: Cambridge.

Hii, L., and Frei, F. (2002). Do Better Customers Utilize Electronic Distribution Channel? The Case of PC Banking. *Management Science , 48* (6), 732-748.

Mills, E. S., and Simenauer, R. (1996). New Hedonic Estimates of Regional Constant Quality House Prices. *Journal of Urban Economics , 39*, 209-215.

Paredes, D., and Aroca, P. (2008). Metodologia para Estimar un Indice Regional de Costo de Vivienda en Chile. *Latin American Journal of Economics , 45* (131), 129-143.

Rosen, S. (1974). Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. *Journal of Political Economy , 72*, 34-55.

Rosenbaum, P. R., and Rubin, D. B. (1983). The central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrica , 70* (1), 41-55.

Rosenbaum, P., and Rubin, D. (1985). Constructing a Control Group Using Multivariate Sampling Methods that Incorporate the Propensity Score. *The American Statistician , 39* (1), 33-38.

Rubin, D. (1976). Matching Methods that are Equal Percent Bias Reducing: Some Examples. *Biometrics , 32*, 109-120.

Rubin, D., and Thomas, N. (2000). Combining Propensity Score Matching with Additional Adjustments for Prognostic Covariates. *Journal of the American Statistical Association , 95* (450), 571-585.

Smith, H. (1997). Matching with Multiple Controls to Estimate Treatment Effects in Observational Studies. *Sociological Methodology , 27*, 325-353.

Thibodeau, T. G. (1989). Housing Price Indexes from the 1974-1983 SIMSA Annual Housing Surveys. *Journal of the American Real Estate and Urban Economics Association , 1* (1), 100-117.

Tritchler, D. (1995). Interpreting the Standardized Difference. *Biometrics , 51* (1), 351-353.

Xu, T. (2008). Heterogeneity in housing attribute prices: A study of the interaction behavior between property specifics, location coordinates and buyers' characteristics. *International Journal of Housing Markets and Analysis , 1* (2), 166-181.