University of Groningen

Controlling omitted variables and measurement errors by means of constrained autoregression and structural equation modeling

Suparman, Yusep

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2015

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):
Suparman, Y. (2015). *Controlling omitted variables and measurement errors by means of constrained autoregression and structural equation modeling: Theory, simulations and application tomeasuring household preference forin-house piped water in Indonesia*. [Thesis fully internal (DIV), University of Groningen]. University of Groningen.

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license. More information can be found on the University of Groningen website: https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment.

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Download date: 21-06-2024

Chapter 5 The Willingness to Pay for In-House Piped Water in Urban and Rural Indonesia[1]

Abstract

This paper analyzes household preferences for in-house piped water in urban and rural Indonesia via a hedonic price model, specified as a constrained autoregression-structural equation model (ASEM). ASEM reduces bias due to time-varying omitted variables and measurement errors. In addition, it provides a convenient way of testing and correcting for endogeneity. On the basis of the Indonesia Family Life Survey data set, we find that on average urban and rural households have the same willingness to pay for in-house piped water i.e. 34.24% of their monthly house rent. For the 25% urban and rural households with lowest expenditure, this percentage is equivalent to 9.41% and 7.57% of their monthly expenditure, respectively. The findings support a need for further investment in in-house piped water in both areas, particularly for the households with the lowest expenditure levels.

Keywords: housing hedonic price model, structural equation model (SEM), constrained autoregression, piped water, urban and rural Indonesia.

1. Introduction

Initiated by the Declaration of the International Clean Drinking Water Decade 1981-1990, parties responsible for promoting health in developing countries have focused on improving domestic water supply. This initiative was followed by including safe

---

[1] This chapter has been published as Suparman Y, Folmer H, Oud JHL (in press) The willingness to pay for in-house piped water in urban and rural Indonesia. *Papers in Regional Science*.

drinking water and basic sanitation into the United Nations' Millennium Development Goals (MGD). Indonesia, as one of the targeted countries, has actively participated in improving its domestic water supply. Currently, 82% of the Indonesian population is served by improved water supply, i.e. piped water (into dwelling, yard or lot, public tap or standpipe), tube-well or borehole, protected dug well or spring, and rain water (UNICEF 2013). Nevertheless, the prevalence of water borne diseases is still high. For instance, annually about 30.3 thousand deaths are estimated to be associated with diarrhea incidence due to the consumption of contaminated water (WHO 2012). In addition to diarrhea, there are various other kinds of illnesses related to water quality such as schistosomiasis, trachoma, ascariasis, trichuriasis and hookworm (Prüss et al. 2002). In total, poor sanitation and hygiene contribute to 120 million disease incidents and 50 thousand premature deaths annually (Pinto 2013).

The high prevalence of waterborne diseases relates to the quality of water extracted from wells and springs. Particularly, current rapid environmental degradation has deteriorated the quality, as well as decreased the availability, of natural clean water. Under these conditions, the extension of piped water into dwellings or yards (denoted in-house piped water in the sequel) is the most effective mode to improve health in developing countries like Indonesia (Hutton and Haller 2004).

Currently, only 20% of the Indonesian population is served by in-house piped water service. The distribution between rural and urban Indonesia, however, is skewed. Of the rural population only 8% is served while the service covers 36% of the urban population (WHO-UNICEF 2013). A similar gap is found in other developing countries. For example, in other South Eastern Asian countries, on average 53% of the urban and 13% of the rural population is served by in-house piped water.

The gap between urban and rural water provision in Indonesia and other developing countries is related to the prevalent policy of domestic water supply. Particularly, the government of Indonesia (GoI) has focused in-house piped water investments mainly on urban areas. In rural areas, it has promoted alternatives, such as public wells, taps or hydrants. For instance, the rural water and sanitation project, ‘PAMSIMAS’[1], focuses on public taps and hydrants in 5000 villages only, which is a small fraction of the 79075 villages in rural Indonesia. Furthermore, of the 4,217 piped-water connections in 2012, only 218 were in-house connections. Most of the project was dedicated to the provision of public taps (3467 units) and public hydrants (532 units) (PAMSIMAS 2013). The prevalent policy is supported by some studies which generally report that urban households have stronger preferences for in-house piped water than rural households. See among others, Casey et al. (2006), McPhail (1993), Nauges et al. (2009), de Oca et al. (2003) and Whittington et al. (1991) for urban studies; North and Griffin (1993), Kihn et al. (2012) and Mbata (2006) for rural studies and Yusuf and Kondouri (2004) for rural and urban Indonesia.

Recent developments indicate that rural households in Indonesia increasingly choose vended water to meet their needs. For example, WHO-UNICEF (2013) and the Indonesia Family Life Survey (IFLS) (Straus et al. 2009) report an increase of purchases of vended water in rural areas from 0.8% in 1997 to 5.8% in 2007. However, compared to in-house piped water, vended water is more expensive while its quality and availability

---

[1] PAMSIMAS (2013) is an Indonesian Government project aimed at halving the number of households without access to drinking water and basic sanitation as determined by the UN MDGs minimum standards. The focus of the project is on villages in rural areas, particularly those with a high poverty rate. The project finances both physical (e.g. public tap and hydrant facilities) and non-physical (e.g. community empowerment and local institution development) investments.

are less.[2] Hence, the increase might indicate a change in the preference for in-house piped water in rural Indonesia with implications for in-house piped water investment policy.

The main purpose of this paper is to shed further light on the preferences for in-house piped water in rural and urban Indonesia and thus on continuation or modification of the PAMSIMAS project. For that purpose, we will estimate the willingness to pay (WTP) for in-house piped water by means of the hedonic price (HP) model on the basis of the IFLS panel data (Straus et al. 2009). We will apply the constrained-autoregression structural-equation approach (ASEM) (Suparman et al. 2013) to reduce bias due to omitted variables and measurement errors which, as in many HP studies, are also present in the IFLS data. We will furthermore use the specific features of ASEM to test and, if necessary to correct, endogeneity of the main variable of interest, i.e. the presence of in-house piped water.

2. The conceptual hedonic price model

The conceptual model presented below is based on a literature review and considerations which are partly inspired by data availability. The conceptual model describes the dependent variable, the explanatory variables and their expected impacts on the dependent variable.

The main variable of interest in this study is the WTP for in-house piped water. Since there are several alternative modes of domestic water provision, we consider the

---

[2] Piped water meets quality standards set and controlled by the government. Although in some areas, piped water may not be fully available for 24 hours, in principle it is distributed such that it meets daily needs. In contrast, for vended water that is usually distributed by chart or truck, no quality standard is applied. It is often directly extracted from wells and contaminated at delivery. Its distribution is dependent on vendors' schedules. In some cases, it is to be ordered.

WTP for in-house piped water relative to WTP for the alternative modes (see below for details).

The basic notion underlying HP analysis is that the value or price of a good, i.e. a house, is a function of the values or prices of its attributes. Because in-house piped water is an attribute of a house, the WTP for it can be extracted from the price differentials between houses with and without in-house piped water while controlling for other house characteristics. That is, the HP method uses systematic variation in house prices, attributable to the presence of in-house piped water, to impute the WTP for it (Haab and McConnell 2002). Accordingly, we use *house value* as the dependent variable in our HP model and estimate the impact of in-house piped water on it, controlling for other house attributes.

It is common in housing HP studies to use actual sale price, or rent, to represent *house value*. However, in Indonesia, as in many other developing countries, appropriate actual price or rent data is difficult to obtain. Often recorded sales in the real estate market do not reflect the actual prices due to, for example, high transaction costs. Similarly for rent which is often artificially low due to rent controls. When recorded sales or rents are absent or unreliable, one may resort to owner appraisal data (e.g. Anselin et al. 2008; Yusuf and Koundouri 2005). Here, we use owner monthly *rent appraisal* to measure *house value*.

As mentioned above, in Indonesia the main sources of domestic water provision, next to in-house piped water, are wells, rivers and lakes, and vendors. Generally, the value of a mode is determined by the following three characteristics: accessibility, availability, and quality. With respect to the first characteristic, in-house piped water has the highest accessibility, since it brings water into a house. Regarding availability, in-house piped water is also preferred to the other modes, since major interruptions in the provision are rare in Indonesia. However, minor interruptions often occur in urban areas during peak hours. Finally, the quality of piped water is substantially better than the quality of water obtained from the other sources. Particularly, it is safe for consumption, if it is boiled. The other sources vary in terms of the three criteria. However, they are always dominated by in-house piped water. So, in-house piped water definitely is a highly desirable attribute. Therefore, we hypothesize that it improves the value of a house. In the HP model we use a dummy variable with value 1 for *in-house piped water* and 0 for the other sources.

In some HP studies, *in-house piped water* is considered to be endogenous (e.g. Vásquez 2013). We test for endogeneity of *in-house piped water* by estimating the covariance between the error and *in-house piped water* as elements of the extended covariance matrix $\boldsymbol{\Psi}^*$ (see section 3). We use z-test to test the significance of this covariance. If the covariance is significant, the hypothesis of endogeneity is accepted, and ASEM will be estimated with the covariance between both terms explicitly included in the model in $\boldsymbol{\Psi}^*$.

In addition to *in-house piped water*, the data set contains two other sets of structural attributes. The first is *house size*, measured by *floor area*. This characteristic is common in most hedonic housing models (e.g. Bin and Polasky 2005; Brasington and Hite 2008; Espey et al. 2007; Kim et al. 2003; Nauges et al. 2009). The second set of attributes is *house condition* which is taken as the sum of five dummy variables representing the following five different house characteristics that are commonly included in HP housing models: *floor* and *wall material* (e.g. Koschinsky et al. 2012; McMillen 2004; Nauges et al. 2009), presence of *toilets* (e.g. Baltagi and Bresson 2011; Espey et al. 2007), *sewage* connection (Engle 1985) and *electricity* connection (e.g. Nauges et al. 2009). If *floor material* is ceramic or tiles, it takes the value 1, otherwise it is 0. If *wall* *material* is brick or cement, it takes the value 1 and 0 otherwise. The dummy variables for presence of *toilets*, *sewage* and *electricity* connections are defined in the usual way: 1 for presence and 0 for absence. The composite variable takes values between 0 and 5. We combine the dummies into a composite variable to reduce multicolinearity.[3] Note that we could have also taken *in-house piped water* as a *house condition* component and obtained its impact on *rent appraisal* by means of the coefficient of house condition in the structural model. However, this approach implies that the attributes are weighted equally. It is satisfactory, if the purpose is to control for the attributes simultaneously. However, if the focus is on the WTP for an individual attribute - in the present case for *in-house piped water* - it should be included as a separate variable. In that case, the WTP is its coefficient in the structural model (see section 3).

The data set contains only one neighborhood characteristic, viz. *median household monthly expenditure* on food and nonfood in the neighborhood where the house is located. It (or a proxy for it) is included in most housing HP studies to indicate the socio-economic status or *community wealth* of a neighborhood (see e.g. Brasington and Hite 2008; Koschinsky et al. 2012; McMillan 2004; Yusuf and Koundouri 2005). We hypothesize that *community wealth*, *house size,* and *house condition* have positive impacts on *house value,* since they reflect house quality dimensions. Note that other important neighborhood characteristics (like presence and quality of schools, safety, accessibility and shopping facilities) and the other house attributes are missing. ASEM will be applied to account for them.

---

[3] Inclusion of the five dummy variables separately leads to strong multicollinearity. Particularly, in the case of separate dummies, the highest condition numbers for multicollinearity for the urban and rural data sets are 69.65 and 29.11, respectively. The introduction of the composite variable reduces the highest numbers to 24.45 and 14.85. Note that strong multicollinearity affects estimation (Greene, 2008).

3. ASEM

In this section we present ASEM and the assumptions underlying the urban and rural HP model. Before going into detail, we point out that we use a linear model with *house value*, *community wealth* and *house size* measured in natural logarithms (log). The choice of a linear model (in the present case SEM) is supported by Cropper et. al. (1988), who suggest that a linear HP model consistently outperforms alternative functional forms, particularly the quadratic Box-Cox model, when some variables are not observed or replaced by proxies.

As mentioned above, the rationale for applying ASEM is twofold: reduction of bias due to (i) measurement errors, and (ii) omitted variables. In addition, we make use of the specific SEM structure to test and, if necessary, to correct endogeneity of *in-house piped water*. Regarding the presence of measurement errors in the explanatory variables, one of the core assumptions of the classical regression model is violated which results in an inconsistent estimator (Wooldridge 2002). Particularly, in the case of simple regression, the estimator has persistent bias toward zero (attenuation) (Greene 2008). In a SEM, the variance of an observed variable is decomposed into (i) a latent variable[4] variance and (ii) a measurement error variance.[5] When a latent variable that is purged of its measurement error is included in a model (see below), the impact of measurement error is eliminated or at least reduced. Identification of both kinds of variances can be done by means of using more than one observed variable as measurements or indicators

---

[4] A latent variable refers to a phenomenon that is supposed to exist but cannot be directly observed. A well-known example is socioeconomic status. A latent variable is measured by means of a set of observed variables (Folmer and Oud 2008). For instance, the latent variable socio-economic status is typically measured by observables like income, education, and profession. For further details, we refer to inter alia Suparman et al. (2013) and the references therein.

[5] From measurement equation (15b) and the zero conditional mean assumption, it follows that the variance of, say, observed variable y_1, equals $\mathrm{var}(y_1) = \mathrm{var}(\eta_1) + \mathrm{var}(\varepsilon_1)$. Hence, the variance of the observed variable falls apart into two components.

of the underlying latent variable. An alternative approach is to apply repeated measurements of a variable as indicators and to take the repeatedly measured variable as the latent variable. Since in the present paper the data analyzed is panel data, the latter option is applied.[6]

Structural equation models with latent variables (SEMs) have been developed to purge latent variables of their measurement errors. Specifically, a SEM is made up of two components: a structural model and two measurement models.[7] The structural model, which is a system of equations, reads:

$$\boldsymbol{\eta} = \boldsymbol{\alpha} + \boldsymbol{\Gamma}\boldsymbol{\xi} + \mathbf{B}\boldsymbol{\eta} + \boldsymbol{\zeta}.^{8} \tag{1}$$

In (1) the vector $\boldsymbol{\alpha}$ contains the intercepts. The matrix $\boldsymbol{\Gamma}$ specifies the relationships between the latent exogenous variables $\boldsymbol{\xi}$ and the latent endogenous variables $\boldsymbol{\eta}$ and the matrix $\mathbf{B}$ the relationships among the latent endogenous variables. The vector $\boldsymbol{\zeta}$ contains structural equation errors with covariance matrix $\boldsymbol{\Psi}$. These errors are assumed to be uncorrelated with $\boldsymbol{\xi}$. The expectation vector and covariance matrix of $\boldsymbol{\xi}$ are $\boldsymbol{\kappa}$ and $\boldsymbol{\Phi}$, respectively.

The measurement models read:[9]

$$\mathbf{x} = \boldsymbol{\tau}_x + \boldsymbol{\Lambda}_x\boldsymbol{\xi} + \boldsymbol{\delta}, \tag{2}$$

$$\mathbf{y} = \boldsymbol{\tau}_y + \boldsymbol{\Lambda}_y\boldsymbol{\eta} + \boldsymbol{\varepsilon}. \tag{3}$$

$\mathbf{x}$ and $\mathbf{y}$ are vectors of exogenous and endogenous observed variables (indicators) of the exogenous latent variables vector $\boldsymbol{\xi}$ and the endogenous latent variables vector $\boldsymbol{\eta}$,

---

[6] Note that the use of repeated measurements as indicators of a latent variable requires constraints on the measurement model parameters. See (10) and the subsection *The HP ASEM*, particularly (15a) and (15b).

[7] It is possible to combine the two measurement models and use one measurement model only (Oud and Delsing 2010).

[8] Note that multicolinearity is mitigated by subsuming highly correlated variables under one and the same latent variable in the structural model (Folmer 1981; Oud and Folmer 2008).

[9] Note that directly observed variables, lacking measurement errors, can be included in the structural model by way of identity relationships in the relevant measurement model.

respectively. The relationships between the observed variables and their respective latent variables are specified in the loading matrices $\mathbf{\Lambda}_x$ and $\mathbf{\Lambda}_y$, respectively. The vectors $\boldsymbol{\tau}_x$ and $\boldsymbol{\tau}_y$ are the intercepts of the measurement models. The vectors $\boldsymbol{\delta}$ and $\boldsymbol{\varepsilon}$ are the measurement errors with covariance matrices $\mathbf{\Theta}_\delta$ and $\mathbf{\Theta}_\varepsilon$, respectively. Often, but not necessarily, these covariance matrices are specified to be diagonal. The vectors $\boldsymbol{\delta}$ and $\boldsymbol{\varepsilon}$ are assumed uncorrelated with their corresponding latent variables.

To test, and correct for endogeneity, we reformulate the standard SEM (1)-(3) as (e.g. Oud and Jansen 2000):

$$\boldsymbol{\eta}^* = \mathbf{B}^*\boldsymbol{\eta}^* + \boldsymbol{\zeta}^* \tag{4}$$

and

$$\mathbf{y}^* = \mathbf{\Lambda}^*\boldsymbol{\eta}^* + \boldsymbol{\varepsilon}^* \tag{5}$$

with

$$\mathbf{y}^* = (\mathbf{y}' \quad \mathbf{x}' \quad \mathbf{1})', \qquad \boldsymbol{\eta}^* = (\boldsymbol{\eta}' \quad \boldsymbol{\xi}' \quad 1)', \qquad \boldsymbol{\varepsilon}^* = [\boldsymbol{\varepsilon} \quad \boldsymbol{\delta} \quad 1]', \qquad \boldsymbol{\zeta}^* = [\boldsymbol{\zeta}' \quad \mathbf{0} \quad 0]',$$

$$\mathbf{\Lambda}^* = \begin{bmatrix} \mathbf{\Lambda}_y & \mathbf{0} & \boldsymbol{\tau}_y \\ \mathbf{0} & \mathbf{\Lambda}_x & \boldsymbol{\tau}_x \\ \mathbf{0} & \mathbf{0} & 1 \end{bmatrix}, \Theta^* = \begin{bmatrix} \mathbf{\Theta}_\varepsilon & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Theta}_\delta & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & 0 \end{bmatrix}, \mathbf{B}^* = \begin{bmatrix} \mathbf{B} & \mathbf{\Gamma} & \boldsymbol{\alpha} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & 0 \end{bmatrix} \text{ and } \mathbf{\Psi}^* = \begin{bmatrix} \mathbf{\Psi} & \mathbf{\Omega} & \boldsymbol{\alpha} \\ \mathbf{\Omega}' & \mathbf{\Phi} & \boldsymbol{\kappa} \\ \boldsymbol{\alpha}' & \boldsymbol{\kappa}' & 1 \end{bmatrix}.$$

The sub-matrix $\mathbf{\Omega}$ of $\mathbf{\Psi}^*$ contains the covariance between the structural error and *in-house piped water*. Note that the elements of $\mathbf{\Omega}$ may not be identified for some models, (e.g. the standard multiple regression model). The HP-ASEM features specified below facilitate identification.

The parameters in the structural model and the measurement models are simultaneously estimated as a function of the elements of the covariance matrix and the mean vector of the observed variables. To see this, write $\boldsymbol{\eta}^*$ as a function of $\boldsymbol{\zeta}^*$ (equation (4)) and substitute it into measurement equation (5). This gives:

$$\mathbf{y}^* = \mathbf{\Lambda}^*\left(\mathbf{I} - \mathbf{B}^*\right)^{-1}\boldsymbol{\zeta}^* + \boldsymbol{\varepsilon}^* . \tag{6}$$

The theoretical covariance matrix of (6) is

$$\mathbf{\Sigma}^* = \mathbf{\Lambda}^*\left(\mathbf{I} - \mathbf{B}^*\right)^{-1}\mathbf{\Psi}^*\left(\mathbf{I} - \mathbf{B}^*\right)^{-1T}\mathbf{\Lambda}^{*T} + \mathbf{\Theta}^* = \begin{bmatrix} \mathbf{\Sigma}_{(\mathbf{yx})} & \boldsymbol{\mu}_{(\mathbf{yx})} \\ \boldsymbol{\mu}_{(\mathbf{yx})}^T & 1 \end{bmatrix} \tag{7}$$

where the last term of (7) follows from the fact that y* is made up of of $\mathbf{y}$ and $\mathbf{x}$.

Specifically:

$$\mathbf{\Sigma}_{(yx)} = \begin{bmatrix} \mathbf{\Lambda}_y\left(\mathbf{I} - \mathbf{B}\right)^{-1}\left(\mathbf{\Gamma\Phi\Gamma}^T + \mathbf{\Psi}\right)\left(\mathbf{I} - \mathbf{B}\right)^{-1T}\mathbf{\Lambda}_y^T + \mathbf{\Theta}_\varepsilon & \mathbf{\Lambda}_y\left(\mathbf{I} - \mathbf{B}\right)^{-1}\mathbf{\Gamma\Phi\Lambda}_x^T \\ \mathbf{\Lambda}_x\mathbf{\Phi\Gamma}^T\left(\mathbf{I} - \mathbf{B}\right)^{-1T}\mathbf{\Lambda}_y^T & \mathbf{\Lambda}_x\mathbf{\Phi\Lambda}_x + \mathbf{\Theta}_\delta \end{bmatrix}$$

and

$$\boldsymbol{\mu}_{(yx)} = \begin{bmatrix} \boldsymbol{\tau}_y + \mathbf{\Lambda}_y\left(\mathbf{I} - \mathbf{B}\right)^{-1}\left(\boldsymbol{\alpha} + \mathbf{\Gamma}\boldsymbol{\kappa}\right) \\ \boldsymbol{\tau}_x + \mathbf{\Lambda}_x\boldsymbol{\kappa} \end{bmatrix}.$$

Equation (7) shows that the theoretical covariance matrix is a function of all the model parameter matrices. Estimation of the theoretical covariance matrix and mean vector, and thus the SEM parameters, is done by minimizing the distance in some metric between (i) the theoretical matrix (7) and (ii) the observed covariance matrix and mean vector calculated from the data. In the present paper maximum likelihood will be applied. For further details see amongst others Jöreskog and Sörbom (1996), Bollen (1989) and Jöreskog et al. (2000).

We now turn to the second rationale of using ASEM: reduction of omitted variables bias. For that purpose, ASEM also exploits the features of panel data. Consider the following (HP) model:

$$p_t = \sum_i \pi_{it} q_{it} + \sum_j \pi_{jt} q_{jt} + o_t . \tag{8}$$

Furthermore, let the variables q_{jt} be omitted. In ASEM, the following constrained autoregressive panel model is estimated to reduce the omitted variables bias:

$$p_t = \rho_{1t} p_{t-1} - \rho_{1t} \sum_i \pi_{i(t-1)} q_{i(t-1)} + \sum_i \pi_{it} q_{it} + \upsilon_t \,. \tag{9}$$

As can be seen in (9), the omitted variables are approximated by the difference between (i) the lagged dependent variable and (ii) the sum of the lagged observed variables weighted by their lagged regression coefficients, where both (i) and (ii) are weighted by the autoregressive parameter of the dependent variable (Suparman et al. 2013). Note that while conventional correction procedures like differencing or fixed and random effects modeling can be applied to correct for time invariant omitted variables only, constrained autoregression allows correction of both time-varying and time-invariant omitted variables.

In addition to the constrained autoregression model (9), we define auxiliary autoregressive models for the three latent explanatory variables as follows:

$$q_{it} = \rho_{0it} + \rho_{it} q_{it-1} + v_t \,. \tag{10}$$

The auxiliary autoregressions enable identification of measurement error variance of a single indicator measurement model in that they add the covariances of such an indicator over different waves to the identification process (see Suparman et al. 2013) and the discussion of the identification of the simplex model in Jöreskog and Sörbom (1996, p. 230-234). Note that the autoregressions turn the house characteristics beyond the first wave into endogenous variables.

The HP ASEM

We will now present the HP ASEM in terms of the parameter matrices of equations (1)-(3). Before going into detail, we make the following two observations. First, to distinguish latent and observed variables, we denote the former with initial capitals (e.g. *House Value*) and the latter with lower initials (e.g. *house value*). Secondly, note that there are four waves of data collection (see section 4). If we used the standard HP model, there would be four HP models, one for each wave. In ASEM, however, there will be three models only because of the autoregression which requires the first wave variables to be exogenous. That is, we have HP models for the second, third, and fourth waves only. Thirdly, since it is directly observed (see below for an explanation), there is no need to specify autoregression for the dummy variable *in-house piped water*.

In SEM notation, we denote log(*rent appraisal*), log(*median households monthly expenditure*), log(*floor area*), and *house condition*[10] in wave-1 as x_1-x_4, respectively, and *in-house piped water* in wave-1 to wave-4 as x_5-x_8, their corresponding latent variables as ξ_1-ξ_8 and their measurement errors as δ_1-δ_8.[11] In addition, we denote log(*rent appraisal*), log(*median household monthly expenditure*), log(*floor area*), and *house condition score* in wave-2 as y_1-y_4, in wave-3 as y_5-y_8, and in wave-4 as y_9-y_{12}. Their corresponding latent variables are denoted as η_1-η_{12} and measurement errors as ε_1-ε_{12}. Accordingly, the exogenous and endogenous observed and latent vector are $\mathbf{x}' = [x_1 \quad \cdots \quad x_8]$, $\mathbf{y}' = [y_1 \quad \cdots \quad y_{12}]$, $\boldsymbol{\xi}' = [\xi_1 \quad \cdots \quad \xi_8]$, and $\boldsymbol{\eta}' = [\eta_1 \quad \cdots \quad \eta_{12}]$.

In SEM notation, the constrained autoregression HP model (9) supplemented with the auxiliary autoregressions of the three house characteristics reads:

---

[10] Note that we do not specify the five house condition variables as reflective indicators of the latent variable *House Condition*, since conceptually they are formative indicators of *House Condition*. Specifying them as reflective indicators would lead to bias due to causality misspecification (Diamantopoulos et al. 2008).

[11] Note that the relationship between the latent and the observed variable *in-house piped water* is identity.

$$\begin{bmatrix}\eta_1\\\eta_2\\\eta_3\\\eta_4\\\eta_5\\\eta_6\\\eta_7\\\eta_8\\\eta_9\\\eta_{10}\\\eta_{11}\\\eta_{12}\end{bmatrix}=\begin{bmatrix}\alpha_1\\\alpha_2\\\alpha_3\\\alpha_4\\\alpha_5\\\alpha_6\\\alpha_7\\\alpha_8\\\alpha_9\\\alpha_{10}\\\alpha_{11}\\\alpha_{12}\end{bmatrix}+\begin{bmatrix}\gamma_{11}&\gamma_{12}&\gamma_{13}&\gamma_{14}&\gamma_{15}&\gamma_{16}&0&0\\0&\gamma_{22}&0&0&0&0&0&0\\0&0&\gamma_{33}&0&0&0&0&0\\0&0&0&\gamma_{44}&0&0&0&0\\0&0&0&0&0&\gamma_{56}&\gamma_{57}&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&\gamma_{97}&\gamma_{98}\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\end{bmatrix}\begin{bmatrix}\xi_1\\\xi_2\\\xi_3\\\xi_4\\\xi_5\\\xi_6\\\xi_7\\\xi_8\end{bmatrix}+$$

$$\begin{bmatrix}0&\beta_{12}&\beta_{13}&\beta_{14}&0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0&0&0&0&0\\\beta_{51}&\beta_{52}&\beta_{53}&\beta_{54}&0&\beta_{56}&\beta_{57}&\beta_{58}&0&0&0&0\\0&\beta_{62}&0&0&0&0&0&0&0&0&0&0\\0&0&\beta_{73}&0&0&0&0&0&0&0&0&0\\0&0&0&\beta_{84}&0&0&0&0&0&0&0&0\\0&0&0&0&\beta_{95}&\beta_{96}&\beta_{97}&\beta_{98}&0&\beta_{910}&\beta_{9,11}&\beta_{9,12}\\0&0&0&0&0&\beta_{10,6}&0&0&0&0&0&0\\0&0&0&0&0&0&\beta_{11,7}&0&0&0&0&0\\0&0&0&0&0&0&0&\beta_{12,8}&0&0&0&0\end{bmatrix}\begin{bmatrix}\eta_1\\\eta_2\\\eta_3\\\eta_4\\\eta_5\\\eta_6\\\eta_7\\\eta_8\\\eta_9\\\eta_{10}\\\eta_{11}\\\eta_{12}\end{bmatrix}+\begin{bmatrix}\zeta_1\\\zeta_2\\\zeta_3\\\zeta_4\\\zeta_5\\\zeta_6\\\zeta_7\\\zeta_8\\\zeta_9\\\zeta_{10}\\\zeta_{11}\\\zeta_{12}\end{bmatrix}.(11)$$

The wave-2 HP model relating to the *House Value* (η_1) is given in row-1 of $\mathbf{\Gamma}$ and $\mathbf{B}$. Row-5 and row-9 of $\mathbf{\Gamma}$ and $\mathbf{B}$, present the wave-3 and wave-4 *House Value* models, respectively. Observe that the instantaneous coefficients are the WTP parameters for the corresponding associated house characteristics.

In addition to the rent appraisal autoregression models, there are auxiliary autoregressions. The autoregression coefficients of *Community Wealth*, *House Size*, and *House Condition* in wave-2, 3 and 4 are γ_{22}, γ_{33} and γ_{44}; β_{62}, β_{73} and β_{84}; and $\beta_{10,6}$,

$\beta_{11,7}$ and $\beta_{12,8}$, respectively. In addition to the structural parameters, we specify the means of latent exogenous variables in the mean vector

$$\boldsymbol{\kappa} = \begin{bmatrix} \kappa_1 & \cdots & \kappa_8 \end{bmatrix}', \tag{12}$$

the variances of the structural errors in the diagonal matrix

$$\boldsymbol{\Psi} = \operatorname{diag}\begin{bmatrix} \psi_{11} & \cdots & \psi_{12,12} \end{bmatrix}, \tag{13}$$

and the variances and covariances of the exogenous latent variables in the symmetric matrix

$$\boldsymbol{\Phi} = \begin{bmatrix} \phi_{11} & & & \\ \phi_{21} & \phi_{22} & & \\ \vdots & \vdots & \ddots & \\ \phi_{12,1} & \phi_{12,2} & \cdots & \phi_{12,12} \end{bmatrix}. \tag{14}$$

We now turn to the measurement models. As mentioned above, each variable is measured by a single observed variable. Hence, each observed variable serves as a reference variable for the underlying latent variable and therefore has unit factor loading and zero intercept. Note that in contrast to the other explanatory variables, we take *in-house piped water* as measured without error because it is unambiguously defined and measured twice: (i) by questioning the interviewee and (ii) via direct observation by the interviewer.[12] Accordingly, we have the following measurement models for the exogenous and endogenous observed variables:[13]

---

[12] The other variables are most likely to be subject to measurement error. Specifically, *wall* and *floor material* may be partly cement and partly bamboo or wood. *Floor area* is also likely to be measured with error because it partly depends on the interviewee's observation and recollection. Even the presence of a *toilet* is subject to measurement error because of the different forms of it may take, especially in lower class dwellings. For these variables cross-validation by the interviewer is not quite feasible. *Monthly expenditure* is a proxy for income and thus subject to measurement error. Similarly, *median monthly household expenditure* is a proxy for *Community Wealth.* Finally, *House Condition* is not only determined by the five attributes but also by other unobserved characteristics. It follows that all variables, except *in-house piped water*, are measured with error.

[13] Note that the parameters in the loading matrices are constrained to be equal to 1 for reasons of identification.

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \\ x_8 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \\ \xi_5 \\ \xi_6 \\ \xi_7 \\ \xi_8 \end{bmatrix} + \begin{bmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{15a}$$

and

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ y_7 \\ y_8 \\ y_9 \\ y_{10} \\ y_{11} \\ y_{12} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \\ \eta_5 \\ \eta_6 \\ \eta_7 \\ \eta_8 \\ \eta_9 \\ \eta_{10} \\ \eta_{11} \\ \eta_{12} \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \\ \varepsilon_6 \\ \varepsilon_7 \\ \varepsilon_8 \\ \varepsilon_9 \\ \varepsilon_{10} \\ \varepsilon_{11} \\ \varepsilon_{12} \end{bmatrix}. \tag{15b}$$

The measurement error covariance matrices (which we take to be diagonal) are:

$$\mathbf{\Theta}_\delta = \mathrm{diag}\left[\theta_{11}^\delta \quad \cdots \quad \theta_{44}^\delta \quad 0 \quad \cdots \quad 0\right] \tag{16}$$

and

$$\mathbf{\Theta}_\varepsilon = \mathrm{diag}\left[\theta_{11}^\varepsilon \quad \cdots \quad \theta_{1212}^\varepsilon\right]. \tag{17}$$

To test and account for endogeneity of *in-house piped water*, we specify the covariances between HP error terms and piped water dummies in the additional parameter matrix:

$$\mathbf{\Omega} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \omega_{61} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \omega_{75} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \omega_{89} & 0 & 0 & 0 \end{bmatrix}. \tag{18}$$

Constraints

Following Flores and Carson (1997), we assume that the valuation of the housing attributes changes proportionally to income (in the present paper to expenditure). This assumption translates into a linear equality constraint such that for each observation time point the WTP of the explanatory variables are proportional to the mean of *household monthly expenditure (constraint 1).* If we denote the wave-*i*/wave-*j* mean expenditure ratio as r_{ij}, the WTP at wave-*i* can be obtained as the product of r_{ij} and the WTP at wave-*j*. Hence, we only need to estimate a model for one wave (benchmark); the models for the other waves can be derived from it via *constraint 1*. The choice of the benchmark is arbitrary, because different benchmarks will lead to the same estimates. Note that *constraint 1* makes it possible to identify the coefficients of the lagged explanatory variables in the second wave models. From (7) it follows that these coefficients are the negative of the product of the autoregressive coefficient and the WTP estimates at the first wave. These latter WTPs can only be estimated as functions of the WTPs in another wave, since there is no model for the first wave.

As an illustration, consider the WTP for *Community Wealth* in wave-2, $\beta_{1,2}$. According to *constraint 1*, $\beta_{1,2}$ is the product of the proportion of mean expenditures in waves-2 and 4, and wave-4 *Community Wealth* WTP $\beta_{9,10}$ (the benchmark). Since the mean expenditure ratio for wave-2 and 4 is r_{24}, we obtain: $\beta_{1,2} = r_{24}\beta_{9,10}$. From Table 1, we obtain the mean *household monthly expenditure* ratios: $r_{14} = 0.7969$, $r_{24} = 1.0553$, and $r_{34} = 0.9070$ for urban areas and $r_{14} = 0.7866$, $r_{24} = 0.9701$ and $r_{34} = 0.9585$ for rural areas.

A second constraint derives from equation (9) which specifies the coefficient of a lagged house characteristic as the negative of the product of its lagged coefficient and the current *House Value* autoregression coefficient.

We also apply some constraints to the measurement model parameters. Since each observed variable in each wave is measured by the same set of items, we assume that the reliabilities of the measurements and the functional relationships between each indicator and its corresponding latent variable are equal over the waves. For a given measurement equation, this assumption implies that the same measurement model applies to all four waves. Hence, the variances of a given measurement error are equal over the waves (*constraint 2*). For (16) and (17) the assumption translates into the following equality constraints:

$$\theta_{kk}^{\delta} = \theta_{kk}^{\varepsilon} = \theta_{ll}^{\varepsilon} = \theta_{mm}^{\varepsilon}, \text{ for } (k,l,m) = (1,5,9),(2,6,10),(3,7,11),(4,8,12). \tag{18}$$

As an illustration of (16), consider *rent appraisal*. For (k,l,m) is $(1,5,9)$ we get $\theta_{11}^{\delta} = \theta_{11}^{\varepsilon} = \theta_{55}^{\varepsilon} = \theta_{99}^{\varepsilon}$. Note that the above constraints increase the efficiency of the estimator by increasing the degrees of freedom. Moreover, they have substantive implications.

We estimate the ASEM HP model by means of the maximum likelihood method[14] in the LISREL-8 software package (Jöreskog and Sörbom 1996). The LISREL package also provides several test statistics and modification indices. The latter can be used to

---

[14] Since we have large sample sizes, the maximum likelihood method provides robust estimates (Chou et al. 1991).

improve the model fit in that they give indications to free fixed or constrained parameters in the model.

4. Data and Empirical Results

4.1. Data

In Indonesia, no data is available which is tailored to the estimation of hedonic housing (HP) models. We therefore use the Indonesia Family Life Survey (IFLS)[15] data. The IFLS is not exclusively nor even purposively designed as a survey to estimate HP models. Hence, many important explanatory variables commonly included in HP models, are missing. However, as explained in the previous section, ASEM has been developed precisely for the purpose of controlling for missing variables to reduce bias of the estimator of the effects of the main variable of interest, the WTP for in-house piped water. Note that the IFLS allows comparing urban and rural livelihoods in that the sample is drawn to adequately represent urban and rural populations (Strauss et al. 2009). The classification of sample units as living in urban or rural areas is based on *Badan Pusat Statistik* (BPS), the Indonesian Statistics Agency. A village is classified as urban or rural according to population density, occupation structure of the population, and presence of public facilities obtained via a specific scoring method (BPS 2013).

The IFLS dataset is a panel data set for rural and urban areas made up of four waves. The first wave was administered in 1993. The same respondents were re-interviewed in 1997, 2000, and 2007 (Strauss et al. 2009). Since, the study covers rent appraisals by house owners, renters were excluded from the sample. Of the 7216 respondents in the first wave, 5423 met this specification with 2259 respondents from urban and 3164 from rural areas respectively.

---

[15] The IFLS is a longitudinal socio-economic and health survey of Indonesian individuals and households. It was conducted by the RAND Institute (Strauss et al. 2009).

Total dropout from the sample due to death or migration was 1492 (688 from urban and 904 from rural areas). In addition, some respondents refused to respond to some questions (non-response), particularly to the *rent appraisal* question (for waves 1-4: 404, 415, 207, and 187 in the urban regions, and 658, 598, 314 and 290 in the rural regions). Both dropouts and non-response result in a missing value problem. To evaluate whether the missing values lead to bias, we perform Little's (1988) MCAR (missing completely at random) test. The test chi-square values for the urban and rural sample are 953.0 and 1436.3, respectively with a 759 degree of freedom.[16] Since the chi-square test tends to reject the null hypothesis under a large sample size (e.g. Bentler and Bonet 1980; Hu et al. 1992; Jöreskog and Sörbom 1996), we applied the relative chi-square test (i.e. chi-square value/degree of freedom) proposed by Wheaton et al. (1977). In the case of this test, the null hypothesis is rejected, if the relative chi-square value is larger than two (Ullman 2007) or three (Kline 1998). Accordingly, we do not reject the null hypotheses and infer that missing values in urban and rural data are MCAR and do not bias the estimators.

Based on preliminary inspection of the data, we decided to exclude from the sample multilevel houses which have different characteristics. Particularly, multilevel houses in Indonesia are exclusively owned by households with a high socioeconomic status. Accordingly, a sample of 2085 rural and 1294 urban middle-to-lower income households resulted.

The IFLS is an ongoing longitudinal survey. New waves will be available in the future. The earliest new wave data set is expected to be available in 2014. The analysis in this paper is based on the waves that were available to date. Note that the analysis can be straightforwardly updated when new waves become available. Particularly, changes in the

---

[16] Since both urban and rural data have the same number of variables as well as missing value patterns, they have equal degrees of freedom.

WTP for in-house piped water can be tested. A Chow or chi-square test can be applied to test if a new wave relates to the same underlying population or not. In the latter case, the proportionality constraint (*constraint 1*) will not hold, a completely new model will be estimated and new WTP estimates will be obtained.

Table 5.1 Descriptive statistics

Variable	Statistic	Urban				Rural			
		1993	1997	2000	2007	1993	1997	2000	2007
rent appraisal	Mean	0.32	0.43	0.30	0.39	0.16	0.21	0.14	0.15
(Rp100,000/month)	s.d.	0.28	0.41	0.27	0.34	0.13	0.18	0.11	0.12
median household expenditure	Mean	1.97	2.52	2.12	2.50	1.19	1.48	1.45	1.54
(Rp100,000/month)	s.d.	0.86	1.22	0.90	2.06	0.46	0.51	0.38	0.57
floor area	Mean	0.53	0.68	0.70	0.75	0.50	0.64	0.67	0.71
(100m^2)	s.d.	0.31	0.31	0.33	0.37	0.28	0.27	0.30	0.32
house condition	Mean	3.23	3.56	3.72	4.03	1.82	2.36	2.67	3.12
(min=0, max=5)	s.d.	1.43	1.32	1.20	1.09	1.36	1.36	1.35	1.29
floor material (tile or ceramic)	Mean	0.36	0.42	0.43	0.64	0.13	0.22	0.24	0.41
(1=yes, 0=no)	s.d.	0.48	0.49	0.50	0.48	0.33	0.41	0.43	0.49
wall material(cement or brick)	Mean	0.70	0.77	0.80	0.84	0.38	0.48	0.55	0.65
(1=yes, 0=no)	s.d.	0.46	0.42	0.40	0.37	0.48	0.50	0.50	0.48
presence of toilet	Mean	0.60	0.71	0.75	0.84	0.35	0.48	0.53	0.66
(1=yes, 0=no)	s.d.	0.49	0.45	0.43	0.36	0.48	0.50	0.50	0.47
sewerage system connection	Mean	0.69	0.71	0.76	0.72	0.41	0.40	0.47	0.46
(1=yes, 0=no)	s.d.	0.46	0.46	0.43	0.45	0.49	0.49	0.50	0.50
electricity connection	Mean	0.88	0.96	0.97	0.99	0.55	0.78	0.87	0.95
(1=yes, 0=no)	s.d.	0.33	0.21	0.16	0.10	0.50	0.41	0.34	0.21
In-house piped water connection	Mean	0.13	0.18	0.23	0.35	0.01	0.05	0.08	0.18
(1=yes, 0=no)	s.d.	0.34	0.39	0.42	0.48	0.11	0.21	0.28	0.39
household expenditure	Mean	2.03	2.69	2.31	2.55	1.26	1.56	1.54	1.60
(Rp100,000/month)	s.d.	1.29	1.82	1.52	1.74	0.75	0.92	0.85	1.07

4.2. Empirical Results

The variables, their means and standard deviations are presented in Table 5.1. The table shows that all structural house characteristics increased in both urban and rural areas during the observation period. Nevertheless, *rent appraisal* as well as *household expenditure* decreased between 1997 and 2000 which is due to the economic crisis that hit Indonesia in 1997. The high inflation rate during the crisis decreased Indonesian households' purchasing power. After the crisis, *rent appraisal* and *household expenditure* increased, although in 2007 they had not reached the 1993 level yet. Accordingly, over the waves *rent appraisal* and *household expenditure* decreased while the house attributes increased. These seemingly confusing relationships are explained by the general relationship between WTP and income (see Flores and Carson (1997) and the discussion of *constraint 1*). Table 5.1 furthermore shows that urban areas have more favorable house characteristics and higher *rent appraisals* than rural areas. Moreover, in the former variability, as measured by the standard deviation, is larger in most cases.

Below, we first estimate separate urban and rural models, as specified in section 3. The estimates and fit measures are presented in Table 5.4 column 1a and 1b. Because of the large sample sizes, we evaluate model fit by means of the Root Mean Square Error of Approximation (RMSEA) rather than by the chi-square test statistic which is inappropriate in the case of large samples.[17] The RMSEA for both models is 0.06, which is well below the critical value of 0.08 for a reasonable model fit (Browne and Cudeck 1993). Another fit measure is the Goodness-of-Fit Index (GFI). It is 0.95 for both models, which is at the lower bound of 0.95 for a good model fit (Hoelter 1983). Other indications

---

[17] Note, however, that in the case of large sample size, the chi-square statistic can be used to test nested hypotheses (Bollen 1989), as shown below.

of good fit are the parameter estimates which are within acceptable ranges and the absence of Heywood cases and implausibly large standard errors.[18]

The measurement model estimates (Table 5.2) show that all the observed variables are highly reliable indicators. Except for *floor area*, the reliability coefficients[19] are higher than 0.8. The reliability of *floor area* for wave-3 and 4 are 0.7 and 0.71, respectively. These reliability levels are also acceptable.[20] Hence, we proceed to the structural model.

Table 5.2 Measurement error variance estimates and reliability coefficients

Variable	Wave	Urban		Rural	
		Error variance	Reliability	Error variance	Reliability
rent appraisal	1	0.16	0.95	0.12	0.97
	2		0.94		0.97
	3		0.94		0.98
	4		0.94		0.98
median household monthly expenditure	1	0.03	0.93	0.03	0.83
	2		0.96		0.88
	3		0.95		0.85
	4		0.96		0.90
floor area	1	0.14	0.95	0.03	0.97
	2		0.89		0.94
	3		0.70		0.94
	4		0.71		0.94
house condition score	1	0.30	0.97	0.46	0.91
	2		0.98		0.94
	3		0.98		0.95
	4		0.98		0.96

---

[18] Since SEM estimation is based on numerical procedures, negative variance estimates (Heywood cases) may result (Haughton et al. 1997). Furthermore, in the case of substantial deviation from normality, maximum likelihood may produce implausibly large standard error estimates (Bentler 1989; Chou et al. 1991).

[19] The reliability is the proportion of an observed variable's variance explained by its latent variable. It equals 1 - (measurement error variance /observed variable variance).

[20] Common practice in the social sciences is that 0.7 is the lower bound for good reliability.

Before going into the detail, note that we only present and discuss the WTP estimates for wave-4. Under the proportionality assumption, the WTP estimates and standard errors for the other waves can be obtained from the wave-4 estimates. The complete set of estimates for all waves is also available at http://blogs.unpad.ac.id/yusepsuparman/research/.

As a first step, we test endogeneity of *in-house piped water*. The estimates of the covariances between the error terms of the various structural models and *in-house piped water* are presented in Table 5.3. We find that all of them are significantly different from zero, except for the wave-2 rural model. Accordingly, we conclude that *in-house piped water* is endogenous. Hence, we continue our analysis by retaining the covariances between the error terms and *in-house piped water* in the models.

Table 5.3 The covariances between HP error term and piped water (endogeneity test)

Statistics	Urban			Rural		
	wave-2	wave-3	wave-4	wave-2	wave-3	wave-4
Estimate	-0.0234	-0.0242	-0.0400	-0.0045	-0.0120	-0.0260
Standard error	0.0062	0.0071	0.0133	0.0034	0.0049	0.0104
p-value	0.0000	0.0000	0.0000	0.1857	0.0143	0.0124

Next we check the proportionality assumption (*constraint 1*). For that purpose, we compare the estimated model under *constraint 1* with the model under the more restrictive assumption of constant WTP for each house characteristic across all four waves. The comparison is for the urban and rural models separately. The urban and rural models under the constant WTP constraint are presented in Table 5.4, column 2a and 2b. The chi-square values under this constraint are 748.64 for the urban model and 998.34 for the rural model. For the models under the proportionality constraint (*constraint 1*) in Table 5.4, columns 1a and 1b, the corresponding chi-square values are 735.74 and 990.71, respectively. Since the models under the constant WTP constraint and the proportionality constraint have the same degree of freedom, we cannot perform a formal chi-square test.[21] However, the chi-square values under the constant WTP constraint are substantially larger than under the proportionality constraint. Hence, we reject the constant WTP assumption. Accordingly, we continue the analysis under the assumption of proportionality.

Table 5.4 The structural ASEM and FEM estimates with goodness of fit Indices

Variables	ASEM (a)		ASEM (b)		FEM (a)		ASEM (a,c)	
	Urban (1a)	Rural (1b)	Urban (2a)	Rural (2b)	Urban (3a)	Rural (3b)	Urban (4a)	Rural (4b)
House Value (lagged)	0.54	0.26	0.0.45	0.24	-	-	0.54	0.26
	(0.05)	(0.04)	(0.04)	(0.04)	-	-	(0.05)	(0.04)
Community Wealth	0.77	0.43	0.86	0.45	0.50	0.43	0.79	0.42
	(0.07)	(0.05)	(0.06)	(0.05)	(0.03)	(0.04)	(0.06)	(0.05)
House Size	0.22	0.10	0.19	0.09	0.10	0.13	0.22	0.10
	(0.06)	(0.03)	(0.06)	(0.03)	(0.02)	(0.02)	(0.06)	(0.03)
House Condition	0.26	0.21	0.29	0.22	0.19	0.13	0.26	0.21
	(0.02)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
in-house piped water	0.40	0.30	0.41	0.32	0.55	0.24	0.34	
	(0.08)	(0.07)	(0.07)	(0.07)	(0.04)	(0.07)	(0.06)	
Unit constant	-1.46	-2.21	-1.78	-2.26	-2.65	-2.77	-1.43	-2.21
	(0.16)	(0.15)	(0.12)	(0.12)	(0.04)	(0.04)	(0.16)	(0.12)
RMSEA	0.06	0.06	0.05	0.05	0.09	0.07	0.06	
GFI	0.95	0.95	0.96	0.96	0.95	0.97	0.96	
Chi-square	735.74	990.71	748.64	998.34	1210.22	669.67	1727.43	
df	143	139	143	139	64	64	283	

Notes: Standard errors within brackets. All estimates are significant at 0.01 level.
(a) Estimated under the proportional WTP constraint (1).
(b) Estimated under the constant WTP constraint.
(c) Estimated under the constraint of equal WTP for *in-house piped water* in urban and rural areas.

---

[21] This is the reason why we speak of a check rather than a test.

To get insight into its performance, we compare ASEM to the fixed effects panel model[22] which is frequently used to correct for omitted variables (Greene 2008). The urban and rural ASEM models are presented in Table 5.4, columns 1a and 1b, and the urban and rural fixed effects models in columns 3a and 3b. The table shows that the fixed effect estimates are far below the corresponding ASEM estimates (attenuation bias). Particularly, the urban fixed effect estimate of *Floor area* is smaller than the corresponding rural estimate. This outcome is highly unreasonable and inconsistent with inter alia Yusuf and Koundouri (2004).[23] Accordingly, we continue our analysis on the basis of the ASEM estimates.

As mentioned in section 1, this paper focuses on the question whether the WTP for piped water in urban and rural areas are equal or not. As the first step, we test the more restrictive hypothesis that the same model holds for both the urban and rural areas, i.e. the WTPs for all house characteristics are equal in both areas. With wave-4 as the benchmark, this hypothesis reads:

$$H_0 : \begin{bmatrix}\beta_{9,10} & \beta_{9,11} & \beta_{9,12} & \gamma_{9,8}\end{bmatrix}^{urban} = \begin{bmatrix}\beta_{9,10} & \beta_{9,11} & \beta_{9,12} & \gamma_{9,8}\end{bmatrix}^{rural} \qquad (19)$$ [24]

To test hypothesis (19) we apply the chi-square likelihood ratio test of a nested model in the SEM multi-group framework (Jöreskog and Sörbom 1996). The test statistic is:

$$\Delta\chi^2 = \chi^2_{H_0} - \chi^2_{H_1} \qquad (20)$$ [25]

---

[22] We estimate the fixed effect model controlling endogeneity of *in-house piped water* by specifying the covariance between the error term and *in-house piped water*, as described in section 3 with respect to ASEM.

[23] For further comparisons of ASEM and alternative approaches to correct for measurement error and omitted variables, including the fixed effect panel approach, we refer to Suparman et al. (2013).

[24] Note that the alternative hypothesis implies that some, or all, of the rural and urban coefficients are different,

[25] Note that under the null we basically have a pooled model and under the alternative hypothesis two separate models.

with $\chi^2_{H_0}$ and $\chi^2_{H_1}$ the overall model fit chi-square values under the null and under the alternative hypothesis, respectively, and df_{H_0} and df_{H_1} their corresponding degrees of freedom. We reject the null at level α if $P\left(\chi^2_{df_{H_0}-df_{H_1}} \geq \Delta\chi^2\right) < \alpha$ with $\chi^2_{df_{H_0}-df_{H_1}}$ the critical chi-square value with $df_{H_0} - df_{H_1}$ degrees of freedom.

The first panel of Table 5.5 shows that the null hypothesis of equal WTP for all house characteristics in both areas (19) is rejected: $\Delta\chi^2$ = 176.19 with *df* =4 which gives a zero *p-value*. Accordingly, we can infer that the WTP for at least one house characteristic in the urban and rural regions differs.

We now turn to the equality of the WTP for *in-house piped water*. The hypothesis reads:

$$H_0 : \gamma_{9,8}^{rural} = \gamma_{9,8}^{urban} \tag{21}$$

vs.

The test procedure for hypothesis (21) is as in the case of hypothesis (19).[26] The chi-square values and their degree of freedom for (21) are presented in the second panel of Table 5. We find that $\Delta\chi^2$ = 0.99 with *df* = 1 which gives a *p-value* = 0.32. Hence, we accept hypothesis (21) and infer that the WTP for *in-house piped water* in urban and rural areas are equal.

Equal WTP for *in-house piped water* across urban and rural areas is not in line with previous literatures which show that in Indonesia the WTP in the former is larger than in the latter (Yusuf and Koundouri 2004). A likely explanation for our finding is that health, which is closely related to the availability of clean and safe drinking water (see section 1), is valued approximately equally in both areas. In this regard, it is important to

---

[26] Note that although the alternative to hypothesis (21) is different from the alternative to hypothesis (19), their associated SEM models are the same and thus their associated overall chi-squares are equal.

note that the availability of alternative sources of safe drinking water, such as wells, springs, lakes and rivers, which used to be far more abundant in rural than in urban areas, have been rapidly declining in the former because of increasing pollution and excessive extraction. Hence, the availability of safe alternatives in both areas tends to converge. In addition, availability and accessibility of in-house piped water tend to converge in both kinds of regions (Yusuf and Koundouri 2004).

We also tested the equalities of the WTP for the other house characteristics. Table 5 shows that all the null hypotheses were rejected. Accordingly, we conclude that the WTPs for *Community Wealth*, *House Size* and for *House Condition* differ between urban and rural areas. These different WTPs are likely to follow from differences in income, education, life styles, and the valuation of private and public goods (Jones et al. 2009).

The urban and rural models with equal WTP for *in-house piped water* but different WTPs for the other house characteristics are presented in Table 4, column 4a and 4b. Note that the Final Models only marginally differ from the ASEM under the proportionality constraint for which the equality constraint

$$\gamma_{9,8}^{rural} = \gamma_{9,8}^{urban} \tag{22}$$

(i.e. equal WTP for *in-house piped water* in both areas) does not apply.

Table 5.5 Equality test of WTP for housing characteristics, wave-4

Null Hypothesis	Statistic	Value
(1) Equal WTPs for all house attributes	$\chi^2_{H_0}$	1902.65
	df_{H_0}	286
	$\Delta\chi^2$	176.19
	$df_{H_0} - df_{H_1}$	4
	p-value	0.00
(2) Equal WTP for *in-house piped water*	$\chi^2_{H_0}$	1727.44
	df_{H_0}	283
	$\Delta\chi^2$	0.99
	$df_{H_0} - df_{H_1}$	1
	p-value	0.32
(3) Equal WTP for *Community Wealth*	$\chi^2_{H_0}$	1751.39
	df_{H_0}	283
	$\Delta\chi^2$	24.93
	$df_{H_0} - df_{H_1}$	1
	p-value	0.00
(4) Equal WTP for *House Size*	$\chi^2_{H_0}$	1758.80
	df_{H_0}	283
	$\Delta\chi^2$	32.35
	$df_{H_0} - df_{H_1}$	1
	p-value	0.00
(5) Equal WTP for *House Condition*	$\chi^2_{H_0}$	1745.74
	df_{H_0}	283
	$\Delta\chi^2$	19.29
	$df_{H_0} - df_{H_1}$	1
	p-value	0.00

Note: The chi-square value under the alternative hypotheses is the sum of the chi-square values of the urban and rural models in Table 4, column 1a and 1b. It equals to 1726.45 with $df = 282$.

The positive signs of the WTP of *in-house piped water* (Table 4, column 4a and 4b) relative to the alternatives sources of wells, rivers and lakes and vendors indicate that in-house piped water is the most desirable type of domestic water provision in urban and rural Indonesia. The estimates indicate that on average, the presence of *in-house piped water* increases the *House Value* by 34.24% which is equivalent to 5.18% and 3.22% of total *monthly household expenditure* (on food and non-food items) in urban and rural areas, respectively. For the 25% of households with the lowest total expenditure levels, these percentages are 9.96% and 5.23%, respectively.[27] These percentages are higher than household expenditure on vended water in Jakarta, Indonesia, in the 1990s. Crane (1994) reports that in the absence of piped water or a public water tap, poor households spend 7.5% of their expenditure on water from vendors. Lovie and Whittington (1993) find that poor households in Jakarta might spend as much as 7% of their income on vended water in the dry season. As mentioned above, a possible explanation for the increase in the WTP for piped water is that the quality of water from alternative sources has substantially deteriorated during the past two decades and more information has become available on the health risks of consuming polluted water. Moreover, compared to vended water in-house piped water has the additional advantage of easier access to water, both for consumption and for other uses such as washing.

5. Conclusion

In this paper we have estimated the willingness to pay (WTP) for in-house piped water in Indonesia by means of a housing hedonic pricing (HP) model. We applied the constrained autoregressive - structural equation approach (ASEM) to mitigate bias due to

---

[27] These percentages are obtained from the WTP for in-house piped water relative to rent appraisal (0.1202) multiplied by the average of 2007 rent appraisal (Table 4), divided by the average of monthly food consumption, multiplied by 100. The averages of monthly food consumption in urban and rural areas are obtained from the 2007 IFLS data. They are 0.4653 and 0.3441 (times Rp100.000), respectively.

time varying omitted variables and to measurement error in the explanatory variables. These types of problems are quite common in HP and other survey based studies.

The major finding is that households in urban and rural areas have the same relative WTP for in-house piped water. Particularly, *in-house piped water* increases the average of house rent appraisal by 34.24% which is equivalent to 5.18% and 3.22% of *monthly household expenditure* (on food and non-food items) in urban and rural areas, respectively. For the 25% households with the lowest expenditure levels, these latter estimates are equivalent to 9.41% and 7.57%, respectively. These results are higher than the outcomes from earlier studies on actual expenditure on vended water by poor households in Indonesia. Possible explanations for the increase are: (i) piped water has better quality, accessibility and availability than vended water; (ii) during the past 20 years, awareness of the importance of piped water has increased and the availability and quality of natural water sources has deteriorated while (iii) people are better informed about the risks of consuming polluted water.

Our results imply that urban and rural households have the same strong relative preference for in-house piped water. The findings are relevant for water provision in Indonesia. Particularly the Government of Indonesia should reconsider its policy to give priority to urban areas. In both rural and urban areas, there is a strong preference for in-house piped water and both areas should be targeted.

Our findings also show that the average WTPs of poor households are well above the World Bank “affordability” benchmark of 5% of total income. Moreover, the relative WTPs of poor households are higher than those of wealthier households. Accordingly, they should be prioritized. However, if the government decided to apply a full recovery scheme, high water tariffs would result which would deter poor households from connecting to the system (Jalan et al. 2009). Instead, they would divert to less safe but affordable modes. Therefore, the government should consider support systems such as subsidies where they are in place and introduce where they are not (see also van den Berg and Nauges 2012). Alternatively, the government might consider a block tariff system with the first block free of charge for the poor (Majumdar et al. 2009).

Increasing the accessibility of piped water in urban and rural areas will generate substantial welfare improvements by reducing morbidity and mortality due to water-borne diseases, particularly for the poor households. In addition, piped water connections may increase households' real income (Nauges et al. 2009) and hence contribute to the reduction of the disparity between the rich and the poor. However, extension of the piped water system will be costly. Therefore, further studies on financing mechanisms of the extensions are required.

References

Anselin L, Gracia NL, Deichmann U, Lall S (2008) Valuing access to water: A spatial hedonic approach applied to Indian cities. Working Paper No. WPS4533, the World Bank

Baltagi BH, Bresson G (2011) Maximum likelihood estimation and Lagrange multiplier test for panel seemingly unrelated regressions with spatial lag and spatial errors: an application to hedonic housing prices in Paris. *Journal of Urban Economic* 69: 24-42

Bentler PM (1989) *EQS structural equation program manual*. BMDP Statistical Software, Lost Angles

Bentler PM, Bonnet DG (1980) Significance test and goodness of fit in the analysis of covariance structures. *Psychological Bulletin* 88: 588-606

Bin O,Polasky S(2005) Evidence on the amenity value of wetlands in a rural setting. *Journal of Agricultural and Applied Economics* 37: 589-602

Bollen KA (1989) *Structural equation with latent variables*. Wiley, New York

BPS (2013) Istilah Statistik - Statistical Terms, Badan Pusat Statistik http://www.bps.go.id/menutab.php?tab=4&tabel=1&kat=1&id_subyek=30&ist=1&var=D (08/14/2013)

Brasington DM, Hite D (2008) A mixed index approach to identifying hedonic price model. *Regional Science and Urban Economics* 38: 271-284

Browne MW, Cudeck R (1993) Alternative ways of assessing model fit. In: Bollen KA, Long JS (eds) *Testing structural equation models*. Sage, Newbury Park

Casey JF, Kahn JR, Rivas A (2006) Willingness to pay for improved water service in Manaus, Amazonas, Brazil.*Ecological Economics* 58: 365-372

Chou C, Bentler PM, Satorra A (1991) Scaled test statistics and robust standarderrors for non-normal data in covariance structure analysis: A Monte Carlo study.*British Journal of Mathematical and Statistical Psychology* 44: 347-357

Crane R (1994) Water market, market reform and the urban poor: Results from Jakarta, Indonesia. *World Development* 22: 71-83

Cropper ML, Deck L,McConnell KE (1988) On the Choice of Functional Forms for Hedonic Price Functions, *Review of Economics and Statistics*70: 668-675

De Oca GSM, Bateman IJ, Tinch R, Moffatt PG (2003) *Assessign the willingness to pay for maintained and improved water supply in Mexico City*. Working paper No ECM 03-11. CSERGE.

Diamantopoulos A, Riefler P, Roth KP (2008) Advancing formative measurement models. *Journal of Business Research* 61: 1203-1318

Engle RF, Lilien DM, Watson M (1985) A dynamic model of housing price determination. *Journal of Econometrics* 28: 307-326

Espey M, Fakhruddin F, Gering LR, Lin H (2007) Living on the edge: residential property values in the urban-rural interface. *Journal of agricultural and applied economics* 39: 689-699

Flores NE, Carson RT (1997) The relationship between the income elasticities of demand and willingness to pay. *Journal of Environmental Economics and Management* 33: 287-295

Folmer H (1981) Measurement of the effect of regional policy instruments by means of linear structural equation models and panel data. *Environment and Planning A* 13: 1435-1448

Folmer H, Oud JHL (2008) How to get rid of W: A latent variable approach to modeling spatially lagged variables. *Environmentand Planning A* 40: 2526-2538

Greene WH (2008) *Econometric Analysis*. Prentice Hall, New Jersey

Haab TC, McConnell KE (2002) *Valuing environmental and natural resources: The economitrics of non-market valuation*.Edward Elgar,Cheltenham

Haughton DMA, Oud JHL, Jansen RARG (1997) Information and other criteria in structural equation model selection.*Communication in Statistics – Simulation and Computation* 26: 1477-1516

Hoelter JW (1983) The analysis of covariance structure: Goodness-of-Fit Indices. *Sociological Methods and Research* 11: 324-44

Hu L, Bentler PM, Kano Y (1992) Can test statistics in covariance structure analysis be trusted?.*Psichological Bulletin* 112: 351-362

Hutton, G. and Haller, L. (2004) *Evaluation of the Cost Benefit of Water and Sanitation Improvements at the Global Level*.Working Paper No WHO/SDE/WSH/04.04, WHO

Jalan J, Somanathan E, Chaudhuri S (2009) Awareness and the demand for environmental quliaty: survey evidence on drinking water in urban India. *Environment and Development Economics* 14: 665-692

Jones N, Malesios C, Botetzagias I (2009) The influence of social capital on willingness to pay for the enviroment among European citizen. *European Societies*, 11: 511-530

Jöreskog K, Sörbom D (1996) *LISREL 8: User's reference guide*. Scientific Software International, Chicago

Jöreskog K, Sörbom D, du Toit S, du Toit M (2000) *LISREL 8: New statistical features*. Scientific Software International, Chicago

Kihn (2012) Valuing rural piped water services in the Mekong Delta, Vietnam: Case study at Phuoc Vinh Dong Commune.In Subramanian V (ed.) *Coastal Environment: Focus on Asian Regions*. Springer, Dordrecht

Kim CW, Phipps TT, Anselin L (2003) Measuring the Benefit of Air Quality Improvement: A Spatial Hedonic Approach.*Journal of Environmental Economics and Management* 45: 24-39

Kline RB (1998) *Principles and practice of structural equation modeling*. The Guilford Press, New York

Koschinsky J, Lozano-Garcia N, Piras G (2012) The welfare benefit of a home's location: an empirical comparison of spatial and non-spatial model estimates. *Journal of Geographical System* 14: 319-356.

Little RJA (1988) A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association* 83: 1198-1202.

Lovie L, Whittington D (1993) Rent-extracting behavior by multiple agents in the provision of municipal water supply: A study of Jakarta, Indonesia.*Water Resources Research* 29: 1965-1974

Majumdar C, Gupta G (2009) Willingness to pay and municipal water pricing in transisition: a case study. *Journal of Integrative Environmental Science* 6: 247-260

Mbata JN (2006) Estimating household willingness to pay for water services in a rural economy: the case of Kanye in southern Botswana.*Development Southern Africa* 23: 29-43

McMillan DP (2004) Airport expansions and property values: The case of Chicago O'Hare Airport. *Journal of Urban Economic* 55: 627-640.

McPhail AA (1993) The "Five Percent Rule" for improved water service: can household afford more? *World Development* 21: 963-973

Nauges, C, Strand, J, Walker I (2009) The value of water connection in Central American cities: a revealed preference study.*Environment and Development Economics* 14: 349-370

North JH, Griffin CC (1993) Water source as a housing characteristic: hedonic property valuation and willingness to pay for water. *Water Resources Research* 29: 1923-1929

Oud JHL, Delsing JMH (2010) Continuous time modeling of panel data by means SEM. Invan Montfort K, Oud JHL, Satorra A (Eds) *Longitudinal research with latent variables*. Springer, Heidelberg

Oud JHL, Folmer H (2008) A structural equation approach to models with spatially dependence.*Geographical Analysis* 40:152-66

PAMSIMAS (2013) Tentang PAMSIMAS – About PAMSIMAS, PAMSIMAS http://new.pamsimas.org/index.php?option=com_k2&view=item&layout=item&id=15&Itemid=129 (08/14/2013)

Pinto R (2013) Results, Impacts, and Learning from Improving Sanitation at Scale in East Java, Indonesia. Water and Sanitation Program: Field Note, The World Bank

Prüss A, Kay D, Fewtrell L, Bartram B (2002) Estimating the Burden of Disease from Water, Sanitation, and Hygiene at a Global Level.*Environmental Health Perspective*110: 537-542

Strauss J, Witoelar F, Sikoki B, WattieAM (2009)Thefourth wave of the Indonesia Family LifeSurvey (IFLS4): Overview and field report. Working Paper No. WR-675/1-NIA/NICHD, RAND Labor and Population.

Suparman Y, Folmer H, Oud JHL(2014) Hedonic Price Models with Omitted Variables and Measurement Errors: A Constrained Autoregression - Structural Equation Modeling Approach with Application to Urban Indonesia. *Journal of Geographical System* 16: 49-70. DOI: 10.1007/s10109-013-0186-3

Ullman JB (2007) Structural equation Modeling. In Tabachnick BG and Fidell LS (eds) *Using multivariate statistics* 5th ed. Allyn Bacon, New York

UNICEF (2013) Access to water and sanitation: a few definitions. UNICEF http://www.unicef.org/wcaro/overview_2570.html (04/21/2013)

van den Berg C, Nauges C (2012) The willingness to pay for access to piped water: a hedonic analysis of house prices in Southwest Sri Langka. *Letter in Spatial and Resource Sciences* 5: 151-166

Vásquez WF (2013) A Hedonic valuation of residential water services. *Applied Economic Perspective and Policy*. DOI: 10.1093/aepp/ppt022

Wheaton B, Muthen B, Alwin DF, Summers G (1977) Assessing reliability and stability in panel models. *Sociological Mthodology* 8: 84-136

WHO (2012) Disease and injury country estimates: Burden of disease. WHO http://www.who.int/healthinfo/global_burden_disease/estimates_country/en/index.html/ (04/21/2013)

WHO-UNICEF (2013) WHO/UNICEF joint monitoring program for water supply and sanitation: Data and estimates. WHO http://www.wssinfo.org/data-estimates/table/ (04/21/2013)

Whittington D, Luria DT, Mu X (1991) A study of water vending and willingness to pay for water in Onitsha, Nigeria. *World Development* 19: 179-198

Wooldridge JM (2002) *Econometric analysis of cross section and panel data.* MIT Press, Cambridge

Yusuf AA, Koundouri P (2004) Household valuation of domestic water in Indonesia: Revisiting the supply driven approach. In Koundouri P (ed) *Econometric informing natural resource management: Selected empirical analysis*. Edward Elgar, Cheltenham