The value of air quality and crime in Chile: A hedonic wage approach

Article *in* Environment and Development Economics · June 2011

DOI: 10.1017/S1355770X10000483 · Source: RePEc

3 authors:

Jorge Dresdner
University of Concepción

Felipe Lavín
University of Desarrollo

Renato Aguilar
University of Gothenburg

Environment and Development Economics 16: 329–355 © Cambridge University Press 2011
doi:10.1017/S1355770X10000483

The value of air quality and crime in Chile: a hedonic wage approach*

FELIPE VÁSQUEZ LAVÍN

Department of Economics, Universidad de Concepción, Casilla 160-C, Barrio Universitario, Concepción, Chile. Tel: 56 41/220 3203. Fax: 56 41/225 45 91. Email: fvasquez@udec.cl

JORGE DRESDNER

Department of Economics, Universidad de Concepción, Casilla 160-C, Barrio Universitario, Concepción, Chile. Email: jdresdne@udec.cl

RENATO AGUILAR

Department of Economics, Gothenburg University, Sweden. Email: Renato.Aguilar@economics.gu.se

Submitted November 1, 2009; revised July 15, 2010; accepted November 26, 2010; first published online 4 February 2011

ABSTRACT. We estimate the implicit prices of the crime rate and airborne pollution in Chile, using spatially compensating price differentials in the housing and labor markets. We evaluate empirically the impact of different estimation strategies for the wage and rent equations, on the economic value of these two amenities. The results show that increments in the crime rate or in air pollution have a negative impact on welfare and that the estimated welfare measures and their variances are sensitive to selection bias, endogenous amenities and clustering effects. In contrast, the welfare measures do not seem to be very sensitive to the simultaneity bias.

1. Introduction

The idea that market wages and rents tend to compensate for differences in local amenities is well established in the economic literature since the work of Rosen (1974, 1979, 1986) and Roback (1982). This idea has been used to value non-market amenities and to measure quality of life between locations in several countries (see Gyourko *et al.*, 1999). Moreover, to understand intercity migration, the value that people give to different amenities seems to be very important. However, there is surprisingly little evidence of compensating price differentials in less developed countries (LDCs). To our knowledge, there exist two studies that attempt to calculate an index of quality of life for LDCs using the compensating differentials approach: one for India (Madheswaran, 2007) and another for Colombia

* We are grateful to the Environmental Economics Unit at the University of Gothenburg, and especially to Thomas Sterner for his support in the initial stages of this collaborative work.

(Pérez and Arias, 2007). One reason for this lack of evidence in LDCs might be the data constraints existing in many of these countries.

In this paper, we examine the existence of price compensating differentials in the housing and labor markets originated by amenities in Chile. Based on countrywide data, we are able to estimate measures for the value of two important amenities in Chile: airborne pollution and crime. In this way, we add to the scanty evidence that exists for compensating differentials for LDCs.

However, measuring compensating differentials requires taking proper account of several methodological issues, such as the potential selection bias in the wage and housing price equations, the interrelation between the housing and labor market decisions, the possibility of endogenous amenities and the impact of clustered amenity data on the significance of the estimated parameters. These issues have been raised in the literature, although they have rarely been considered in the empirical implementation of the compensating differentials model for developed countries, and have not been considered in the LDCs studies. To our knowledge, there is no study that considers all these issues simultaneously. As it can be seen from the results we present in this study, the consideration of these issues is relevant because they have an important impact on the estimates of the value of amenities and their standard errors.

Four results can be derived from our application. First, there is evidence of selection bias both in the wage and housing prices (rent) equations, indicating that amenity value estimates that do not consider this could be biased. Second, our results suggest that proper valuation of amenities requires considering the correlation between error terms in the wage and housing price equations. Third, endogeneity affects significantly the magnitude and standard errors of the parameter and welfare measures. Fourth, limitations in information about amenities affect the significance of the parameters associated with these amenities due to clustering. Thus, our results shed some light on methodological issues that researchers should take into account if they want to extract meaningful information about the value of amenities using a hedonic wage model.

The next section briefly presents the theoretical frame for compensating differentials. A review of the Chilean case and the data used in this study follows. Then we discuss the econometric issues related to the model, and means of analyzing the empirical evidence. Finally, we present the results and close the paper with some conclusions.

2. Theoretical frame

The hedonic wage theory postulates that spatial differences in wages and land prices will reflect differences in the non-wage characteristics of the jobs and locations. In particular, different levels of amenities, *ceteris paribus*, should be reflected in the gap of equilibrium wages and land prices obtained by workers–consumers and firms with different preferences and technologies. These amenities might include characteristics that are completely exogenous to the economic agents, such as climate components (precipitation, temperature, humidity, etc.), as well as other characteristics that result from agents' decisions, such as urban conditions and environmental quality. When an individual chooses a place to live or work, he or she indirectly affects other aspects of the community, such as population density, average education and income, number of cars in the location, etc. In the same way, when firms make location decisions, they also affect some characteristics of the neighborhood through emissions in the environment, changes in labor demand, etc. Conditions like air quality and crime rates generate price differentials among different locations, to the extent that these characteristics vary among regions.

In the hedonic wage model, from the worker's perspective, living location and job decisions are connected. A decision to move to a job in another city will also be a decision to change residence to that city. If an environmental amenity does have an effect on job and housing location characteristics, then these decisions will not be independent. Theoretically, job and living location decisions are modeled together in the formal hedonic wage model (Rosen, 1979; Marin and Psacharopoulos, 1982; Blomquist *et al.*, 1988; Roback, 1988). The principal result of this process is that, in equilibrium, the derivatives of the wage and housing price function with respect to the environmental amenity can be used to measure the amenity's marginal (implicit) price (Roback, 1982).

The formal model used here comes from Roback (1982). We assume that workers–consumers maximize utility of a composite good (X), land, (l^c) and an amenity (s) subject to an income constraint made up of wages (w) and non-wage income (I). The price of X will be the *numeraire*, and the price of land is r. Spatial equilibrium of households can be characterized by

$$V(w, r; s) = c, \tag{1}$$

where $V(w, r; s)$ is the indirect utility function associated with this problem and c is a constant. The derivative of the utility function will be $V_s > 0$ ($V_s < 0$) if utility increases (respectively decreases) with amenity consumption, and we assume that $V_r < 0$ and $V_w > 0$ as usual.

On the other hand, firms maximize profits given a constant return to scale production function with land (l^p) and the total number of workers (N) as inputs. In equilibrium, the unit cost function for the firm follows:

$$C(w, r : s) = 1, \tag{2}$$

with $C_r = l^p/X > 0$, $C_w = N/X > 0$, and $C_s > 0$ if the amenity is unproductive and $C_s < 0$ if the amenity is productive.

Considering equations (1) and (2), we can solve the model for the equilibrium wage and land price levels. Roback (1982) shows that the implicit price (p_s^*) associated with the amenity s is

$$p_s^* = -\frac{V_s}{V_w} = l^c \frac{\partial r}{\partial s} - \frac{\partial w}{\partial s}. \tag{3}$$

The expected results for the land price and wage derivatives in this expression will depend on whether the amenity is productive or unproductive for firms and whether it is utility increasing or not for households.

This basic model has been amended in the literature in several ways. Hoehn *et al.* (1987) and Blomquist *et al.* (1988) incorporate intra-urban commuting costs, total amount of land available and population density into the model. In their model, the effects on the endogenous variables depend on a complex interaction of factors. In particular, the effects on $\partial w/\partial s$ and $\partial r/\partial s$ are ambiguous and depend on several assumptions about the effect of changes in the amenity level on firms and individuals, along with agglomeration effects. Ultimately, though, the estimation of the welfare measure is essentially the same as Roback (1982), and a pure consumption amenity is expected to have a positive price. Gyourko and Tracy (1991) expand the Roback model to include local produced amenities, other than housing. They argue for the importance of including local taxes and local public unions to obtain unbiased estimation of amenity prices. Gabriel *et al.* (2003) expand the basic two-equation specification to a three-equation system, with an additional relation for the price of other local produced amenities. Finally, Bayer *et al.* (2009) introduce migration costs in the wage hedonic model, and show that failing to take account of these costs might lead to biased estimates of amenity prices.

One important point to bear in mind in the empirical implementation of the model is that, from a theoretical point of view, land price and wage determination are interconnected. That is, individuals take decisions on living and working simultaneously, and this connects price determination in both markets. This should be reflected in the empirical estimation strategy. Furthermore, since some amenities are likely to be correlated with unobserved local economic factors that affect land prices, there is a potential endogeneity problem in the estimation of the model (Bayer *et al.*, 2009).

3. Compensation differentials in Chile and data description

There are few references on hedonic prices and amenities for the Chilean case. Figueroa and Lever (1992) estimated a hedonic price housing model using information from the house market in Santiago. Even though the model successfully predicted the value of houses, it did not include welfare estimates for any amenity. Wunder and Gutierrez (1992) proposed a strategy for evaluating and estimating Box–Cox models but, again, no amenities were considered. The only attempt to estimate welfare measures for amenities is Figueroa *et al.* (1996), which estimates a hedonic price function and obtains welfare measures for air quality (a 50% reduction in airborne pollution). However, there is no attempt to measure amenity prices with a wage hedonic model.

The Chilean economy has been a pioneer among LDCs in releasing markets for price determination. Already in 1973 consumer prices were released, and since then subsequent reforms have increased the role of market forces in determining equilibria (Ffrench-Davis, 2003). Specifically, the labor market was released in 1982 (Mizala and Romaguera, 2001), and the land market in 1983 (Aguilar and Dresdner, 2006). Thus, free markets have been playing a significant role in resource allocation in Chile for a long period, seen from the perspective of LDCs. In this sense, Chile constitutes a good case of a less developed country for trying to measure amenity prices through compensating price differentials.

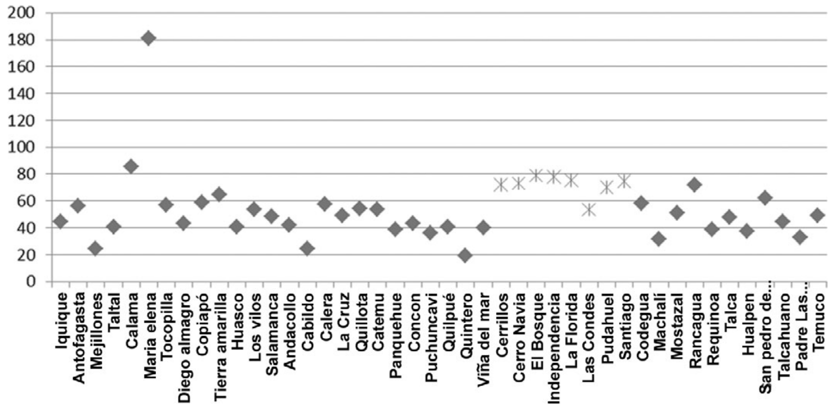

Figure 1. *Yearly average PM_{10} (mg/m^3) by commune in Chile, 2006*

According to the survey of the Centro de Estudios Públicos (CEP), in the year 2006, two of the most important issues in past years for Chilean citizens have been unemployment and crime (see CEP, 2006). Moreover, environmental concern has increased dramatically in the priority scale of Chileans in the last two decades. Therefore, it seems natural to include these variables in the analysis of compensating price differentials.[1]

However, although information on unemployment and crime is good, information on environmental conditions is more difficult to find. We had access to a data set with communal information on environmental conditions for the year 2006.[2] However, the availability of environmental data restricted the number of locations that can be included in the analysis. We use data on only 44 of the 225 communes in the country because these 44 are the only communes with regular monitoring stations collecting comparable environmental information. Specifically, we have data on air quality, defined in terms of particulate matter (PM_{10}).[3] This data were collected both from the Environmental Health Service for Metropolitan Regions (SESMA) and the National Committee for the Environment (CONAMA).

The available information on environmental conditions is distributed over very large distances,[4] but also with a concentration of some data points around the metropolitan area (Santiago) (see figure 1). The geographical dispersion helps to estimate more accurately the

[1] Potentially, other amenities could have been included, such as weather and climate conditions. However, we couldn't obtain this data on a communal basis.

[2] Communes in Chile constitute the smallest political–administrative units.

[3] Unfortunately, we haven't been successful in obtaining information on other pollutants for a significant number of communes. To our knowledge this information exists only for a small number of locations and, therefore, we could not explore the possibility of creating an index of pollutants.

[4] One should note that Chile is a very long and narrow country. Its continental length is approximately 4,300 km, with a medium breadth of 177 km.

compensating differentials, but at the same time, it increases the importance of migration costs. We have discussed the characteristics of the monitored communes with the person responsible for the research area of CONAMA. We have learned that there is no clear pollution pattern in the monitored communes, which is due to at least two reasons. First, the monitoring stations have been started in different years over a long period. In some cases, original high pollution levels have been drastically reduced, so the pollution levels in these communes were low in 2006. Second, some monitoring stations are privately owned and have been started as part of an investment project, to assure that the environmental norms are observed. So there is not necessarily a direct link between pollution levels and presence of a monitoring station. Third, financial restrictions have prevented the national authorities from placing public monitoring stations in all places they could suspect to find high pollution levels. Thus, there might be some places with a high level of pollution that do not have a monitoring system.

Data on the number of reported crimes per thousand inhabitants and the unemployment rates in each commune, as well as on individual and housing characteristics for workers and families, were obtained from the National Survey on Socioeconomic Characterization (CASEN) for the year 2006. This is a large survey developed by the Chilean government to monitor social conditions around the country, and therefore has statistical validity for all large cities. It includes information on a sample of households, including 268,873 individuals. It has data on family structure, housing, health, education, labor and income conditions of the families as well as individual data on wages, self-reported housing rents, labor and residential characteristics, and some firm characteristics. We do not have information on land or housing prices, so we use the available information on housing rents.[5] This information includes imputed self-reports made by housing owners. However, tenants and some housing owners do not report imputed rents in our data. To take account of this fact, in the estimations we control for potential selection bias in the rent equation.

The sample of individuals was sorted by age and commune of residence. We selected all individuals of labor age (18 years old or more) that lived in the communes with available environmental information. Finally, our sample was restricted to 27,833 individuals.

Since in general, communes in our sample are separate urban conglomerates, we do not consider agglomeration effects. Moreover, the Chilean tax system is highly centralized and there is very little role for local taxes on household budgets. The existing property taxes are very small and are calculated on administrative (non-market) property values. Local public unions have no power whatsoever in collective bargaining. Therefore, effects on amenity values as the ones proposed by Gyourko and Tracy (1991) do not seem to be relevant in the Chilean case. Finally, we do not have data on other local public goods (excluding housing), so we are

[5] There exists some information on housing transaction prices for some cities. But this information is incomplete and only exists for some of the communes covered in this study.

unable to estimate a model of three prices (wages, rents and the price of other public goods), as suggested by Gabriel *et al.* (2003). Therefore, we estimate the traditional Roback's two-price model.

Table 1 presents the list of all variables used in the estimations, a description of the main characteristics of the data, and their sources.

4. Econometric estimation

In order to calculate implicit prices for amenities, we must obtain estimates of $\partial w/\partial s$ and $\partial r/\partial s$. We estimate the outcome of both the labor and housing markets using the data described in the previous section. The model includes the following two equations:

$$\ln w_{ij} = \beta_0 + \beta_1 H_{ij} + \beta_2 A_j + \varepsilon_{ij}, \tag{4}$$

$$\ln r_{hj} = \gamma_0 + \gamma_1 Z_{hj} + \gamma_2 A_j + \xi_{hj}, \tag{5}$$

in which $\ln w_{ij}$ is the logarithm of the wage rate for individual i in location j, H_{ij} is a vector of individual characteristics for individual i in location j, $\ln r_{hj}$ is the logarithm of rents for house h in location j, Z_{hj} is a set of housing attributes, A_j is a vector of location characteristics, including amenities (s), such as crime rate and airborne pollution, $(\beta_0, \beta_1, \beta_2)$ and $(\gamma_0, \gamma_1, \gamma_2)$ are vectors of parameters to be estimated and ε_{ij} and ξ_{hj} are error terms.

A correct estimation of equation (4) requires solving the potential selection bias generated by the fact that we cannot observe wages for people in the sample that do not participate in the labor force (Killingsworth, 1983). Following Heckman's (1979) two-step model, a solution to this problem is to first estimate a labor participation equation using a probit model, and then from this estimation calculate the inverse Mills ratio, which is included as an additional explanatory variable in equation (4).

A similar correction is required in the rent equation as a consequence of the available data. Our data contain only imputed value for rent, and several people in the sample did not report any value for the rent of their house. This can give rise to a selection problem, if people that did not report rent values are, in a statistical sense, distinct from those that did report this value. In order to solve this selection problem we follow Berger *et al.* (2008) and estimate a two-step Heckman model, where in the first step we estimate a probit model for the decision of reporting a value for the house rent. The calculated inverse Mills ratio from this regression was included as an additional regressor, in equation (5). It is important to stress that not controlling for this selection bias potentially can distort the estimated amenity values.

The idea behind the wage hedonic model is that decisions in the labor and housing markets are connected and, as far as they involve interregional migration decisions, should be rather simultaneous. Thus, it seems natural to think of simultaneous estimation of the model represented in equations (4) and (5). Since Roback (1982) developed the model, many authors have applied it to estimate the value of environmental amenities. Most of them have ignored the possible correlation between the error terms in these

Table 1. *Variables and data source, Chile, 2006*

	Description	*Source*	*Mean*	*Min*	*Max*	*SD*	*Cases*
Dependent variables							
Log of wage (*lnw*)	Chilean pesos	CASEN	8.446965	3.101993	14.2089	0.8388726	15,129
Log of rent (*lnr*)	Chilean pesos	CASEN	9.052321	5.60993	12.29454	0.8278151	19,351
Employed (EMP)	Dummy	CASEN	0.5435634	0	1	0.4981076	27,833
Declared imputed rent (DECL)	Dummy	CASEN	0.6952538	0	1	0.4603081	27,833
Explanatory variables							
Age (AGE)	Number of years	CASEN	42.86847	18	102	17.25403	27,833
Education (EDUC)	Number of completed years	CASEN	10.06649	0	20	4.264671	27,795
Marital status (MS)	Dummy (1 if married, 0 otherwise)	CASEN	3.427442	1	9	2.648193	27,833
Head of household (HH)	Dummy (1 if head of household, 0 otherwise)	CASEN	0.3799806	0	1	0.4853904	27,833
Gender (SEX)	Dummy (1 if male, 0 otherwise)	CASEN	0.4819818	0	1	0.4996842	27,833
Tenure (TEN)	Dummy (1 if holds a long-term contract, 0 otherwise)	CASEN	0.339525	0	1	0.4735566	27,833
Social security (SS)	Dummy (1 if affiliated to social security, 0 otherwise)	CASEN	0.3831423	0	1	0.4861613	27,833
Number of rooms (ROOM)	Number of rooms in the dwelling	CASEN	4.854955	1	54	1.549501	27,819
Household size (SIZE)	Number of people in the household	CASEN	4.329788	1	17	1.924635	27,833
Water closet (WC)	Dummy (1 if dwelling has connection to sewing system, 0 otherwise)	CASEN	0.9186218	0	1	0.27342	27,833
Type of floor (FLOOR)	Dummy (1 if floor of higher quality material, 0 otherwise)	CASEN	0.6266662	0	1	0.4836983	27,833
Type of roof (ROOF)	Dummy (1 if roof of higher quality material, 0 otherwise)	CASEN	0.8791363	0	1	0.3259747	27,833

Type of wall (WALL)	Dummy (1 if wall of higher quality material, 0 otherwise)	CASEN	0.8040456	0	1	0.3969408	27,833
Household income (INCOME)	Monthly income of all household members in thousand of Chilean pesos	CASEN	651,439	0	9,907,434	777,389.8	27,833
Branch 1: agricultural and fishing (B1)	Dummy	CASEN	0.0090335	0	1	0.0946173	15,830
Branch 2: mining (B2)	Dummy	CASEN	0.1475047	0	1	0.3546196	15,830
Branch 3 (B3)	Dummy	CASEN	0.0584965	0	1	0.2346874	15,830
Branch 4: electricity, gas and water (B3)	Dummy	CASEN	0.1197726	0	1	0.3247057	15,830
Branch 5: construction (B4)	Dummy	CASEN	0.0075174	0	1	0.086379	15,830
Branch 6: commerce (B5)	Dummy	CASEN	0.097789	0	1	0.2970385	15,830
Branch 7: Transp., storage and comm. (B6)	Dummy	CASEN	0.1862919	0	1	0.3893543	15,830
Branch 8: financial services (B7)	Dummy	CASEN	0.0724574	0	1	0.2592519	15,830
Branch 9: social and com. services (B5)	Dummy	CASEN	0.0637397	0	1	0.2442964	15,830
Branch 10 (B5)	Dummy	CASEN	0.2373973	0	1	0.4255012	15,830
Particulate material (PM_{10})	Communal average over the year (ug/m^3)	CONAMA	55.67158	19.06	180.75	22.64901	44
Particulate material, winter (PM10WIN)	Communal average over April–Sept. (ug/m^3)	CONAMA	64.3593	18.7	207.895	28.33963	44
Days over the norm (PM10ABOVE)	Number of days PM_{10} is over the norm	CONAMA	5.391436	0	35	7.843964	44
Crime (CRIME)	Average communal rate over the year (total crimes per 1,000 inhabitants)	SINIM	21.82481	0	50.46744	9.665627	44
Unemployment rate (UR)	Average communal rate over the year (percentage)	SINIM	7.529407	3.351955	12.22707	2.161192	44

Population (POP)	Number of inhabitants (in hundred thousands)	SINIM	0.8408757	0.0323	2.92492	0.7879838	44
Urban (URBAN)	Dummy (1 if urban, 0 otherwise)	SINIM	0.8505012	0	1	0.3565859	44
Urban area (AURBAN)	Percentage of urban and industrial area	SINIM	0.1930211	0	1	0.3040081	44
Agricultural area (AGAREA)	Percentage of agricultural land	SINIM	0.1641975	0	0.8582	0.2241208	44
Grassland area (GAREA)	Percentage of grassland and shrubbery land	SINIM	0.3148686	0	0.9776	0.2871101	44
Forest area (FAREA)	Percentage of forest land	SINIM	0.0721515	0	0.319	0.0898044	44
Land without vegetation (NVEG)	Percentage of area without vegetation	SINIM	0.2369752	0	1	0.3529011	44
Water expenditure (CWATER)	Percentage of communal water expenditure over total communitary services expenditures	SINIM	0.0411357	0.0006	0.1829	0.0380607	44
Public lighting expenditure (ALUM)	Percentage of public lighting expenditure over total communitary services expenditures	SINIM	0.2536435	0.1067	0.6155	0.1111005	44
Cleaning services expenditures (CCLEAN)	Percentage of communal cleaning services expenditures over total communitary services expenditures	SINIM	0.3404905	0	0.546	0.1396106	44
Parks and forest expenditures (PARKS)	Percentage of communal expenditure in parks over total communitary services expenditures	SINIM	0.1491507	0	0.3572	0.0910847	44

equations.[6] The usual practice has been to estimate two separate equations, one for wage and another for rent, using a linear ordinary least squares (OLS) or a nonlinear Box-Cox transformation. Some examples of such applications include papers by Roback (1982, 1988), Hoehn *et al*. (1987), Blomquist *et al*. (1988), Clark and Kahn (1988), Deller and Ottem (2001), Gabriel *et al*. (2003), Deitz and Abel (2008) and Buettner and Ebertz (2009), among others.

For LDCs, Madheswaran (2007) and Pérez and Arias (2007) do not consider the decision problem of housing location. They calculate willingness to pay functions for environmental amenities by estimating the hedonic wage model using only an equation for earnings. That is, they estimate only partial willingness to pay functions.

A joint estimation of the two-equation system is important, because it allows a more efficient estimation of the standard errors of the parameters and gives us the covariance between estimators in both equations. This covariance is required in order to evaluate the statistical significance of the welfare measure associated with an amenity. Earlier applications of the model do not provide any estimation of the welfare measure's variance or ignore the covariance effect.

Even when the explanatory variables from one equation are truly exogenous to the other equation, one should attempt to model the interrelation between the rent and wage determination. Thus, one can hypothesize that the model is a seemingly unrelated simultaneous equations system (SUR). If we assume that the disturbances are normally and identically distributed with

$$\begin{bmatrix} \varepsilon_{ij} \\ \varepsilon_{hj} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{bmatrix}\right), \tag{6}$$

then our model is similar to a bivariate Tobit model.[7]

This model implicitly assumes that the mechanism (regressors) determining whether or not a person is employed is the same as the mechanism that determines wages, and that the mechanism that determines whether a person self-reports rent for a house is the same as the mechanism determining rents. In order to avoid these constraints, besides using the bivariate Tobit model, we also use Heckman's (1979) two-step procedure model in which the mechanism determining employment (or rents self-report) can differ from the mechanism determining wages (or rents). To apply this procedure, with the simultaneous equations, we estimate the model in two steps. In the first step we estimate a participation equation and a reporting equation using a probit model for the whole sample. These equations permit us to obtain the inverse Mills ratio for each case, denoted as λ_i and κ_h, which are used as additional explanatory

[6] An exception is Berger *et al.* (2008). They estimate a seemingly unrelated system for the wage and housing price equations. Our estimation strategy differs from their paper as we additionally control for selection bias in the wage equation with simultaneous estimation, and for potential endogeneity in the amenities.

[7] The Tobit structure emerges from the correction of the selection problem discussed previously.

variables in the wage and rent equations, respectively. In the second step, SUR is applied simultaneously to both equations, but only for individuals with positive earnings and rents. The equations for the second step are

$$\begin{aligned} lnw_{ij} &= \beta_0 + \beta_1 H_{ij} + \beta_2 A_j + \beta_3 \lambda_i + \varepsilon_{ij}, \\ lnr_{hj} &= \gamma_0 + \gamma_1 Z_{hj} + \gamma_2 A_j + \gamma_3 \kappa_h + \xi_{hj}. \end{aligned} \tag{7}$$

To summarize, we used three different estimation models: (1) Heckman's two-step model, using OLS separately for each equation in the second step; (2) Heckman's two-step model, using a joint estimation of the wage and rent equations (SUR model) in the second step; and (3) a Tobit simultaneous equation system (BTobit model). The last two models enable us to estimate the covariance among parameters in both equations.

Two other issues are considered in the estimation process. First, following Bayer *et al.* (2009), we considered the possible endogeneity of the amenities in each of these three estimation approaches. Bayer *et al.* (2009) claim that some amenities might be correlated with 'unobserved local economical factors', which also affect the explanatory variables, generating an endogeneity problem. We focus our analysis on the possibility that the variables crime rate, particulate matter and unemployment rate could be correlated with the error terms of the corresponding wage or rent equations. Another possible source of endogeneity could come from a simultaneity bias, since unobserved components from omitted variables, such as school quality and weather conditions, could affect wages or rents and amenities at the same time. We test the exogeneity hypothesis using the Hausman test. This test has been discussed in the literature, for example in Wooldridge (2002). Then, we use instrumental variable estimation for the three models mentioned above; that is, we estimate a two-stage least squares (2SLS), three-stage least squares (3SLS) and a simultaneous Tobit model with endogeneity.

Second, we address the effect of clustering associated with amenities. Clustering occurs because the information about amenities is, generally speaking, at the city level, while the equations are estimated using micro-level data (consumer level). This cluster effect reduces the precision of the estimates and the significance of the parameter associated with amenities (Moulton, 1990).

Finally, the last issue considered in this section involves the calculation of the welfare measures and their variance. Our estimated equations differ from those by Roback (1982) because we estimate the logarithm of the hourly wage income instead of the total wage income, and the logarithm of land rental cost per month instead of the per area unit rental. For empirical calculations, we adjust equation (3), and the welfare measure is given by

$$\hat{p}_s^* = \hat{\gamma}_s \bar{r} - \hat{\beta}_s \cdot \bar{w} \cdot \bar{n}. \tag{8}$$

In this formula, r is the monthly rental cost per household, $\hat{\gamma}_s$ is the derivative of the logarithm of rents with respect to amenity s, w is the individual labor income, n is the number of employed persons per household and $\hat{\beta}_s$ is the derivative of the logarithm of wages with respect to amenity *s*. A bar over a variable indicates that it is the average over the sample and a hat indicates the estimated value. Note that $\hat{\gamma}_s$ and $\hat{\beta}_s$ can be obtained directly from the estimations of equation (7). The simultaneous estimation methods give us the estimated covariance between $\hat{\gamma}_s$ and $\hat{\beta}_s$, which is used to calculate the variance of expression (8).

5. Econometric results

To summarize our estimation strategy, we specify and estimate a two-equation model with censored dependent variables using household information; one equation explains rents while the other explains earnings. We evaluate the effects that different treatments of the error correlation structure of these equations have on welfare measures and also evaluate the applicability of this model to a developing country with limited data (clustered data) on environmental amenities. The equations are estimated with and without consideration of endogeneity of amenities in the model.

Moreover, we have estimated the models using different ways to measure the air quality variable. Specifically, we have made regressions using yearly average PM_{10} (μg/m^3), which is the average of all daily values of PM_{10} measures from all monitoring stations within the commune, winter average PM_{10}, which is the same measure but only for the winter period between April and September, and the number of days in the year over the norm (which is 150 μg/m^3). The qualitative results do not differ between alternative measures.[8]

5.1. *'Naïve' estimation*

The first set of results presents the 'naïve' estimators, that is, the results for the regression analysis that does not take into account the possible endogeneity of the amenity variables. The first method is the no-correlated two-step Heckman model for both wage and rent equations. First we estimate participation equations in the labor market and the housing market (see tables A1 and A2 in the appendix available at http://journals.cambridge.org/EDE). The former controls for selection bias in the employed, and the latter represents an equation explaining who in the sample is more likely to provide information about housing rents.[9] From these two equations we calculate two inverse Mills ratios, one for each equation (denoted as lambda and kappa in our regressions, respectively). In the second step we estimate independently an OLS regression for each dependent variable incorporating the Mills ratio as an explanatory variable. The second method estimates the same equations

[8] We include tables A3 and A4 with welfare measures in the Appendix (available at http://journals.cambridge.org/EDE) for each of these last two cases. The full set of results is available upon request.

[9] The results for the labor participation equation, in general, are significant and have the expected signs. See tables A1 and A2 in the Appendix (available at http://journals.cambridge.org/EDE). Perhaps the only result that deserves a comment is the nonsignificant result for civil status. This is presumably due to the weight that married women have in the inactive labor force, and the low probability they have to participate in the labor force, once married.

but using a SUR model in the second step, that is, considering the correlation of the error terms. Finally, we estimate a BTobit model in which both participation and intensity decisions are estimated jointly and determined by the same explanatory variables.[10] Results are presented in tables 2 and 3. In the wage equation (table 2), the dependent variable is the logarithm of the hourly wage rate (*lnw).* Explanatory variables are age (*AGE*), age squared (AGE^2), years of education (*EDUC*), head of household (*HEAD*), gender (*SEX*), tenure (*TEN*, a dummy variable that takes the value of one when people have a long-term contract) and social security (SS, a dummy variable taking the value one if the individual is affiliated to a social security company).[11] In addition, we have particulate matter in the commune (PM_{10}), communal crime rate (*CRIME*), communal unemployment rate (*UR*), total population (*POP*, in thousands of people) and population square (POP^2), urban commune (*URBAN*, a dummy variable taking the value one if the commune is categorized as urban according to the 2002 population census), communal vegetation (*VEG*, measured as the percentage of the commune's surface without grassland or shrubbery land), water expenditure (*CWATER*, measured as the percentage of communal water expenditure over total communitarian services expenditures), public lighting expenditure (*CALUM*, measured as the percentage of communal street lighting expenditure over total communitarian services expenditures), cleaning services expenditures (*CCLEAN*, measured as the percentage of communitarian cleaning services expenditures over total communitarian services expenditures), parks and forest expenditures (*CPARKS*, measured as the percentage of communitarian forest and parks expenditures over total communitarian services expenditures) and the estimated inverse Mills ratio (*LAMBDA*) as explanatory variables in the wage equation. PM_{10} and *CRIME* correspond to the evaluated amenities. Moreover, there is a presumption that labor and housing markets adjust slowly, especially in LDCs. Following Berger *et al.* (2008), to consider the possibility of adjustment costs, we included *UR* in the empirical specification as a control for disequilibrium situations. We include *POP* and POP^2 as proxy variables to control for unmeasured commune scale-related amenities and disamenities. Finally, we include *URBAN*, *VEG*, *CWATER*, *CALUM*, *CCLEAN* and *CPARKS* as control variables for communal attributes that might affect the agents' location decisions.

[10] To run the SUR model we had to match the observations in the wage and rent equations. After the matching, the total sample was reduced to 9,633 individuals.

[11] In principle each dependent worker in Chile is automatically affiliated to a social insurance company by his employer. However, independent workers must affiliate themselves to these companies. Some of these workers affiliate and others don't. Moreover, informal workers (without a formal labor contract) are usually not affiliated. Thus, some workers are affiliated and others are not. Usually, this variable is used as a measure of the quality of the activity/firm which generates the job opportunity. Formal works are better than informal ones. We include this variable to pick up potential heterogeneity between these jobs/workers.

Table 2. *Wage equation for Chile, 2006*[a] *(Dependent variable: log of hourly wages)*

Estimation method	*OLS*		*SUR*		*TOBIT*	
Variable	*Coeff.*	*t*	*Coeff.*	*t*	*Coeff.*	*t*
AGE	0.079	6.150	0.076	7.630	0.513	41.950
AGE^2	−0.001	−5.160	−0.001	−6.380	−0.006	−44.090
EDUC	0.125	31.640	0.125	39.390	0.104	11.050
HEAD	0.339	8.220	0.351	10.770	1.408	17.040
SEX	0.363	6.530	0.356	8.520	2.444	32.910
TEN	−0.356	−12.740	−0.370	−17.720	4.453	41.700
SS	0.172	6.000	0.171	7.860	2.511	23.630
PM_{10}	0.003	5.240	0.002	4.200	5.100E−05	0.030
CRIME	−3.72E−04	−0.340	−0.002	−1.820	0.001	0.270
UR	−0.050	−11.660	−0.050	−9.570	−0.086	−5.170
POP	0.110	2.830	0.095	2.230	0.502	3.200
POP^2	−0.032	−2.220	−0.037	−2.670	−0.099	−1.620
URBAN	0.089	2.930	0.119	2.520	−0.246	−2.500
VEG	−0.034	−0.700	−0.167	−2.950	−0.475	−3.040
CWATER	−1.304	−2.500	−0.038	−0.070	4.105	2.490
CALUM	−1.565	−9.210	−1.542	−9.190	−2.397	−4.380
CCLEAN	−1.270	−11.790	−1.129	−10.680	−0.478	−1.450
CPARKS	−0.382	−2.640	−0.325	−2.080	−2.141	−3.890
LAMBDA	0.575	4.140	0.542	5.300	–	–
CONSTANT	5.761	14.020	5.806	17.770	−10.598	−22.830
	N	14,832	N	9,633	N	27,310
	F(19; 14,812)	178.2	chi-squared	5,053.85	Wald	17,040.34
	R-squared	0.3404	R-squared	0.3464	Likelihood	−119,594.13

[a]Robust standard errors for all equations and marginal effects reported for tobit model.

Table 2 presents the outcomes of the estimations of the wage equation, expressed in terms of marginal effects, for each of the different treatments of the correlation between error terms in the equations.[12] Most marginal effects show a rather high level of significance, except for *CRIME* in the OLS case, and PM_{10}, *CRIME* and cleaning services expenditures in the BTobit case. All marginal effects associated with the Mincerian wage model were estimated at satisfactory levels of significance and have the expected sign. Moreover, the magnitude of the estimated returns to education is similar to previous finding for the Chilean economy (see e.g., Fuentes *et al.*, 2005).

The marginal effect corresponding to the environmental variable (PM_{10}), when significant, was estimated positive as expected. A positive estimation for this effect is consistent with an unproductive amenity in the Roback model and in line with previous empirical results (see Roback, 1982; Hoehn *et al.*, 1987; Blomquist *et al.*, 1988; Berger *et al.*, 2008). However, the marginal effect associated with *CRIME* was not significantly estimated in all the equations.

The unemployment rate, estimated at a reasonable level of significance in each model, has a negative effect. This indicates that higher communal unemployment will have a negative effect on communal wages. In our interpretation, this is to be expected if the labor market is functioning correctly, since this means that the incentives are to move from the high unemployment areas to the low unemployment areas, thus equalizing wage and unemployment differences. The population controls are also significant, showing a positive nonlinear relation between wages and population size.

The lambda parameter, indicating the sample selection effect, is positive, relatively large and has a rather high level of statistical significance. This gives evidence that sample selection is an important issue to consider for wage determination in the Chilean case.

The dependent variable of the rent equation (table 3) is the logarithm of rental monthly costs (*lnr*), and the explanatory variables are the number of people in the family (*NUMPER*), number of rooms in the residence (*ROOM*), years of education (*EDUC*), water closet (*WC*, a dummy variable that is one if the house has connection to a sewer system and zero otherwise), floor (*FLOOR*, a dummy variable that is one if the house has a floor of higher quality material and zero otherwise), roof (*ROOF*, a dummy variable that is one if the house has a roof of higher quality material and zero otherwise), wall (*WALL*, a dummy variable that is one if the house has walls of higher quality material and zero otherwise), household income (*INCOME*), particulate matter in the commune (PM_{10}), communitarian crime rate (*CRIME*) communitarian unemployment rate (*UR*), total population (*POP*), population square (POP^2), urban commune (*URBAN*, a dummy variable taking the value one if the commune is categorized as urban according to the 2002 population census), communal vegetation (*VEG*, measured as the percentage of the commune's surface without grassland and shrubbery land), water expenditure (*CWATER*, measured as

[12] All equations are estimated using robust standard errors.

Table 3. *Rent equation for Chile, 2006*[a]

	OLS		SUR		TOBIT	
Estimation method	*Coeff.*	*t*	*Coeff.*	*t*	*Coeff.*	*t*
NUMPER	−0.268	−71.900	−0.261	−92.310	−0.015	−0.770
ROOM	0.058	5.450	0.046	13.980	0.457	18.560
EDUC	0.036	24.830	0.041	29.150	−0.191	−21.420
WC	0.268	8.770	0.246	6.190	0.932	6.000
FLOOR	0.188	14.880	0.165	12.130	0.924	10.920
ROOF	0.111	6.000	0.116	6.320	0.350	3.000
WALL	−0.055	−3.630	−0.030	−1.790	1.096	10.900
INCOME	1.65E − 07	22.450	1.47E − 07	30.600	5.12E − 07	10.290
PM_{10}	0.004	9.540	0.003	8.340	−0.052	−27.120
CRIME	0.002	2.970	0.002	3.020	0.005	1.030
UR	−0.017	−6.390	−0.017	−5.070	0.150	8.450
POP	0.287	13.030	0.300	10.920	0.670	3.930
POP^2	−0.085	−10.460	−0.086	−9.680	−0.038	−0.570
URBAN	0.086	3.760	0.026	0.780	−0.024	−0.210
VEG	0.233	6.710	0.177	4.850	−1.782	−10.290
CWATER	−1.871	−5.440	−1.586	−4.310	6.486	3.540
CALUM	−1.839	−19.090	−1.870	−17.040	0.863	1.490
CCLEAN	−1.357	−20.640	−1.450	−20.960	0.645	1.830
CPARKS	−1.098	−12.690	−1.136	−11.240	−1.420	−2.410
KAPPA	−0.668	−19.400	−0.632	−18.970	–	–
CONSTANT	10.201	117.370	10.258	118.070	3.176	7.400
	N	18972	N	9,633	N	27,310
	F(20; 18,951)	874.55	chi-squared	18,288.83	Wald	17,040.34
	R-squared	0.4336	R-squared	0.6581	Likelihood	−119,594.13

[a]Robust standard errors for all equations and marginal effects reported for tobit model.
Dependent variable: Log of monthly rent.

the percentage of communal water expenditure over total communitarian services expenditures), public lighting expenditure (*CALUM*, measured as the percentage of communal street lighting expenditure over total communitarian services expenditures), cleaning services expenditures (*CCLEAN*, measured as the percentage of communal cleaning services expenditures over total communitarian services expenditures), parks and forest expenditures (*CPARKS*, measured as the percentage of communal forest and parks expenditures over total communitarian services expenditures) and the inverse Mills ratio (*KAPPA*). The variables *WC*, *FLOOR*, *ROOF* and *WALL* are indicators of the quality of the house. Thus, their expected values are all positive.

Table 3 presents the results of the estimations for the rent equation. All marginal effects are estimated with a rather high level of statistical significance except for urban commune in the SUR model, and crime, population square (POP^2), urban commune and public lighting expenditures in the BTobit model. Moreover, the *KAPPA* variable is significant, indicating that the control for selection bias is relevant in this case. Note that the amenities, both airborne particulate matter (PM_{10}) and *CRIME*, are estimated positive and with a high level of statistical significance in the OLS and SUR models. However, particulate matter and crime change sign and are nonsignificant, respectively, with the BTobit model.

Using the results in tables 2 and 3, we estimate the monthly marginal value of PM_{10} and *CRIME* using equation (8). The mean monthly rent is around CLP 43,869 and the mean level of monthly earnings is CLP 246,123.[13] Table 4, panel A, presents the marginal value of PM_{10} and *CRIME* and their standard deviations measured by the different estimation methods. In the OLS model, we assume zero covariance since no covariance can be obtained from the econometric estimation. In all cases, the marginal value of PM_{10} is negative, as we expect for a disamenity.[14] According to these estimates, the monthly value of a marginal improvement in air quality lies between CLP 492 and CLP 1,299 (US$ 1–3). The estimated marginal value of *CRIME* is only significantly different from zero at 5% for the SUR model, but with an unexpected positive sign.

The results show that simultaneous estimation with SUR does not change significantly in comparison with the OLS estimates. This can be seen by comparing the results of the estimated amenity prices between these two methods. The differences emerge when we compare these results with those obtained with the BTobit model. There are two econometric differences between the SUR model and the BTobit model that can explain

[13] In 2006, US$ 1 was around CLP 530.

[14] We run the regressions with PM_{10} values for the winter season (*PM10WIN*) and number of days PM_{10} is over the norm within the year (*PM10ABOVE*). With *PM10WIN* the results are qualitatively the same. With *PM10ABOVE* the results are the same, except for the OLS estimate for the marginal value of PM_{10}, which is not significant at traditional confidence levels. See tables A5 and A6 in the Appendix available at http://journals.cambridge.org/EDE.

Table 4. *Marginal value (in pesos) of* PM$_{10}$ *and crime in Chile, 2006*

	Panel A: not controlled by endogeneity			*Panel B: controlled by endogeneity*		
		PM$_{10}$				
	OLS	*SUR*	*BTOBIT*	*2SLS*	*3SLS*	*BTOBIT_end*
Wage equation parameter	2.65E − 03	2.49E − 03	2.90E − 05	6.97E − 03	1.45E − 02	7.46E − 04
Variance	2.56E − 07	3.53E − 07	9.37E − 07	2.77E − 06	4.14E − 06	5.02E − 06
Rent equation parameter	3.67E − 03	3.22E − 03	−2.94E − 02	−3.30E − 02	1.25E − 02	−3.07E − 02
Variance	1.48E − 07	1.49E − 07	1.18E − 06	2.24E − 04	4.14E − 06	6.46E − 06
Covariance	0	1.93E − 08	−2.90E − 08	0	6.22E − 07	−1.91E − 07
Marginal value	**−491.86**	**−472.61**	**−1,299.02**	**−3,162**	**−3,027**	**−1,528**
Standard deviation	125.77	145.73	244.27	774	495	566
t-value	−3.91	−3.24	−5.32	−4.09	−6.11	−2.70
			Crime			
	OLS	SUR	BTOBIT	2SLS	3SLS	BTOBIT_end
Wage equation parameter	−3.72E − 04	−2.28E − 03	7.27E − 04	−1.94E − 02	5.95E − 02	3.78E − 02
Variance	1.20E − 06	1.57E − 06	7.25E − 06	2.62E − 05	9.12E − 05	2.16E − 04
Rent equation parameter	1.94E − 03	2.44E − 03	2.91E − 03	−5.84E − 02	4.37E − 02	−1.81E − 01
Variance	4.25E − 07	6.51E − 07	7.96E − 06	1.33E − 03	4.37E − 05	2.77E − 04
Covariance	0	8.60E − 08	−2.66E − 07	0	1.37E − 05	−8.38E − 06
Marginal value	**176.5**	**668.0**	**−51.5**	**2,201.2**	**−12,737.1**	**−9,307.6**
Standard deviation	270.94	307.31	678.50	2,035.62	2,304.75	3,615.27
t-value	0.65	2.17	−0.08	1.08	−5.53	−2.57

the divergences. First, the BTobit model should be more efficient than the SUR model that has been corrected by selection bias using a two-step Heckman model. In contrast, the BTobit model was estimated using maximum likelihood estimation in just one step. Of course, this fact does not affect the consistency of the SUR estimates. The second difference is that the explanatory variables in the participation equation (first stage) differ from the explanatory variables in the wage and rent equations (second stage) for the SUR model. This is not the case in the BTobit model, in which both decisions use the same explanatory variables. This difference can explain the changes in the estimated amenity prices between the SUR regression and the BTobit regression.

5.2. Endogeneity issues

The results presented in the previous section are not corrected by endogeneity. Bayer *et al.* (2009) claim that amenity variables might be correlated with the error term because of some unobserved effect over economic variables that also affect the dependent variables, i.e., wage and rents. If this is the case, then the estimated parameters might be biased. We focus our analysis on the possibility that the crime rate, particulate matter and unemployment rate variables could be correlated with the error terms of the corresponding wage or rent equations. We tested for endogeneity of these variables using the Hausman test (Wooldridge, 2002) and found that the joint null hypothesis that the amenity variables are exogenous was rejected at 1% for PM_{10}, *UR* and *CRIME* in both the wage and rent equations. Therefore, we used an instrumental variable estimation procedure to control for this endogeneity. To select instruments we used the strategy chosen by Timmins and Murdock (2007). We selected exogenous amenities in neighboring communes as instruments. These variables should be correlated with particulate matter, crime and unemployment in each commune. When agents choose to settle in one commune, they consider, at least partially, the amenity levels in all neighboring communes. Thus, for example, a lower value of public lighting in commune j will induce more agents to settle in commune i and this will affect the equilibrium values of particulate matter, crime and unemployment in this commune. However, the value of amenities of other communes will not affect the determination of wages and rents in the selected commune i. Therefore, they should not be included in the structural equations that determine these variables. Moreover, there is no reason why the values of these amenities should be correlated with unobserved local amenities in commune i. In fact, to the degree that they are truly exogenous, as assumed, they should be uncorrelated with the error terms in the wage and rent equations. Thus, the conditions for valid instruments, at least in principle, are fulfilled.

To construct the instruments we calculated the average value for these explanatory variables in all the neighboring communes. For the wage equation we used variables urban commune (*URBAN*), communal vegetation (*VEG*), public lighting expenditure (*CALUM*), parks and forest expenditures (*CPARKS)* of the neighboring communes to construct instruments, while for the rent equations the variables communal vegetation (*VEG*), public lighting expenditure (*CALUM*), parks and forest expenditures (*CPARKS*), and water expenditure (*CWATER)* were used.

Notwithstanding, we also used two empirical procedures to test whether the instruments were correctly chosen. When the model is overidentified, it is possible to use the Sargan–Hansen test (Cameron and Trivedi, 2006), which tells us whether the chosen instruments, Z, are correlated with the error term ε. The null hypothesis is $E(Z\varepsilon) = 0$. If we reject the hypothesis, then our instrumental variables (IV) estimation is inconsistent.[15]

Moreover, to test for instrument relevance, we used the Anderson canonical correlation test (Hall *et al.*, 1996), whose null hypothesis is that all the coefficients of the instruments are equal to zero in the auxiliary regression equation. The results obtained for the Sargan–Hansen and Anderson tests suggest that our instruments are relevant and uncorrelated with the error terms.[16]

Then we estimated the 2SLS, the 3SLS and the BTobit models corrected by endogeneity. In the last case we estimated the auxiliary equations and used the predicted value of amenities in the BTobit regression.

Table 4, panel B, presents the estimated marginal amenity prices considering endogeneity.[17] The main effect of correcting for endogeneity was an increase in the absolute value for the marginal effects in the parameters of the wage and rent equations, together with a reduction in the standard error and higher statistical significance. This turned into significantly higher marginal value estimates for both airborne pollution and crime in almost all cases, as can be seen in table 4. The only nonsignificant result appeared in the case of crime in the OLS model. In comparison with the results not corrected for endogeneity, now all significantly estimated marginal prices have the expected value. Moreover, in all other cases the estimated implicit prices increased importantly, as compared with the results that did not control for endogeneity. For example, the 3SLS estimates for PM_{10} increased 5 times. Thus, controlling for endogeneity does have an important impact on the magnitude of the estimated amenity prices. This result is consistent with what Bayer *et al.* (2009), with a different model, obtained when controlling for endogeneity in the amenities. According to the estimates controlled for endogeneity, the monthly value of a marginal improvement in air quality lies between CLP 1.529 and CLP 3.163 per μg/m^3 (grossly between US$ 3 and US$ 6) and in security between CLP 9.308 and CLP 12.737 per reduction in one per thousand crimes (grossly between US$ 17.50 and 24).

[15] See Cameron and Trivedi (2006) for further details about the foundations of the test.

[16] For the 2SLS estimation of both wage and rent equations, the Sargan–Hansen test was 0.319 (*p*-value .573) and 0.104 (*p*-value .747), respectively, and the Anderson tests were all significant at 1%.

[17] The estimation results for the wage and rent equations are respectively presented in tables A3 and A4 in Appendix (available at http://journals.cambridge.org/EDE).

5.3. Clustering effect

Finally, we evaluated how sensitive our results are to the effect of clustering, especially in the significance of the welfare measures. Clustering arises when an individual level of explanatory variable is regressed against some aggregate data at a city or community level. This is the case for both the PM_{10} and *CRIME*, which are available only at a communal level. As shown by Kloek (1981) and Moulton (1990), clustering reduces the variance of the coefficient associated with the aggregate explanatory variable. Therefore, it will also affect the variance of the marginal value measures. We deal with the clustering effects in two ways. First, we apply the clustering correction proposed by Kloek (1981) and Moulton (1990). Second, we predict PM_{10} values for communes where this information is lacking. With these predicted values we re-estimate the whole model for a larger sample (147 communes) and again apply the Moulton–Kloek correction. In table 5 we present the estimated OLS results for the amenity variables when the model is estimated with both the original (44 communes) and the predicted samples (147 communes), with and without the Moulton–Kloek correction. When we use the original sample the cluster-corrected results show no significance for the estimated amenities at traditional levels. However, when we use the extended (predicted) sample, the estimated amenities become significant at traditional levels, showing the impact of the small original sample on the estimated variance. We re-estimated the standard deviation only for the OLS model considering the Moulton–Kloek corrections. The first two columns of table 5 present our previous estimates where we used robust standard errors and the new estimates where the variances are corrected for clustering, respectively. In the cluster-corrected results, as expected, the standard errors of the variables aggregated at the communal level increased, making the *t*-test lower and reducing the statistical significance of these variables and the welfare measures. These results show that there is no evidence of significant estimated marginal values for the amenities. It is important to mention that most of the literature that has used this type of estimation has concluded that the welfare measures are statistically significant, without controlling for clustering effect. This issue should probably be present in all developing countries because of the lack of enough variability in the measures of airborne pollution.[18] If policy makers decide, for example by political priorities or budget constraints, to monitor only cities where pollution is very high and/or where population is also large, this turns into very limited variability of pollution indicators among cities and forces researchers to discard from the data all the cities where environmental information is not available.

The only solution to the clustering problem is to increase the total number of observations (clusters) included in the regression. This implies that either we have to wait until public and private institutions collect more and better information on pollution levels, or to find a way to

[18] We have enough information on crime rates by commune in Chile. It is the environmental information that creates the constraint by reducing the number of communes and the variance in the environmental indicator.

Table 5. *Marginal value (in pesos) of* PM_{10} *and crime in Chile with cluster effect, 2006: original and imputed values*

	Original data		*Imputed data (min)*		*Imputed data (average)*	
	PM₁₀					
	OLS	*OLS_cluster*	*OLS*	*OLS_cluster*	*OLS*	*OLS_cluster*
Wage equation parameter	2.65E − 03	2.65E − 03	3.77E − 03	3.77E − 03	1.78E − 03	1.78E − 03
Variance	2.56E − 07	1.80E − 06	1.37E − 07	1.15E − 06	1.67E − 07	1.12E − 06
Rent equation parameter	3.67E − 03	3.67E − 03	5.60E − 03	5.60E − 03	1.72E − 03	1.72E − 03
Variance	1.48E − 07	5.86E − 06	6.44E − 08	6.44E − 08	7.45E − 08	3.64E − 06
Covariance	0	0	0	0	0	0
Marginal value	**−491.86**	**−491.86**	**−681.30**	**−681.30**	**−362.87**	**−362.87**
Standard deviation	125.77	347.30	91.89	263.91	101.42	273.14
t-value	−3.91	−1.42	−7.41	−2.58	−3.58	−1.33
	Crime					
	OLS	OLS_cluster				
Wage equation parameter	−3.72E − 04	−3.72E − 04	1.54E − 03	1.54E − 03	2.82E − 03	2.82E − 03
Variance	1.20E − 06	7.83E − 06	3.57E − 07	5.64E − 06	3.38E − 07	5.78E − 06
Rent equation parameter	1.94E − 03	1.94E − 03	6.18E − 03	6.18E − 03	8.12E − 03	8.12E − 03
Variance	4.25E − 07	1.36E − 05	1.24E − 07	6.01E − 06	1.19E − 07	5.39E − 06
Covariance	0	0	0	0	0	0
Marginal value	**176.51**	**176.51**	**−107.57**	**−107.57**	**−339.06**	**−339.06**
Standard deviation	270.94	707.53	147.82	536.29	143.90	543.25
t-value	0.65	0.25	−0.73	−0.20	−2.36	−0.62

impute pollution levels for communes without information. We used two alternative ways to predict the missing pollution values. First, for communes without information we imputed the lowest level of pollution of all communes inside the province (upper administrative division). This assumption follows from the case when only critical pollution communes are monitored and the most probable missing pollution level is close to the minimum one within the province. This procedure will probably increase variability in the environmental variable. However, this procedure will be wrong if the implicit assumption of low pollution level for the locations with missing information is incorrect. Second, we used the mean value of the pollution level in the province as the imputed value. The second assumption follows from the idea that in several cases the pollution level of neighboring communes is not very different from the pollution level of communes for which the data exist. This might be the case in big urban areas where neighboring communes have similar environmental conditions (e.g., Santiago). The basic idea of these exercises is not to try to capture the actual pollution levels of the communes without information, but only to try to ascertain how the estimated variance for our amenity price estimates could change if we had access to a more complete sample.[19]

With the predicted PM_{10} for those communes with missing information, we re-estimated the wage and rent equations. We did this for both robust standard errors and cluster standard errors (we increased the clusters from 44 to 147). The estimated marginal values for PM_{10} and *CRIME* are presented for the OLS and OLS cluster-corrected results in columns 1 and 2 of table 5. Moreover, in the next four columns the results for the OLS and OLS cluster-corrected results for the imputed sample with both the methods discussed above are presented. We can see that the effect of clustering in all the cases is reduction in the level of significance of the coefficients. However, when we make the regressions with the expanded sample with imputed values, the coefficient associated with PM_{10} becomes statistically significant. This is the result of adding more communes into the regression analysis. As a result, the standard deviations of the estimated amenity values are reduced.

6. Conclusions

In this paper, we estimated the value of amenities for Chile using the wage hedonic approach. We used different econometric methods to estimate a system of equations composed of the hedonic wage equation and rent and selection equations. From this system, estimates of welfare measures for crime and pollution disamenities were calculated. Our results suggest that air quality and crime affect wage and rent differentials in Chile. Moreover, the monthly value for air pollution is significant and in our best estimate ranges between US$ 3 and US$ 6 per $\mu g/m^3$, depending on the estimation method. Using an aggregated measure of crime, the models give

[19] These exercises, meant to control for the clustering effect, constitute an informal way to evaluate consistency and preference-based sorting because of limited information about environmental conditions in the sample, in the estimated results.

the monthly value of one in thousand crimes as between US$ 17.50 and US$ 24.

Our results indicate that the selection bias is important for both the wage and the rent equations. The application of the procedure for controlling this bias in the rent equation is, in our case, specific to the way our data were collected. However, the control for selection bias in the wage equation is general to all cases. Thus, one lesson we can extract from these results is that this should be a standard procedure when estimating the wage hedonic model.

Furthermore, our results suggest that the simultaneity bias is not very important for the estimated amenity values. To the degree that this result can be extended to other studies, it should increase confidence in the estimates obtained by a great number of papers that do not control for simultaneity in the wage and housing price equations.

Moreover, attention should be paid to the endogeneity of the amenities when the welfare measures estimated in the model will be used in a cost-benefit analysis. Our results suggest that ignoring the endogeneity tends to cause an important downward bias in the welfare measures.

Finally, our results show that when one controls for the clustering effect of the amenity variables, the variance of the estimates increases greatly, turning nonsignificant the willingness to pay measures. We show also that this is a consequence of the low variability in our data set, and that increasing the number of observations, by way of predicting missing values, is a route to circumvent this problem. Especially in the case of LDCs with data shortages, this might be a useful methodology to develop to increase variability in the data sets.

References

Aguilar, R. and J. Dresdner (2006), 'The discrete charm of Chilean agriculture', in P. Aroca and G. Hewings (eds.), *Structure and Structural Change in the Chilean Economy*, New York: Palgrave MacMillan, pp.143–168.

Bayer, P., N. Keohane, and C. Timmins (2009), 'Migration and hedonic valuation: the case of air quality', *Journal of Environmental Economics and Management* **588**: 1–14.

Berger, M., G. Blomquist, and K. Sabirianova (2008), 'Compensating differentials in emerging labor and housing markets; estimates of quality of life in Russia', *Journal of Urban Economics* **63**: 25–55.

Blomquist, G., M. Berger, and J. Hoehn (1988), 'New estimates of quality of life in urban areas', *The American Economic Review* **78**: 89–107.

Buettner, T. and A. Ebertz (2009), 'Quality of life in the regions. Results for German counties', *The Annals of Regional Science* **43**: 89–112.

Cameron, A.C. and P. Trivedi (2006), *Microeconometrics: Methods and Applications*, New York: Cambridge University Press.

CEP (2006), *Encuesta Nacional de Opinión Pública Junio – Julio 2006*, Santiago, Chile: Centro de Estudios Públicos [Online], http://www.cepchile.cl, last accessed November, 18, 2010.

Clark, D. and J. Kahn (1988), 'The two-stage hedonic wage approach: a methodology for valuation of environmental amenities', *Journal of Environmental Economics and Management* **16**: 106–20.

Deitz, R. and J. Abel (2008), 'Have amenities become relatively more important than firm productivity advantages in metropolitan areas?', *Federal Reserve Bank of New York*, Staff Rep. No. **344** (September).

Deller, S. and T. Ottem (2001), 'Crime and the quality of life in Wisconsin counties', *Agricultural and Applied Economic Staff Papers*, University of Wisconsin, p. 31
Ffrench-Davis, R. (2003), *Entre el Neoliberalismo y el Crecimiento con Equidad*, 3rd ed., Santiago, Chile: LOM Ediciones.
Figueroa, E. and G. Lever (1992), 'Determinantes del precio de la vivienda en Santiago: una estimación hedónica', *Estudios de Economía* **19**(1): 69–83.
Figueroa, E., J. Rogat, and L. Firinguetti (1996), 'An estimation of the economic value of air quality improvement program in Santiago de Chile', *Estudios de Economía* **23**: 101–14.
Fuentes, J., A. Palma, and R. Montero (2005), 'Discriminación salarial por genero en Chile: una mirada global', *Estudios de Economía* **32**(2): 133–157.
Gabriel, S., J. Mattey, and W. Wascher (2003), 'Compensating differentials and evolution in the quality of life among U.S. States', *Regional Science and Urban Economics* **33**: 619–649.
Gyourko, J., M. Kahn, and J. Tracy (1999), 'Quality of life and environmental comparisons', in E. Mills and P. Cheshire (eds.), *The Handbook of Applied Urban Economics*, Amsterdam, the Netherlands: North Holland, pp. 1413–1454.
Gyourko, J. and J. Tracy (1991), 'The structure of local public finance and the quality of life', *The Journal of Political Economy* **99**: 774–806.
Hall, A.R., G.D. Rudebusch, and D.W. Wilcox (1996), 'Judging instrument relevance in instrumental variables estimation', *International Economic Review* **37**(2): 283–298.
Heckman, J. (1979), 'Sample selection bias as a specification error', *Econometrica* **47**(1): 153–61.
Hoehn, J., M. Berger, and G.C. Blomquist (1987), 'A hedonic model of interregional wages, rents and amenity values', *Journal of Regional Science* **27**(4): 605–620.
Killingsworth, M. (1983), *Labor Supply. Cambridge Surveys of Economic Literature*, Cambridge, UK: Cambridge University Press.
Kloek, T. (1981), 'OLS estimation in a model where a microvariable is explained by aggregates and contemporaneous disturbances are equicorrelated', *Econometrica* **49**(1): 205–207.
Madheswaran, S. (2007), 'Measuring the value of statistical life: estimating compensating wage differentials among workers in India', *Social Indicators Research* **84**: 83–96.
Marin, A. and G. Psacharopoulos (1982), 'The reward for risk in the labor market. Evidence from the United Kingdom and a reconciliation with other studies', *Journal of Political Economy* **90**(4): 827–53.
Mizala, A. and P. Romaguera (2001), 'La legislación laboral y el mercado del trabajo: 1975–2000', in R. Ffrench-Davis and B. Stallings (eds), *Reformas, Crecimiento y Políticas Sociales en Chile desde 1973*, Santiago, Chile: CEPAL, pp. 201–230.
Moulton, B. (1990), 'An illustration of a pitfall in estimating the effects of aggregate variables on micro units', *Review of Economics and Statistics* **72**: 334–338.
Pérez, C. and F. Arias (2007), 'Diferencias salariales asociadas a atributos ambientales en trece ciudades Colombianas: una estimación de salarios hedónicos', *Lecturas de Economía* **66**: 119–146.
Roback, J. (1982), 'Wages, rents, and the quality of life', *The Journal of Political Economy* **90**(6): 1257–1278.
Roback, J. (1988), 'Wages, rents, and amenities: differences among workers and regions', *Economic Inquiry* **26**: 23–41.
Rosen, S. (1974), 'Hedonic prices and implicit markets: product differentiation in pure competition', *Journal of Political Economy* **82**(1): 34–55.
Rosen, S. (1979), 'Wage-based indices of urban quality of life', in P. Mieszkowski and M. Straszheim (eds.), *Current Issues in Urban Economics*, Baltimore, MD: John Hopkins University Press, pp. 74–104.

Rosen, S. (1986), 'The theory of equalizing differences', in O. Ashenfelter and R. Layard (eds.), *Handbook of Labor Economics*, Vol. 1, Amsterdam, the Netherlands: North-Holland, pp. 641–692.

Timmins, C. and J. Murdock (2007), 'A revealed preference approach to the measurement of congestion in travel cost models', *Journal of Environmental and Economic Management* **53**: 230–249.

Wooldridge, J. (2002), *Econometric Analysis of Cross Section and Panel Data*, Cambridge, MA: The MIT Press.

Wunder, D. and H. Gutierrez (1992), 'Modelos tipo Box-Cox: proposición de una estrategia de estimación', *Revista de Análisis Económico* **7**(2): 79–102.