Australian National University

Working Papers in Trade and Development

The cost of floods in developing countries' megacities: A hedonic price analysis of the Jakarta housing market, Indonesia

José Cobián Álvarez

and

Budy P. Resosudarmo

April 2018

Working Paper No. 2018/05

Arndt-Corden Department of Economics

Crawford School of Public Policy

ANU College of Asia and the Pacific

The cost of floods in developing countries' megacities: A hedonic price analysis of the Jakarta housing market, Indonesia[1]

José Cobián Álvarez

Crawford School of Public Policy
Australian National University
E-mail: u5633161@alumni.anu.edu.au

and

Budy P. Resosudarmo

Indonesia Project
Arndt-Corden Department of Economics
Crawford School of Public Policy
Australian National University
E-mail: budy.resosudarmo@anu.edu.au

Abstract

***Abstract**:* Although many megacities in developing countries experience floods that affect a large number of people, relatively few empirical studies have evaluated the costs involved. This paper estimates such costs by conducting a hedonic price analysis of the impacts of floods on the housing market in Jakarta. A robust regression technique on a simple linear transformation model, and a maximum likelihood estimation technique on the spatial lag version of the simple linear transformation model are utilized to estimate the correlation between the level of the 2007 floods and monthly housing rental prices in Jakarta, Indonesia. This paper concludes that in developing countries' megacities the total cost of floods is not as considerable as the total estimated cost of making the city of Jakarta flood-free.

Key words: Environmental economics, hedonic price analysis, spatial analysis, flood.

JEL Code: Q51; Q54; R32; O21

[1] The authors would like to thank M. Agung Widodo for managing the IFLS data set for this paper. All mistakes are the authors' responsibility.

The cost of floods in developing countries' megacities: A hedonic price analysis of the Jakarta housing market, Indonesia

1. Introduction

Climate change is causing an increase in extreme weather and climate events. Developing countries are particularly vulnerable due to their geographic exposure, poverty, high dependence on agriculture, rapid population growth and limited capacity to cope with an uncertain climate. This leads to increased human exposure to natural disasters such as heatwaves, droughts, storms and floods, which are becoming more frequent as the world gets warmer (Stern, 2007). Among these major weather events, floods have been recognised as the major cause of economic damage worldwide which, in turn, affects a large number of people (UNISDR, 2002). More specifically, this phenomenon has become an annual event over the past few decades in many developing countries' megacities, and has heavily impacted Asia, where there is a large concentration of people in urban areas. (World Resources Institute, 2015).

In 2014, the level of urbanization in developing countries was approximately 48.4%, and the proportion of people living in urban areas in the Asian region was approximately 47.5% (UN, 2014). Urbanization in developing countries has brought on urban management challenges related to the lack of physical infrastructure and inadequate urban services (Cohen, 2004). In some cities, urban expansion has been unplanned or inadequately managed, leading to rapid sprawl, pollution, and environmental degradation, accompanied by unsustainable production and consumption patterns (UN, 2014).

An apparent lack of capability in managing urban development, as a result of high rates of urbanization and large populations, along with increasing climate variability and rising sea levels are typically suspected as the main causes of these floods. It is not uncommon that these floods annually cause serious natural disaster events in developing countries (UN and WB, 2010). A study undertaken by the World Resources Institute (2015) considered Indonesia to be one of the countries with the greatest number of people exposed to flood risk, ranking 6th out of 164 countries in 2010. Jakarta comprises the largest urban area in Indonesia with a population density of approximately 14 thousand persons per km^2 (Yusuf and Resosudarmo, 2009).

In Jakarta, the cause of flooding is due not only to increasing climate variability and rising sea levels, but also to the extensive use of ground water, which has caused subsidence in several areas (World Bank, 2011). Flooding is an annual disaster event in Jakarta and most of the time affects a significant number of residents in the city. For example, the 2007 floods were one of the more significant, inundating almost 36% of Jakarta city, in some areas to a depth of seven meters, resulting in over 70 deaths and 340,000 displaced people (Jha et al., 2012). In 2013 and 2014, Jakarta was again hit by major floods (Budiyono et al., 2016).

Due to growing concern over the impact of floods on Jakarta, local government and non-government organisations have been developing several intervention programs, including better managing the risk of disaster, and the resettlement of urban poor populations at the lower end of the scale, up to reducing greenhouse gas emissions (Baker, 2011). Several of these activities are as follows. Since 2012, with World Bank support, the Jakarta government has developed projects to dredge a number of vital floodways and retention basins, and has rehabilitated embankments and mechanical equipment that are part of Jakarta's flood management system. This includes work on 11 floodways or canals, comprising a total length of 67.5 kilometers, and four retention basins covering an area of

65 hectares. About 42 kilometers of embankments were rehabilitated or constructed within these floodways and retention basins (World Bank, 2016).

However, some constraints have proved to be an obstacle to the success of these initiatives, such as much-needed upgrades to city infrastructure, the significant lack of research and data regarding floods to support decision-making, and the absence of community engagement — both government and community — to take necessary action. The cost of the projects needed to mitigate floods in Jakarta is also not trivial. The Jakarta Water Management Agency estimated the city needs Rp. 118 trillion (USD 9.2 billion) — approximately twice the total revenue of Jakarta government in 2015 — to make Jakarta flood-free (Tambun et al., 2015). Therefore, reducing the flood risk in Jakarta still remains a challenge to be tackled by the Indonesian government, as a key priority within disaster management.

As has been mentioned already, although flooding is a significant occurrence for consideration by any government in developing countries, there has been little research and limited evidence of evaluating the cost to their megacities. Most research has focused on flood risk in developed countries, particularly the United States of America, and has studied the impact of flooding on the price differential of property values and their relation to insurance costs (Carbone et al., 2006; Bin and Landry, 2013; Bin and Polasky, 2004; Bin et al., 2008; Atreya et al., 2013).

Until recently, only a few studies have analysed the economic damage and loss due to flooding. Budiyono et al. (2015) identified areas of highest risk and assessed Jakarta’s risk using the damage scanner model. They also found the annual expected damage due to river flooding in Jakarta is approximately US$ 321 million per year, and obtained new estimates of economic exposure values for different land use classes (industry and warehouse, commercial and business, planned house, and density urban). In the same vein, the study undertaken by Wijayanti et al. (2017) measured flood damage in Jakarta but distinguished between residential and business sectors, with reported values of US$ 1.3 million and US$ 9.2 million in 2013, respectively. Wahab and Tiong (2017) proposed a multi-variate residential flood loss estimation model to estimate direct tangible loss to buildings and contents for the residential sector after the 2013 January floods. The results show that as water flood level (expressed in water depth) increases, the building structure and contents losses (expressed in terms of US$) tend to rise, but the tangible loss for the residential sector in Jakarta city is greater in higher than in lower income areas.

In an attempt to fill the recognised research gap, this paper will apply a hedonic property value analysis using data obtained from a combination of the Indonesian Family Life Survey (IFLS), and flood-level data in Jakarta obtained from the United Nations Department of Safety and Security (UNDSS). The main objective is to analyze whether major flood events are directly correlated with property values in Jakarta.

The paper is divided into five sections; the background and motivation for the research; the use of the hedonic property value method in previous studies; a description of the methodology and data utilized in the paper; the empirical results obtained from the data; the policy implications; and the concluding statements.

2. Literature Review

The hedonic price method provides an intuitive analytical tool for studying the effects of property attributes and spatially integrated amenities on housing prices. Lancaster (1966) pioneered the development of its theoretical foundations, derived from the theory of consumer demand. The central assumption is that consumer utilities are not based on the goods *per se*, but instead on the individual “characteristics” of goods — their composite attributes. Although Lancaster (1966) was the first to discuss hedonic utility, there was nothing about pricing models and the properties of market equilibrium. To fill this gap, Rosen (1974) studied the demand–supply interaction in which they bid (consumers) and offer (suppliers) the combination of attributes and prices of the goods that keep the market in equilibrium.

Additionally, Rosen’s (1974) studies form the basis for using the hedonic property price model to estimate the value of environmental amenities. The argument is that the attributes of residential properties — recognised as heterogeneous goods, such as structural, neighborhood and environmental characteristics — are reflected in the price differentials that affect lessee preferences in a market clearing equilibrium condition (Rosen, 1974). The advantage of using this method over other preference estimation techniques is that it makes use of actual market transactions to recover value estimates for non-market attributes (Bin et al, 2008).

Since then it has been widely utilized in environmental economics literature to estimate the price difference between residential properties located within or outside floodplain regions. Some of them can be seen in Table 1. Most of these studies demonstrate a negative relation between the housing prices and flood events, whereby the properties located in the floodplain are likely to be impacted by a price decrease, in comparison to those properties located in non-floodplain areas. Further to this, following a flood phenomenon, owners of houses located in floodplain areas are forced to pay an increased insurance premium. Skantz and Strickland (1987) note that house-price reactions to flood events initially declined and later regained their lost value due to the market forgetting about the flood event.

Using a semi-logarithmic functional form for hedonic property value analysis, they found there was no immediate decline in flooded-home prices after the flood event. This was due to the flood insurance premium being subsided by the federal government. A year later, when the government cut the economic support, floodplain houses experienced a decrease in property values.

Bin and Polasky (2004) also utilized the hedonic property price function to estimate the flood hazard effects on property values in Pitt County, North Carolina. The methodology used an OLS regression analysis which found that after Hurricane Floyd in 1999, houses located in a floodplain were impacted by a price discount. The marginal effect estimated for the property values located in the floodplain was approximately $ 7,463, i.e. the property value in the floodplain was lowered by that amount of money.

This formed the basis for the study undertaken by Bin and Landry (2013), which re-examined and compared findings with a previous flooding event regarding lessee preferences in a market clearing equilibrium condition — 1996 Hurricane Fran — using difference-in-difference (DID) and spatial effect models (spatial lag and spatial error). They found that average real property values decreased by approximately 5.7% after Hurricane Fran compared to approximately 8.8% after Hurricane Floyd; however, in between both hurricanes, they increased by approximately 2.2%. This price increase is due to the lessee becoming more insensitive to flooding events since the perception of flood risks and cost associated with it are not persistent over time.

Most of the published literature analyzing the relationships between floods and hedonic property value concerns the US (Table 1). There are some studies regarding other developed countries, such as the Netherlands (Daniel et al., 2007) and New Zealand (Samarasinghe and Sharp; 2008); and very few on developing countries. Among the few is a study by Rabassa et al. (2016) which attempts to determine whether flood events are associated with property values in La Plata city, Argentina. Using data from land parcel sales in 2004, they found that property sale prices were affected by a discount of approximately 17.3% for properties located in flood-prone areas, as opposed to those situated outside of the floodplain.

Another important characteristic regarding the most recent studies is their coverage of flood events not occurring on an annual basis. Floods in the southern part of the US might happen frequently, but only once every few years, so that the prices immediately following the shock may not yet be equilibrium prices.

This paper will apply the hedonic price method to see whether the annual flood events have an impact on the housing value, measured by monthly rental property price, in Jakarta, Indonesia. Since this is an annual event, though the size might vary annually, we can expect the housing rental market to be in its equilibrium condition.

3. Study Area and Data Sources

The city of Jakarta, the capital of Indonesia, is the study area of this paper. Jakarta has been one of the fastest-growing megacities in the world. Approximately 6.5 million people resided in this city in 1980 compared to more than 10 million people in 2016 (CEIC, 2017). The city lies on a low, flat alluvial plain formed by the mouth of the Ciliwung River (main river) where it meets Jakarta Bay. This river travels through the middle of the city and divides it into western and eastern areas. The Pesanggrahan and Sunter are less turbulent rivers and cross the western part of Jakarta. Thus, most of the city is prone to swampy and flooded conditions, especially during the rainy season (typically from October to April). Those parts of the city further inland are slightly higher but are also at risk of experiencing flood events (Baker, 2011).

Figure 1 shows a map of the study area and the flood water levels during the February 2007 flood event per sub-district level. As seen, locations with the highest flood level (dark red) are adjacent to the Ciliwung, Pesanggrahan and Sunter Rivers, especially in the southern area of Jakarta. However, the area of the city with more water coverage was northeast Jakarta, which includes the subdistricts of Kelapa Ganding, Pulo Gadung, Cakung, Danau Sunter, Kemoyoran, Tanjung Priok and Cilincing.

The map (Figure 1) and the data for the flood water levels by 'village' or *kelurahan* in Jakarta were taken from the United Nations Department of Safety and Security (UNDSS), which surveyed Jakarta in February 2007. The city is divided into five districts (known as *kotamadya*), which divide into 42 subdistricts (known as *kecamatan*). Each subdistrict is comprised of approximately 2 to 5 kelurahan. The UNDSS collected and reported the water levels of the 2007 Jakarta flood from news sources (radio and television), and United Nations Staff Reports to UNDSS Offices and Police Stations[2].

The flood water level to be studied in this paper (which is in Figure 1) corresponds to the water level (measured in centimeters) registered immediately following the flood event on 6 February 2007. This information was gathered at the 'village' level. For our analysis in this paper, we calculate the weighted average flood water level in each subdistrict. The

[2] https://trip.dss.un.org/dssweb/WelcometoUNDSS/tabid/105/

'village' area (measured in square meters) within each subdistrict is used as a weight to estimate the average water level for each subdistrict. The reasons for aggregating the flood information at subdistrict level are as follows. First, floods in one kelurahan will certainly affect their neighbouring kelurahan; second, floods are typically managed at subdistrict level; and third, for security reasons, household information only contains coded locations at the subdistrict level.

The other data used for this paper is cross-sectional, extracted from the 2007 Indonesia Family Life Survey (IFLS) dataset. The dataset contains information on monthly house rent, housing characteristics and neighborhood characteristics[3]. There are as many as 1,573 observations for the city of Jakarta. This sample arguably represents the population of Jakarta.

The variables selected for the hedonic price analysis are those commonly used in hedonic property value studies (Yusuf and Koundouri, 2005; Yusuf and Resosudarmo. 2009) and are available in the IFLS dataset. Monthly house rental price expressed in rupiahs (Indonesian currency) is used as a proxy of housing value. Meanwhile, housing characteristic variables are house size (expressed in square meters); number of rooms; wall, roof and floor materials; water source availability; and yard at the house. The wall, roof and floor materials are dummy variables which have been assigned a value of one if they are constructed from a reasonably durable material; i.e. cement/brick for walls, concrete/roof tiles for roof and cement/stone for floor, or otherwise they are given a value of zero. Water source is also a dummy variable of 1 if there is a water source inside the house, or otherwise zero. The existence of a yard is valued as 1, otherwise as zero. These variables are expected to bear a positive relationship to the monthly house rent.

[3] https://www.rand.org/labor/FLS/IFLS.html

We also include neighborhood characteristics at the kelurahan level in Jakarta, namely unemployment rate, percentage of people with a university education, whether or not public transport is accessible and whether a house is located along a river basin, distance from the district centre, and the traffic congestion level. The information extracted is at the 'village' (*kelurahan*) level.

The variables for the unemployment rate, the distance to the centre of Jakarta, and the settlements along riverbanks are expected to be negatively associated with the dependent variables; the variables for the percentage of people in the 'village' with a university education, and the accessibility of public transport, are estimated to be positively related to monthly housing rent.

The environmental variable includes the 2007 flood experience in Jakarta, recorded as the water level measured in centimeters, and it is expected to be positively associated with house rent.

Table 2 provides a detailed description and summary of the variables that are utilized in the hedonic price model.

4. Methodology

Considering the hedonic hypothesis as a basis that the goods are valued by the utility of their attributes or characteristics, Rosen (1974) developed a heterogeneous product model whereby the implicit prices or values of the attributes are estimated, and the sum of which equal the observed transaction price. This implies that they cannot be traded separately, but jointly commercialized in a unique market as a single good.

The hedonic price emerges from the interaction of buyers and sellers and represents a market clearing equilibrium, based on the following assumptions: (1) continuity in the levels of attributes — the amount and qualities of the attributes associated with the heterogeneous products are reflected in price differentials; and (2) full information about prices and attributes (Rosen, 1974).

Since the pioneering work by Rosen (1974), many studies have applied the hedonic price model to assess the attributes of land and residential properties, but also to estimate the value of environmental amenities — non-market characteristics that also affect house prices. These related to aesthetic sights and their closeness to recreational sites such as parks, and beaches, as well as the quality of the environment in terms of air, water and noise pollution.

According to this method, the hedonic price function is typically represented as:

$$P_i = f(s, n, l, e)$$

where P_i is the price of property i which is a function of structural characteristics (e.g. house size, number of rooms, quality of walls), s; neighborhood characteristics (for example, ethnic composition, crime rate, flow of traffic), n; location characteristics (e.g. proximity to economic centres, distance to highways, accessibility to public transport), l; and environmental characteristics (such as air pollution and flooding), e. Therefore, characteristics that generate benefits for households, such as a larger number of rooms or home size, increase the property's price; while characteristics that imply costs for households, such as a neighborhood with a high crime rate, reduce the property's price. This method also makes inferences about non-observable values of different attributes in the housing market such as air pollution and flooding.

Given the basis of the method is to find what portion of the price is determined by the hedonic variable, we obtain the environmental attribute (which is flooding) by calculating the partial derivative of the price with respect to the variable e, $\partial P_i / \partial e$. It gives us the marginal implicit value for an additional unit of the environmental asset, and thus enables an estimate of its monetary value.

The theoretical model specified in equation (1) will be utilized to calculate the implicit price (or discount) of the flood event on rent price as the environmental variable considered in this study. The model is an ordinary least square (OLS), which is commonly utilized in a hedonic property value analysis. In addition, a robust regression technique is applied to the equation to produce a robust estimate of variance and to ensure that coefficients estimated are more efficient (Hubert, 1973):

$$y = \beta_0 + \boldsymbol{x_1} \boldsymbol{\beta_1} + \boldsymbol{x_2} \boldsymbol{\beta_2} + f \beta_3 + \varepsilon \tag{1}$$

where y is the logarithmic form of the monthly rent of the house which is the proxy for housing value, $\boldsymbol{x_1}$ is a vector of housing attribute variables and $\boldsymbol{x_2}$ is a vector of neighborhood characteristics. The variable f is the logarithm form of the flood water level. Meanwhile ε is the error term.

In this hedonic analysis, it is assumed that the lessee makes a rental decision accepting all the housing characteristics, and so the property value is a function of the heterogeneous characteristics of the property. It is suspected that this hedonic property value function is a nonlinear function of its characteristics and many of the variables involved are not normally distributed; and so, a transformation function technique is usually adopted. The Box–Cox transformation model is most commonly used in hedonic price analysis (Cropper et al., 1988; Yusuf and Resosudarmo, 2009).

However, in this paper, we adopt a simpler transformation technique; we transform and normalize the dependent variable (monthly housing rent), the continuous explanatory variables (house size, distance to the business centre and congestion level), and the study variable (flood level) using the logarithmic functional form.

Previous studies show the dependent variable log-transformed due to the significant variation in the housing price variable (Skantz and Strickland 1997; Bin and Polansky 2004; Daniel et al. 2007; Bin et al. 2008b; Samarasinghe and Sharp 2010; Pope 2008; Kousky 2010; Bin and Laundry 2013). Taking the logarithm of the explained variable minimizes the possibility of heteroscedasticity (Gujarati 1995; Wooldridge 2003) or corrects for it between house price (or house rent) and the residuals (Basu and Thibodeau, 1998).

In addition, a better R^2 value is obtained when we consider the average room size in the OLS model, instead of the number of rooms. Therefore, the average size of a room is measured using a variable proxy (size/room).

Implicit in this model is the assumption that the differential effect of the housing characteristics (house size, rooms, wall, roof and floor materials, water access, and yard house) are constant across the flood water level, and the differential effect of the flood event is also constant across the property’s attributes. That is to say, if the mean housing rental price is higher for a large than for a little house, this is so whether the house is located in a flood-plain area or not. Likewise, if for example a house in a flood-plain area has a lower mean rental price, this is so whether it is an apartment or a condominium.

As mentioned before, in this paper, an average flood water level is used for subdistrict areas. One reason for analyzing subdistrict areas is to take into account the impact of nearby flooding on the value of property in a certain area. A subdistrict in the Jakarta context is relatively large enough; however, there is still a possibility that average flood water levels in neighboring subdistricts affect the property value in a subdistrict (Yusuf and Resosudarmo, 2009).

Anselin (1988) introduced a concept of spatial dependence to determine the relationship among the property values in neighboring locations. Several studies have incorporated this analysis to estimate the real impact of all the housing attributes — such as Daniel et al. (2007), Bin et al. (2008), Cho et al. (2009), Samarasinghe and Sharp (2010), Bin and Landry (2013) and Rabassa et al. (2016) — which suggests the presence of this spatial effect in a cross-sectional hedonic price analysis. Ignoring this estimation, the resulting coefficients from the OLS model could be inefficient or inconsistent (Anselin, 1988).

To capture the neighboring spillover effect, this research paper uses the spatial lag model[4] proposed by Anselin (1988) and adopted by various studies (Leggett and Bockstael, 2000; Brasington and Hite, 2005; Daniel et al., 2007; Bin et al., 2008; Cho et al., 2009; Yusuf and Resosudarmo, 2009; Samarasinghe and Sharp, 2010; Bin and Landry, 2013; Rabassa et al., 2016).

This assumes that the housing rental price depends both on its characteristics (structural and neighborhood) and on neighboring house rental prices; i.e. the spatial lag model includes

[4] Similarly, a spatial error model can be considered, which supposes that spatial dependence arises due to measurement errors or some omitted variables that are correlated and vary spatially. The Lagrange Multiplier (statistic=26.489; p-value=0.000) and the Robust Lagrange Multiplier (statistic=17.467; p-value=0.000) tests show spatial error dependence.

the spatially-weighted sum of neighboring house rental prices as the independent variable in the functional form of the housing price formation:

$$y = \beta_0 + \rho \boldsymbol{W y} + \boldsymbol{x_1 \beta_1} + \boldsymbol{x_2 \beta_2} + f \beta_3 + \varepsilon \tag{2}$$

where ρ is the spatial dependence parameter and $\boldsymbol{W}$ is an nxn standardized spatial weight matrix (where *n* is the number of observations). The spatial matrix, $\boldsymbol{W}$, tells us whether any pair of observations are neighbors. If, for example, house *i* and *j* are neighbors, then $\mathrm{w}_{i,j} = 1$ and zero otherwise, for all i ≠ j. Please note that $\mathrm{w}_{i,i} = 0$ for all i.

Whether any pair of houses is neighboring in this paper is determined by them sharing some common borders (contiguity). The spatial weight matrix is usually standardized, such that every row of the matrix is summed to 1. This enables us to interpret the spatial lag term in a spatial model as a simply spatially-weighted average of neighboring house prices. The spatial lag model will be estimated using a maximum likelihood (ML) regression technique (Anselin, 1988).

Results and Discussion

Table 3 shows the results of estimating the basic and spatial lag models; i.e. equations (1) and (2), respectively. From the result for the spatial lag model, it can be seen that the ρ estimate is significant at 5%; and by comparing results for the basic and spatial lag models, it can also be seen that while most coefficients are almost similar, the coefficients for roof material, water source, house yard, public transport access, distance to business centre, traffic flow and flood water level are relatively different. These results indicate that spatial dependence plays an important role in the process of formulating housing rental prices in the Jakarta housing market; i.e. estimated coefficients of the basic model are likely to be inefficient or inconsistent. The results from the spatial lag model are argued to be superior to those of the basic model.

As can be seen in table 3, the value of the adjusted R^2 in the OLS model is 0.423, whereas for the Spatial Lag model the value of the variance ratio is 0.431 and the squared correlation value is 0.432. This indicates a better fit of the model to the observed data. The Moran's I statistic shows a negative spatial autocorrelation in house rental prices in 2007, denoting that observations with similar rental prices are dissimilar when compared.

In order to evaluate the functional form of equation (1), we use the regression specification error test (RESET test). A linear regression model is correctly specified when there are no omitted variables, i.e. the null hypothesis equals zero. In this study, the OLS regression produces an F-statistic of 10.93 with a p-value of 0.0000, which indicates a specification problem. This indicates that additional explanatory variables are required to be included in the OLS model, since housing rent prices could also be affected by characteristics; however, we put as many available control variables as possible in the model, and conducted a spatial analysis.

Let us observe the results for the spatial lag model. Five out of seven house structural characteristics, i.e. house size, size of rooms, wall and floor materials and house yard, are positively associated with the house rental price. This is as expected. Estimated coefficients for these variables are strongly significant at the 1% level, except for the house yard, which is not significant at a conventional level. The other two estimated coefficients, i.e. roof material and water source, are negatively related to the dependent variable but not statistically significant; however, the negative signs are unexpected.

An explanation for this could be the greater presence of houses with non-cement/brick walls and without a water source in the sample, whereby a better quality of wall material and access to water in the house may not necessarily be a good thing for lessees. Interestingly, according to the study undertaken by the World Bank (2015), 28.9 million units, or 45% of existing homes in Indonesia, are considered substandard according to one of the following factors: poor quality housing material (such as roof, wall and floor); do not have access to basic utilities (water and sanitation); or, are overcrowded (less than 7.2 m^2 per capita). Houses with substandard roofing material (defined as asbestos and fiber/palm) amount to 9% or 5.9 million units, while those with no access to a water supply represent 14% or 8.8 million units.

All estimated coefficients for neighborhood qualities have the expected sign. Three out of six comply with expectations and are statistically significant correlated with housing rent price; i.e. with the coefficient for the percentage of people with a university degree, the distance to the centre and the house located along a river basin, all significant at the 1% level. The coefficient for the percentage of people with a university degree is positively related to housing rental price. The distance to the centre of Jakarta is negatively associated with housing rental price; meaning the closer the house is to the business centre, the higher the rental price charged to the tenant. Finally, the closer the house is to the river basin, the lower the housing rental price.

On the main variable of analysis in this paper, namely flooding, it can be seen that the coefficient of the flood variable is negative and statistically significant at 5%. This coefficient suggests that a higher flood water level is associated with a lower housing value.

Further analysis is needed to understand the full association between flood water level and housing rental prices. The first derivation of the spatial lag model for the hedonic housing value is as follows:

$$\frac{\partial y}{\partial f} = \beta_3 \tag{3}$$

Let us insert in equation (3) the average value of $\boldsymbol{x_1}$, then we have $\frac{\partial y}{\partial f} = -0.1279$. This indicates that, on average, an increase in flood water level of 1% is associated with a 0.128% lower housing value.

Inserting the average flood water level in Jakarta in 2007, which was approximately 42.33cm, it can be roughly concluded that flooding in Jakarta lowers the monthly housing value by Rp. 619 thousand or approximately 12.8% of the average housing rental price in Jakarta. Comparing this to previous studies for other countries recorded in Table 1, it can be seen that 12.8% is within the range of results from previous literature.

If this Rp. 619 thousand can be interpreted as the average monthly willingness of a household to 'permanently' get rid of the cost of flooding, i.e. the capitalized marginal willingness to pay (MWTP), and assuming that there are approximately 10 million people or 1.8 million households in Jakarta having houses with an average lifetime of 25 years and a discount rate of 5% annually, it can be estimated that the total willingness of all households in Jakarta to permanently get rid of the cost of flooding is approximately Rp 40.5 trillion or approximately 7.2% of Jakarta's GDP in 2007.

The formula to calculate the capitalized MWTP is as follows:

$$W = \sum_{t=0}^{25} (1+r)^t \, (1.w) \tag{4}$$

where W is the capitalized marginal willingness to pay; i.e. how much a household is willing to pay for a 'permanent' (typically 25 year) reduction of a unit of pollutant, and w is the marginal willingness to pay per month, i.e. marginal effect of hedonic equation, while r is a discount rate of 5% and t is year.

It can be interpreted that the reconstruction and rehabilitation process in developing countries' megacities after a flood is more likely to take considerably longer than in developed countries, due to the gap between the flood cost and the flood-risk management values. This can be seen in the case of Jakarta where the regional government needs Rp 118 trillion to make the city flood-free, while the total flood cost is Rp 40.5 trillion. The difference reveals the difficulty for Jakarta local authorities to recover from flooding, making the post-disaster reconstruction slower. By comparison, in developed countries it is generally much faster to restore and rebuild after floods because of a better flood prevention system, developed infrastructure, and emergency-response plans (Laframboise and Loko, 2012; UN, 2016).

It is important to note that house characteristics variables are mainly represented by dummies (wall, roof and floor materials, water availability, and house yard), so that it is most likely that the OLS model has a certain degree of multicollinearity. This might explain the low significance levels and opposite signs obtained from a linear regression. Variance inflation factors (VIFs) are used to test for multicollinearity among the independent variables. VIF is an index that shows the existence of multiple correlation coefficients between a single variable and the rest of the independent variables, and indicates the magnitude of the inflation in the standard errors.

According to Gujarati (1995), multicollinearity may be a problem if the VIF is greater than 10. In this study, only two variables (LOG(Traffic) and LOG(Distance)) were found to have a VIF value above 2.0. Overall, the mean of the VIF values for all of the variables was 1.50 for the OLS regression. This means that there is no multicollinearity or no correlation between ε_i and the independent variables.

Floods and Human Health Conditions

In an attempt to understand why people's attitudes differ regarding houses in flood prone areas and those that are not; i.e. in general, people place less value on houses in flood prone areas, this paper explores the relationship between human health conditions and housing characteristics, including flood water levels. Human health indicators utilized in this paper are number of restricted activity days (or number of days with daily activities disrupted due to feeling sick in the past 4 weeks) and the case of depression (whether or not there is a member in the household who has suffered from depression in the past week; one if yes there is, and otherwise zero), both of which are available in the 2007 IFLS dataset.

Number of Restricted Activity Days

We adopt the basic and spatial lag models as shown in equations (1) and (2), respectively, where y is the logarithmic form of the number of restricted activity days. Table 4 presents the results of estimating the relationship between number of disrupted daily activities due to poor health and flood water level, implementing the basic and spatial lag models.

Comparing results using the basic model and the spatial model in Table 4, it can be seen that implementing a spatial lag model is not needed, since the estimated ρ is not significant at the convenience level, i.e. there is no evidence of spatial dependence in the data, and the tests based on the principle of Lagrange Multiplier Lag and Robust Lagrange Multiplier Lag are not statistically significant at any level of significance. In other words, apart from the level of flooding in the area, floods in neighboring areas are not associated with this type of human health condition.

Flood coefficients in number of restricted activity days are significant at 10% and have the expected sign. A higher level of flood water level is associated with a higher number of restricted activity days. Using a similar strategy to that in equation (3), we estimate the overall association between poor health conditions and flood water levels. In the case of restricted activity days, the result is $\frac{\partial y}{\partial f} = 0.0713$; i.e. an increase in flood water level of 1% is associated with a 0.07% higher number of restricted activity days in a household.

People suffering depression

The dependant variable is a binary response nominal variable: this is the probability of a household member (PD) having suffered depression in the last week). It is a binary response nominal variable since it only takes the values 0 and 1. PD = 1 if a household member has felt depressed in the last week, and PD = 0, otherwise. The probit model is represented as follows:

$$Probit\ (PD) = \boldsymbol{\beta_0} + \boldsymbol{x_1\beta_1} + \boldsymbol{x_2\beta_2} + f\ \beta_3 + \varepsilon \qquad (5)$$

where PD is the probability of the member of the family having suffered depression in the last week, and the explanatory variables are elements from equation (1). The output of this non-linear model must be converted into marginal effects, so that it is possible to measure the impact of the independent variables on the dependent variable.

Table 5 presents the average marginal effects (AME, marginal effect at the mean of each independent variable). The flood coefficient is significant at 5% and has the expected sign. A higher level of flood water level is associated with a higher incidence of feeling depressed in a household. However, if observations of the dependent variable are similar to those in nearby locations, spatial dependence exists within the data, and ignoring it (as in OLS model) will result in inconsistent and inefficient estimated coefficients (Anselin, 1998). In this paper, the presence of this spatial effect in the binary probit model is tested.

Spatial probit models have been estimated using the ML method (McMillen, 1992; Murdoch, Sandler, and Vijverberg, 2003), and generalized method of moments estimators (GMM) (Pinkse and Slade, 1998). However, the traditional ML is less practical because the likelihood function involves n integrals and the determinant of the $n \; x \; n$ matrix (where n is the sample size), which implies a computational burden (Pinkse and Slade, 1998).

Pinkse and Slade (1998) proposed a two-step GMM estimation for the spatial-error probit model. The advantage of using this method lies in the fact that it does not rely on the assumption of normally distributed errors, and it does not require the calculation of the determinants and inverses of $n \; x \; n$ matrices because it is based on the two-stage least squares technique (Klier and McMillen, 2008). Hence, we conduct spatial analysis by estimating a spatial-error probit model, and this is discussed in more detail in Appendix A.

The result suggesting the presence of a spatial-error effect in the probit model is also shown in Table 5. The λ estimate is statistically significant at 1%, which means there are spatial dependence effects in the observations. In other words, the probability of having a household member feel depressed rises in flood-prone areas but people living in neighboring areas also suffer the same symptoms. Column 2 contains estimates that are corrected for spatial error probit correlation using the nearest-neighbor weights.

Numbers in brackets are Robust Standard Errors; AME is Average Marginal Effect (evaluated at the mean of each independent variable).

Using equation (3), we estimate the overall association between human health conditions and flood water levels. The result is $\frac{\partial y}{\partial f} = 0.087$ in the case of feeling depressed; i.e. a 10cm increase in flood water level is associated with 8.7% increase in the probability of having a depression incident in the household.

From this result, it can be argued that houses in flooding areas have been associated with a worsening in human health conditions. This could be a reason why people put a lower value on a house in a flood prone area than one in an area that does not flood, given similar characteristics of the house and the neighborhood.

Furthermore, in general, by using these results it can be estimated that flooding in Jakarta is associated with approximately 1 restricted activity day and approximately a 14% probability of diagnosable depression in a household annually; or, in total, flooding in Jakarta is associated with approximately 2.4 million cases of restricted activity days and approximately 257 thousand cases of depression symptoms annually.

Conclusion

This study is an attempt to estimate the cost of flooding in developing countries' megacities by conducting a hedonic price analysis of the Jakarta housing market. It estimates the correlation between levels of flooding and monthly housing rental prices in Jakarta in 2007.

Data on the flood water levels by 'village' or *kelurahan* in Jakarta were obtained from the United Nations Department of Safety and Security (UNDSS), which collected and reported the water levels of the 2007 Jakarta flood from news sources (radio and television), and United Nations Staff Reports to the UNDSS Office and Police Stations. Data on monthly housing rental prices and other information related to house and neighborhood characteristics are taken from the IFLS for 2007.

The empirical results indicate that a one percent high flood water level is associated with a 0.128% lower monthly housing rental price; or, on average, flooding in Jakarta is associated with lowering monthly housing values by approximately Rp. 619 thousand. Furthermore, if this number can be interpreted as an average monthly willingness of a household to 'permanently' get rid of the cost of flooding, and assuming that there are approximately 1.8 million residential houses in Jakarta with an average lifetime of 25 years and an annual discount rate of 5%, the cost of flooding for households in Jakarta is approximately Rp 40.5 trillion or approximately 7.2% of Jakarta's GDP in 2007.

This paper also found that a lowering in human health conditions could be the reason that households put less value on houses located in flood prone areas compared to those on higher land. This paper estimates that a one percent higher flood water level is associated with a 0.07% higher number of restricted activity days in a household, and a 10 cm increase in flood water levels is associated with an 8.7% increase in the probability of suffering a depression symptom in the household. In general, using this result, it can be estimated that flooding in Jakarta is associated with approximately 1 restricted activity day and approximately a 14% probability of having a depression symptom in a household annually; or, in total, flooding in Jakarta is associated with approximately 2.4 million cases of restricted activity days and approximately 257 thousand cases of depression annually.

The Jakarta Water Management Agency has estimated the city needs Rp 118 trillion to make Jakarta flood-free (Tambun et al., 2015). This number is higher than the paper's estimate of the cost of flooding for households in Jakarta; i.e. approximately Rp 40.5 trillion. It can be seen, therefore, that it will be challenging for the Jakarta government to extract resources from its society to fund projects to eliminate flooding in the city. External resources from the central government are most likely needed to resolve the problem of flooding in Jakarta.

Appendix A. Spatial analysis of the case of a discrete variable

The spatial analysis of the probit model is viewed as an approximation because the structure of the spatial dependence is rarely known; however, what is known is that the errors tend to be correlated over space. We start with the structural model for the latent variable of the spatial-error probit that takes the following form:

$$y^{\bullet} = X\beta + u \tag{6}$$

$$u = \lambda W u + \varepsilon \tag{7}$$

where $y^{\bullet}$ is regarded as an unobserved scalar (latent variable), W is an $n\ x\ n$ spatial weights matrix, λ is the spatial autoregressive parameter, y is the observed value of the limited-dependent variable, and X is a matrix of explanatory variables.

Latent-variable $y^{\bullet}$ –a continuous variable $y^{\bullet} \in (-\infty, \infty)$ – links to the observed binary – outcome y through the measurement equation (observed model):

$$y = \begin{cases} 1\ if\ \ y^{\bullet} > 0 \\ 0\ \ if\ y^{\bullet} \leq 0 \end{cases} \tag{8}$$

with $\varepsilon \sim N\ (0, \sigma_\varepsilon^2 I_n)$, equations (6) and (7) can be written in a reduced form as:

$$y^\bullet = X\beta + (I_n - \rho W)^{-1}\varepsilon \tag{9}$$

where u (a vector of errors) has mean zero and the variance-covariance matrix is proportional to $(I - \lambda W)^{-1}(I - \lambda W')^{-1}$, i.e. $u = MNV\ (0, \Omega)$, in which the diagonal elements of $E(uu') = \Omega$ contains σ_i^2 vary across observations. This implies both heteroscedasticity and autocorrelation for u, unless λ=0, and makes estimates inconsistent.

The matrix notation $E(uu') = \Omega$, with a diagonal matrix as follows:

$$u = \begin{bmatrix} \sigma_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_n \end{bmatrix}$$

Equation (8) can also be written as:

$$\begin{aligned} Pr(y_i = 1) &= Pr(y_i^\bullet > 0 | X_i, w_{ij}y_i{}^\bullet) \\ &= Pr(X_i\beta + u_i > 0 | X_i, w_{ij}y_i{}^\bullet) \\ &= Pr(-u_i \leq X_i\beta | X_i, w_{ij}y_i{}^\bullet) \\ &\cong \phi(X_i\beta) \end{aligned} \tag{10}$$

Equation (10) is the probability for the single i-th observation when $y_i = 1$, $Pr(y_i = 1)$, so that we have a probit model, where ϕ represents the cumulative normal distribution function, so that ϕ(.) ∈ [0,1]. Since equation (6) induces a non-spherical disturbance, in which u is distributed n dimensional multivariate normal (*MNV*), intuitively, ε is also *MNV* with mean 0 and non-spherical variance-covariance $\sigma_\varepsilon^2 I$.

Due to the presence of heteroscedasticity ($\sigma_i^2$is not constant), the spatial error probit model for the single i-th observation is as follows:

$$Pr\left(y_i = 1|X_i, w_{ij}y_i^{\bullet}\right) = \phi\left(\frac{X_i\,\beta}{\Omega_{ii}}\right) \tag{11}$$

The marginal distribution of the *MNV*, denoted by $\phi_i(.)$, requires to evaluate simultaneously the joint probabilities at each location. The unknown parameters β and λ can be obtained by maximizing the log likelihood function as follows:

$$l(\beta, \lambda) = \ln[L(\beta, \lambda)] = \sum_{i=1}^{n} y_i\,\{\ln F(X_i\beta) + (1 - y_i)\ln[1 - F(X_i\beta)]\} \tag{12}$$

But now that we have the normal distribution in equation (11), the log-likelihood of y_i given β and λ, can be written as:

$$l\left(\beta, \lambda\middle|X_i, w_{ij}y_i^{\bullet}\right) = \sum_{i=1}^{n} y_i\,ln\phi\left(\frac{\mathrm{X}_i'\,\beta}{\Omega_{ii}}\right) + \sum_{i=1}^{n}(1 + y_i)\,ln\phi\left(1 - \frac{\mathrm{X}_i'\,\beta}{\Omega_{ii}}\right) \tag{13}$$

However, given the heteroscedasticity and the non-independence of the u_i, under a multivariate distribution, the joint distribution is not the product of the n marginal distributions, so that a binary probit observation does not maximize the sum of logs of n additively separable one-dimensional probabilities; instead, it maximizes the log of one non-separable n dimensional distribution (Franzese et al., 2010).

Due to the problem of endogeneity and non-sphericity of the variance-covariance matrix, Pinkse and Slade (1998) suggested the GMM estimation because it does not rely on the normally distributed errors and reduces the computational burden of the MLE estimation. As stated by Arbia (2014), they introduced the generalized errors in the context of the probit model because the vector (β, λ) is not known. Then, the i-th generalized error is given by:

$$\tilde{u}_i = \mathrm{E}[u_i|y_i, \beta, \lambda] = \frac{\phi\left(\frac{\mathrm{X}_i' \beta}{\sigma_i}\right)}{\phi\left(\frac{\mathrm{X}_i' \beta}{\sigma_i}\right)\left\{1 - \phi\left(\frac{\mathrm{X}_i' \beta}{\sigma_i}\right)\right\}}\left\{y_i - \phi\left(\frac{\mathrm{X}_i' \beta}{\sigma_i}\right)\right\} \tag{14}$$

Pinkse and Slade (1998) consider a set of k instruments that might include the explanatory variables, which are arranged in a n-by-k matrix Z. Since Z contains instruments that are exogenous, they suggest the moments condition as follows (Arbia, 2014):

$$E(Z'\tilde{u}) = 0 \tag{15}$$

Using equation (14), the moments condition for the i-th condition is:

$$E = \left[z_i \frac{\left\{y_i - \phi\left(\frac{\mathrm{X}_i' \beta}{\sigma_i}\right)\right\}\phi\left(\frac{\mathrm{X}_i' \beta}{\sigma_i}\right)}{\phi\left(\frac{\mathrm{X}_i' \beta}{\sigma_i}\right)\left\{1 - \phi\left(\frac{\mathrm{X}_i' \beta}{\sigma_i}\right)\right\}}\right] = 0 \tag{16}$$

where z_i indicates the i-th row of a matrix of instruments Z.

Finally, the following equation shows the set of conditions of the GMM estimation:

$$m(\beta, \lambda) = \frac{1}{n}\sum_{i=1}^{n} h_i \frac{\left\{y_i - \phi\left(\frac{\mathrm{X}_i' \beta}{\sigma_i}\right)\right\}\phi\left(\frac{\mathrm{X}_i' \beta}{\sigma_i}\right)}{\phi\left(\frac{\mathrm{X}_i' \beta}{\sigma_i}\right)\left\{1 - \phi\left(\frac{\mathrm{X}_i' \beta}{\sigma_i}\right)\right\}} \tag{17}$$

Under the GMM estimation, the number of moments conditions is greater than the number of parameters (β) to be estimated, so the set of values for β and λ minimizes the following equation:

$$m(\beta,\lambda)'M^{-1}m(\beta,\lambda) = min \tag{18}$$

where M is a positive- definite matrix, which defines the weights assigned to each sample moments $m(\beta,\lambda)$. This proves the consistency and asymptotic normality of the GMM procedure and derived the variance-covariance matrix $\Omega = (I_n - \rho W)^{-1}(I_n - \rho W^T)^{-1}$ of the unknown parameter vector of β and λ.

The GMM procedure implies a computational burden for large sample sizes since it requires the variance-covariance matrix Ω (that has a complex form) to be inverted in each iteration of the parameter λ.

However, the advantage of this procedure is that it provides the asymptotic variance of the estimator for a binary spatial error model, and also develops the hypothesis test for spatial error correlation (Calabrese and Elkink, 2013) based on the test on $\tilde{u}_i$ which is the generalized residuals corrected for heteroscedasticity.

To obtain the estimates and test the presence of spatial dependence for our spatial probit model using GMM, we follow the steps suggested by Arbia (2014), assuming an initial value for the vector $|\lambda| < 1$ before starting the iterative search of a solution. In that sense, we use a starting value of $\lambda_0 = 0.4$. This is explained in two ways: (1) when λ_0 takes values between 0.7 and 0.9, the computation time for each iteration increases in comparison with λ_0 ranging between 0 and 0.6; (2) the LOG(flood) estimates are approximately the same; however, as λ_0 increases, LOG(flood) estimates vary significantly.

References

Anselin, L. 1988, *Spatial Econometrics: Methods and Models,* Kluwer Academic Publishers, Boston.

Arbia, G. 2014, *A primer for spatial econometrics with applications in R*, Palgrave Macmillan, Hampshire.

Atreya, A., Ferreira, S. & Kriesel, W. 2013, ‘Forgetting the flood?: an analysis of the flood risk discount over time’, *Land Economics*, vol. 89, no.4, pp. 577-596.

Baker, J. 2011, *Climate Change, Disaster Risk, and the Urban Poor: Cities Building Resilience for a Changing World*, World Bank, Washington.

Basu, S. and Thibodeau, T.G. 1998, ‘Analysis of spatial autocorrelation in house prices’, *Journal of Real Estate Finance and Economics*, vol. 17, pp. 61–85.

Bin, O., Crawford, T., Kruse, J. & Landry, C. 2008, ‘Viewscapes and flood hazard: coastal housing market response to amenities and risk’, *Land Economics*, vol. 84, no. 3, pp. 434-448.

Bin, O. & Landry, C. 2013, ‘Changes in implicit flood risk premiums: empirical evidence from the housing market’, *Journal of Environmental Economics and Management*, vol. 65, no. 3, pp. 361–376.

Bin, O. & Polasky, S. 2004, ‘Effects of flood hazards on property values: evidence before and after hurricane Floyd’, *Land Economics*, vol. 80, no. 4, pp. 490-500.

Bin, O., Kruse, J. & Landry, C. 2008, ‘Flood hazards, insurance rates and amenities: evidence from the coastal housing market’, *Journal of Risk and Insurance*, vol. 75, no. 1, pp. 63-82.

Brasington, D.M. & Hite, D. 2005, ‘Demand for environmental quality: a spatial hedonic analysis’, *Regional Science and Urban Economics*, vol. 35, no. 1, pp. 57–82.

Budiyono, Y., Aerts, J.C.J.H., Brinkman, J., Marfai M.A., & Ward, P.J. 2015, ‘Flood risk assessment for delta mega–cities: a case study of Jakarta’, *Natural Hazards*, vol. 75, pp. 389-413.

Budiyono, Y., Aerts, J.C.J.H., Tollenaar, D. & Ward, P.J. 2016, ‘River flood risk in Jakarta under scenarios of future change’, *Natural Hazards and Earth System Sciences*, vol. 16, pp. 757-774.

Calabrese, R. & Elkink, J.A. 2013, ‘Estimators of binary spatial autoregressive models: A Monte Carlo study’, *Journal of Regional Science*, vol. 54, no. 4, pp. 664–687.

Carbone, J., Hallstrom, D. & Smith, V. 2006, ‘Can natural experiments measure behavioural responses to environmental risks?’, *Environmental & Resource Economics*, vol. 33, no. 3, pp. 273-297.

CEIC Data 2017, see CEIC Data Company Limited.

Cohen, B. 2004, ‘Urban Growth in Developing Countries: A Review of Current Trends and a Caution Regarding Existing Forecasts’, *World Development*, vol. 32, no. 1, pp. 23-51.

Cropper, M., Deck, L. & McConnell, K. 1988, ‘On the choice of funtional form for hedonic price functions’, *The Review of Economics and Statistics*, vol. 70, no. 4, pp. 668-675.

Cho, S., Clark, C., Park, W. & Kim, S. 2009, ‘Spatial and temporal variation in the housing market values of lot size and open space’, *Land Economics*, vol. 85, no. 1, pp. 51-73.

Daniel, V., Florax, R. & Rietveld, P. 2007, ‘Long term divergence between ex-ante and ex-post hedonic prices of the Meuse River flooding in The Netherlands’, Discussion Paper, European Regional Science Association, viewed 14 April 2016, <http://www.tbm.tudelft.nl/fileadmin/Faculteit/CiTG/Over_de_faculteit/Afdelingen/Afdeling_Waterbouwkunde/sectie_waterbouwkunde/people/personal/gelder/publications/citations/doc/citatie_806.pdf>.

Franzese, R.J., Hays, J.C. & Schaeffer, L.M. 2010, ‘Spatial, temporal, and spatiotemporal autoregressive probit models of binary outcomes: estimation, interpretation, and presentation’, Annual Meeting of the American Political Science Association, viewed 9 October 2017, https://www.researchgate.net/publication/228237952_Spatial_Temporal_and_Spatiotemporal_Autoregressive_Probit_Models_of_Binary_Outcomes_Estimation_Interpretation_and_Presentation>.

Gujarati, D.N. 1995, *Basic Econometrics*, McGraw-Hill International Editions, New York.

Huber, P.J. 1973, ‘Robust regression: Asymptotics, conjectures and monte carlo’, *The Annals of Statistics*, vol. 1, no. 5, pp. 799-82.

Indonesia's urban studies 2014, ‘The megacity of Jakarta: problems, challenges and planning efforts’, viewed 25 June 2016, <http://indonesiaurbanstudies.blogspot.com .au/2014/03/the-megacity-of-jakarta-problems.html>.

Jha, K., Bloch, R. & Lamond, J. 2012, *Cities and Flooding: A Guide to Integrated Urban Flood Risk Management for the 21st Century*, World Bank, Washington.

Kousky, C. 2010, ‘Learning from extreme events: risk perceptions after the flood’, *Land Economics*, vol. 86, no. 3, pp. 395-422.

Kousky, C. & Walls, M. 2014, ‘Floodplain conservation as a flood mitigation strategy: examining costs and benefits’, *Ecological Economics,* vol. 104, pp.119-128.

Laframboise, N. and Loko, B. 2012, ‘Natural disasters: mitigating impact, managing risks’, *IMF Working Paper*, WP/12/245.

Lancaster, K.J. 1966, ‘A new approach to consumer theory’, *Journal of Political Economy*, vol. 74, no. 2, pp. 132–157.

Leggett, C.G. & Bockstael, N.E. 2000, ‘Evidence of the effects of water quality on residential land prices’, *Journal of Environmental Economics and Management*, vol. 39, no. 2, pp. 121–144.

McMillen, D.P. 1992, ‘Probit with spatial autocorrelation’, *Journal of Regional Science*’, vol. 32, issue 3, pp. 335–348.

McMillen, D.P. & Klier, T. 2008, ‘Clustering of auto supplier plants in the united states: generalized method of moments spatial logit for large samples’, *Journal of Business & Economic Statistics*, vol. 26, no. 4, pp. 460–471.

Murdoch, J., Sandler, T. & Vijverberg, W. 2003, ‘The participation decision versus the level of participation in an environmental treaty: a spatial probit analysis’, *Journal of Public Economics*, vol. 87, issue 2, pp. 337-362

Pinkse, J. & Slade, M.E. 1998. ‘Contracting in space: An application of spatial statistics to discrete-choice models’, *Journal of Econometrics*, Elsevier, vol. 85, pp. 125–154.

Pope, J. 2008, ‘Do seller disclosures affect property values? buyer information and the hedonic model’, *Land Economics*, vol. 84, no. 4, pp. 551-572.

Rabassa, M.J. & Zoloa, J.I. 2016. ‘Flooding risks and housing markets: a spatial hedonic analysis for La Plata City’, Environment and Development Economics, viewed 9 June 2016, < http://www.saerargentina.com.ar/trabajos/8%20-%20Zoloa%20--%20 Flooding%20risk%20La%20Plata.pdf>.

RAND, ‘The Indonesia family life survey’, viewed 25 June 2016, <http://www.rand.org/ labor/FLS/IFLS.html>.

Rosen, S. 1974, ‘Hedonic prices and implicit markets: product differentiation in pure competition’, *Journal of Political Economy*, vol. 82, no. 1, pp. 34–55.

Samarasinghe, O. & Sharp, B. 2010, ‘Flood prone risk and amenity values: a spatial hedonic analysis: a spatial hedonic analysis’, *Australian Journal of Agricultural and Resource Economics*, vol. 54, no. 4, pp. 457-475.

Skantz, T. & Strickland, T. 1987, ‘House Prices and a Flood Event- An Empirical Investigation of Market Efficiency’, *The Journal of Real Estate Research*, vol. 2, no. 2, pp. 75-83.

Stern, N. 2007, *The Economics of Climate Change: The Stern Review*, Cambridge University Press, Cambridge.

Tambun, L.T., Lumanauw, N. & Marhaenjati, B. 2015, ‘Jakarta’s flood problem totally solvable, for $9,2b’, *Jakarta Globe*, 11 February 2015.

United Nations Chronicle 2016, *Economic recovery after natural disasters*, viewed 23 December 2017, <https://unchronicle.un.org/article/economic-recovery-after-natural-disasters>.

United Nations Department of Economic and Social Affairs, Population Division 2014, *World Urbanization Prospects: The 2014 Revision*, viewed 10 June 2016, <https://esa.un.org/unpd/wup/Publications/Files/WUP2014-Highlights.pdf >.

United Nations Office for Disaster Risk Reduction 2002, *Guidelines for Reducing Flood Losses*, viewed 14 March 2016, < http://www.unisdr.org/files/558_7639.pdf>.

United Nations & World Bank 2010, *Natural Hazards, Unnatural Disasters: The Economics of Effective Prevention*, World Bank, Washington, D.C.

Wahab, R. & Tiong, R. 2017, 'Multi-variate residential flood loss estimation model for Jakarta: an approach based on a combination of statistical techniques', Natural Hazards vol. 86, pp. 779–804.

Wijayanti, P., Zhu Tiong, R., Hellegers, P., Budiyono, Y. & Van Ierland, E. C. 2017, 'Estimation of river flood damages in Jakarta, Indonesia', Natural Hazards vol. 86, pp. 1059–1079.

Wooldridge, J. 2003. *Introductory Econometrics: A Modern Approach*. 2d ed. Cincinnati: Thompson and Southwestern.

World Bank 2011, *Climate Change, Disaster Risk and the Urban Poor: Cities Building Resilience for a Changing World*, World Bank, Washington, D.C.

World Bank 2015, *Indonesia: A Roadmap for Housing Policy Reform*, World Bank, Washington D.C.

World Bank 2016, 'Feature story: Keeping Indonesia's capital safer from floods', World Bank, Jakarta, 8 January 2016 <http://www.worldbank.org/en/news/feature/2016/01/08/keeping-indonesias-capital-safer-from-floods>.

World Resources Institute 2015, 'World's 15 countries with the most people exposed to river floods', viewed 14 March 2016, <http://www.wri.org/blog/2015/03/world%E2%80%99s-15-countries-most-people-exposed-river-floods>.

Yusuf, A.A., Koundouri, P. 2005, 'Willingness to pay for water and location bias in hedonic price analysis: evidence from the Indonesian housing market', *Environment and Development Economics*, vol. 10, no. 6, pp. 821–836.

Yusuf, A. & Resosudarmo, P. 2009, ‘Does clean air matter in developing countries' megacities? a hedonic price analysis of the Jakarta housing market, Indonesia’, *Ecological Economics*, vol. 68, no. 5, pp. 1398-1407.

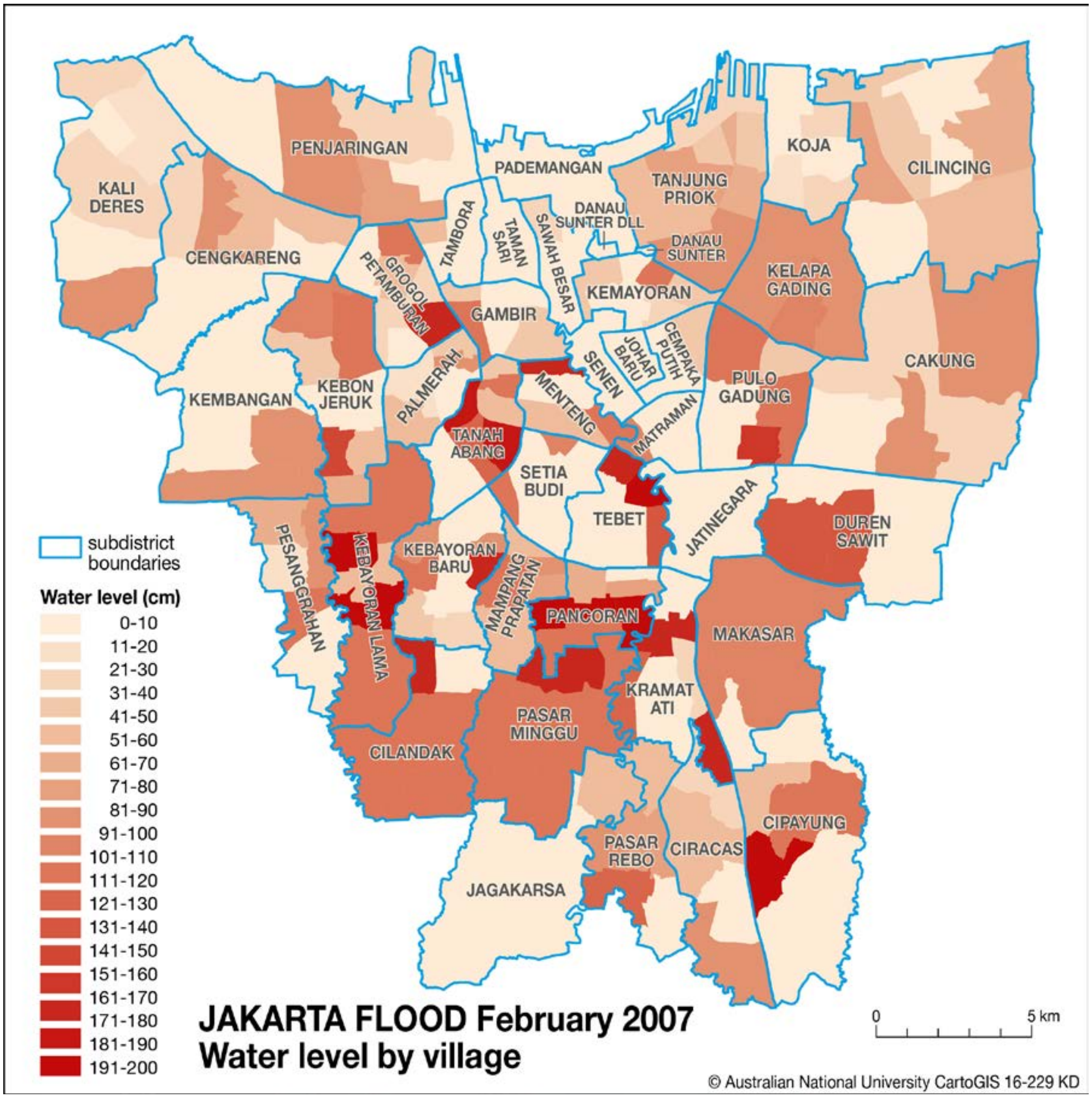

Figure 1. Map of Jakarta after the flood disaster in 2007
Source: United Nations Department of Safety and Security (UNDSS), 2007.

Table 1. Summary of existing hedonic price studies related to flood events

No	Author (publication year)	Method	Location	Results
1.	Skantz and Strickland (1987)	OLS and Event Study hedonic	TX, USA	Negative; not significant
2.	Bin and Polasky (2004)	D-D hedonic	NC, USA	-5.7%; significant
3.	Carbone et al. (2006)	D-D hedonic	FL, USA	-20% to -30%; significant
4.	Daniel et al. (2007)	OLS and Spatial hedonic	Netherlands	-7% to -13%; significant
5.	Bin et al (2008a)	Spatial hedonic	NC, USA	-11%; significant
6.	Bin et al. (2008b)	Spatial hedonic	NC, USA	-7.3%; significant
7.	Pope (2008)	Spatial FE hedonic	NC, USA	-4%; significant
8.	Samarasinghe and Sharp (2008)	Spatial hedonic	New Zealand	-6.2%; significant
9.	Kousky (2010)	D-D hedonic	MI, USA	-2.6%; significant
10.	Bin and Landry (2013)	D-D hedonic	NC, USA	-5.7% and 8.8%; significant
11.	Kousky and Walls (2013)	Simulation	MI, USA	-0.7%; not significant
12.	Rabassa et al. (2013)	OLS and Spatial hedonic	Buenos Aires, Argentina	-17.3%; significant

Table 2. Summary statistics of variables in the hedonic equation

	Mean	Std. deviation
Dependent variable		
Monthly rent (million rupiahs)	5.672	21.240
Housing characteristics		
House size (m^2)	74.524	189.086
Number of rooms	5.112	3.162
Wall material is cement/brick (1,0)	0.879	0.326
Roof material is concrete/roof tiles (1,0)	0.477	0.500
Floor material is cement/stone (1,0)	0.846	0.362
Water source inside (1,0)	0.553	0.497
House yard (1,0)	0.285	0.451
Neighborhood characteristics		
Unemployment rate at the neighb. (pct)	5.548	3.359
People w. univ. educ. the neighb. (pct)	9.310	10.970
Accessible by public transport (1,0)	0.757	0.429
Distance from district centre (km)	6.970	6.508
Traffic (hourly number of vehicles passing by)	5.590	3.209
House located along river basin (1,0)	0.330	0.470
Environmental variable		
Flood in water level (cm)	42.326	23.136

Note: Number of observations is 1,573.
Source: 2007 Indonesian Family Life Survey (IFLS) and United Nations Department of Safety and Security (UNDSS).

Table 3. Results of basic and spatial lag models

Monthly rent	Basic		Spatial lag	
Housing characteristics				
LOG(House Size)	0.9928	***	0.9973	***
	(0.0504)		(0.0492)	
Size/rooms (m2)	1.6779	***	1.6757	***
	(0.6398)		(0.5548)	
Wall is cement/brick (1,0)	0.2914	***	0.3068	***
	(0.1026)		(0.1162)	
Floor is ceramics/stone (1,0)	0.4974	***	0.4918	***
	(0.0938)		(0.1069)	
Roof is concrete/roof tiles (1,0)	-0.1274	*	-0.1094	
	(0.0740)		(0.0757)	
Water source inside (1,0)	-0.0464		-0.0149	
	(0.0785)		(0.0782)	
House yard (1,0)	0.0404		0.0270	
	(0.0908)		(0.0863)	
Neighborhood characteristics				
Unemployment rate (%)	-0.0259	**	-0.0201	
	(0.0129)		(0.0125)	
People w. univ. education (%)	0.0402	***	0.0469	***
	(0.0086)		(0.0096)	
Public transport access (1,0)	0.0432		0.0012	
	(0.0916)		(0.1021)	
LOG(Distance)	-0.2379	***	-0.2446	***
	(0.0657)		(0.0624)	
LOG(Traffic)	-0.1431		-0.1183	
	(0.1003)		(0.1061)	
House located along river basin (1,0)	-0.2094	***	-0.2664	***
	(0.0782)		(0.0795)	
Environmental variable				
LOG(Flood)	-0.1131	**	-0.1279	**
	(0.0567)		(0.1613)	
Constant	10.4006		10.7902	
Rho	n/a		-0.1901	**
Number of observations	1231		1231	
R-squared	0.4291		n/a	
Variance ratio	n/a		0.431	
Squared corr.	n/a		0.432	
Moran's I statistic	-8.120		n/a	
LM Lag	9.366	***	n/a	
RLM Lag	0.344		n/a	

Note: ***Significant at 1% level. **Significant at 5% level. *Significant at 10% level. Numbers in brackets are standard deviations.

Table 4. Number of restricted activity days and flood level

Restricted days	Basic		Spatial lag	
Housing characteristics				
LOG(House Size)	0.1884	***	0.1913	***
	(0.0430)		(0.0449)	
Size/rooms (m2)	-0.3610		-0.3298	
	(0.2992)		(0.4246)	
Wall is cement/brick (1,0)	-0.0688		-0.0754	
	(0.1047)		(0.1009)	
Floor is ceramics/stone (1,0)	-0.0937		-0.1047	
	(0.0957)		(0.0896)	
Roof is concrete/roof tiles (1,0)	-0.0465		-0.0428	
	(0.0641)		(0.0639)	
Water source inside (1,0)	0.0749		0.0735	
	(0.0639)		(0.0638)	
House yard (1,0)	-0.2072	**	-0.2118	***
	(0.0806)		(0.0771)	
Neighborhood characteristics				
Unemployment rate (%)	-0.0313	***	-0.0304	***
	(0.0104)		(0.0107)	
People w. univ. education (%)	-0.0215	***	-0.0198	***
	(0.0072)		(0.0075)	
Public transport access (1,0)	0.1350		0.1275	
	(0.0881)		(0.0874)	
LOG(Distance)	-0.0500		-0.0569	
	(0.0544)		(0.0551)	
LOG(Traffic)	0.0371		0.0060	
	(0.0849)		(0.0903)	
House located along river basin (1,0)	0.0572		0.0645	
	(0.0673)		(0.0651)	
Environmental variable				
LOG(Flood)	0.0713	*	0.1995	*
	(0.0408)		(0.0459)	
Constant	1.0690		1.4015	
Rho	n/a		-0.2067	
Number of observations	727		727	
R-squared	0.0731		n/a	
Variance ratio	n/a		0.075	
Squared corr.	n/a		0.076	
Moran’s I statistic	-2.997		n/a	
LM Lag	1.707		n/a	
RLM Lag	0.015		n/a	

Note: ***Significant at 1% level. **Significant at 5% level. *Significant at 10% level. Numbers in brackets are standard deviations.

Table 5. Depression and flood level

Depression	Probit (AME)		Spatial Probit (GMM)	
Housing Characteristics				
LOG(House Size)	0.0211		0.0663	
	(0.0201)		(0.0558)	
Size/Room (m2)	0.1899		0.5098	
	(0.2127)		(0.5844)	
Wall is cement/brick (1,0)	0.0167		0.0588	
	(0.0461)		(0.1269)	
Floor is ceramics/stone (1,0)	-0.0091		-0.0332	
	(0.0429)		(0.1181)	
Roof is concrete/roof tiles (1,0)	-0.1322	***	-0.3390	***
	(0.0300)		(0.0793)	
Water source inside (1,0)	0.0115		0.0622	
	(0.0301)		(0.0856)	
House yard(1,0)	-0.0092		-0.0369	
	(0.0338)		(0.0934)	
Neighborhood characteristics				
Unemployment rate (%)	0.0010		0.0123	
	(0.0048)		(0.0091)	
People w. univ. education (%)	-0.0045		-0.0087	
	(0.0036)		(0.0095)	
Public transport access (1,0)	0.0326		0.0556	
	(0.0395)		(0.1103)	
LOG(Distance)	-0.0303		-0.0591	
	(0.0242)		(0.0559)	
LOG(Traffic)	-0.0231		-0.1431	
	(0.0410)		(0.0905)	
House along the river	0.0140		0.0123	
	(0.0295)		(0.0806)	
Environmental variable				
LOG(Flood)	0.0475	**	0.0870	*
	(0.0221)		(0.0508)	
Constant	n/a		-0.2983	
Rho	n/a		0.8069	***
Number of observations	1,230		1,230	
Wald chi-squared(10)	30.23		n/a	
Prob > chi-squared	0.0071		n/a	
Log pseudolikelihood	-788.71		n/a	
Pseudo R-squared	0.02		n/a	

Note: ***Significant at 1% level. **Significant at 5% level. *Significant at 10% level.