ARTICLE

<https://doi.org/10.1057/s41599-020-0444-1>

OPEN

Land and building separation based on Shapley values

Ünsal Özdilek ¹✉

ABSTRACT The total value apportionment between land and building components remains an international issue both in theory and in practice. There are several concepts and methods of value separation, each leading to approximate estimations and therefore to divergent opinions about their reliability. In this paper, we present an alternative method of value apportionment based on Shapley's scheme of values, well recognized in the coalitional game theory. The practicality of this method is verified using observed prices of 14,715 residential properties sold during the year 2019 over all the 27 districts in Montreal (Canada). This unique data comes with detailed information about the essential attributes of the land and the building components. The empirical results of the method presented in this work are in line with practical expectations of total and separate values, either taken case-by-case or in aggregation per district. They are indeed encouraging when compared to the results of two other independent methods (i.e., the city evaluations and the OLS predictions) for the same properties. The results are interesting not only regarding the separation of value but also in several other related aspects. For instance, land values are often close to or even higher than the building values. This shows a phenomenon of building depreciation and land value appreciation. Some districts seem to favor the quality of the building, others being influenced by the location and quality of the land. Interestingly, in contrast to what is believed in practice, a good quality parcel of land does not necessarily have a good quality building according to the results.

¹Department of Strategy, Business School, University of Quebec, Montreal, Canada. ✉email: ozdilek.unsal@uqam.ca

Introduction

The separate estimation of the land and its building in practice is more challenging than global value estimation. In theory, there are diverging opinions on the separability/inseparability of value between the two components, raising concerns especially about the imputation of revenues to its agents of production and the impact of different systems of taxation. Classically, the inseparability thesis was assumed by Ely (1922), Russell (1945), Ratcliff (1950), Fisher (1958), and Dorau and Hinman (1969), and it is restated in some of the contemporary literature (Mills, 1998; Connellan (2004); Fischel, 2000; Kitchen, 2003; Hendriks, 2005). The consensus is that apportionment is not practicable (and is useless) because land and improvements are merged together like in an “omelet”. However, in contrast, several early (George, 1881; Smith, 1886; Marshall, 1922) and contemporary authors (Brueckner, 1986; Rice, 1982) defended separability. As these explanations are essentially theoretical, some responses are expected from those working in the field. Perhaps because of the same theoretical uncertainties, priority is usually given to global value estimation, which is often needed for taxation (Andelson, 2000).

Importance of value separation. The land and the building components (of a property) each have specific characteristics making them types of capital that are more different than alike and thus requiring separate values (Gaffney, 1994; Gihring, 1999). Separate values are relevant everywhere and concern all types of properties, especially those in residential urban areas for which the land market is scarce and classical methods are ineffective. Land and building values coexist and shape the dynamics of real estate markets in a specific way. In developed urban areas, land value progressively catches up to the value of the improvements, and sometimes goes much higher (Bourassa et al., 2011). Many practical situations require separate estimations, for example, city and capital gains taxation, tax incidences (Gloudemans, 2001), challenges by contributors in courts (Keligian, 1994), applicability of the cost method for the portion of improvements (Appraisal Institute, 2013), decisions about option value and amortization (Dye and McMillen, 2007), volatility of land and building values over time (Davis and Palumbo, 2008), management of insurances and mortgages (Guofang et al., 2003), sound land use policies and practices (Dancaescu, 2000), and preparation of land value maps (Ohno, 1985).

Contribution to the issue of price separation. The value scheme of Shapley (1953), combined with fuzzy logic, has been thoroughly examined in coalitional game theory models. It is almost absent in real estate analysis and valuation. The main contribution of this work is the application of Shapley value concept in real estate valuation as an alternative method to separate the total value of a residential property between its land and building components, following a game theory framework. We also contribute in several other ways by extending the convenient Shapley method, notably with the integration of the fuzzy logic set of values in the Gaussian function in a way that reflects the capacity of subsets of property attributes. These capacities are optimized under several logical constraints and an error minimization process.

Shapley method offers an occasion to appreciate related, important issues of value separation, such as the optimal timing and location to rebuild/change the existing structure/usage (for instance, when land values are significantly higher than the depreciated building values). The method is helpful for investors and planners to appreciate per district if the quality of land and improvements goes together; some districts may favor the quality

of the building, others being influenced by the quality of the location.

The use of the game theory basis, Shapley’s global value and fuzzy logic contribute to reviving interest on the separation issue, especially in regions where the land market is scarce. Direct attention to the value separation difficulties and the subtleties within Shapley’s special method make it more palpable in practice, in contrast to an approach that takes them for granted or leaves them to speculation. The Shapley method tangibly promotes a coherent system of separate and total value estimations for professionals working in both case-by-case and mass evaluation contexts. The following describes the conceptual bases of this method, tested with a practical application in the case of the housing market in Montreal (Canada).

Value separation approaches

Each of the three possible value apportionment theories in Table 1 supports various methods in practice. In fractional apportionment theory (FAT), the deduction of the cost portion from the total price allows one to derive an estimation of the land value. The allocation and abstraction techniques are preferred if there is a shortage of land sales in built-up areas (they establish a typical land-to-building value ratio). The extraction technique estimates land value by deducting the estimated depreciated cost of the building from the total observed price. In rent apportionment theory (RAT), which supports the income method, subdivision, land residual and ground rent techniques are also traditionally used in practice (appropriate for stable income generating properties). The most reliable technique, supported by price apportionment theory (PAT), is the direct comparison when data on land sales is available.

Comparing available approaches. The FAT techniques are approximate and essentially subjective as they proceed ad hoc almost by a rule of thumb and some simple assumptions (Mills, 1998). Techniques in RAT are also approximate, except for the land residual in the income capitalization approach (ICA) which is quite consistent in the case of income generating properties. Originally defined by Ricardo (1817) in an agricultural production context, the land residual bases were later taught with the classical urbanization models, notably by Alonso (1964), Mills (1967), and Muth (1969). The related contemporary option value concept is about the dialectics between the rent of a land parcel (empty or improved) and the type/amount of capital to be

Table 1 Various price apportionment methods.		
Apportionment theories	Methods/models	Approaches
1. Fractional apportionment theory (FAT)	1.1 Allocation	Price, cost
	1.2 Abstraction	
	1.3 Extraction	
2. Rent apportionment theory (RAT)	2.1 Subdivision	Income
	2.2 Land residual	Income
	2.3 Ground rent	
3. Price apportionment theory (PAT)	3.1 Direct comparison	Price
	3.2 OLS modeling	
	3.3 Geometric modeling	
4. Coalitional apportionment theory (CAT)	4.1 Shapley value	Price, cost ^a , income ^a
^a Not yet verified.		

optimally realized on top of it (Quigg, 1993; Yavas and Sirmans, 2005; Clapp and Salavei, 2010). The OLS approach, which refers essentially to the hedonic modeling tradition (Sirmans et al., 2005), proposes an interesting solution. However, it is limited to the independency hypothesis between the attributes, the accuracy of their isolated contributions and the thinness of the constant term. Elsewhere, the geometric approach applies to mass appraisals and its performance increases with the size of the data base (Özdilek, 2016).

In comparison to existing approaches using the basis of the utility maximization theory (Sirmans et al., 2005), the “Shapley value” method rests on the foundations of the coalitional apportionment theory (CAT), as suggested in Table 1. Besides its application in real estate in general, the development of this theory for valuation is particularly appealing. Although we focus on the practicality of the Shapley method, basic conceptual insights are provided in the next section. The empirical verification of this method in this work involves only the sales comparison approach (SCA), but we believe it has potential to conveniently address the separation of income and cost attributes with the ICA and the cost summation approach—CSA, respectively (for a description of these approaches, see Appraisal Institute, 2013).

Recent statistical approaches typically fall within either the value prediction or explanation schemes. The predictive methods aim at a reliable estimation of the total value, while the explicative methods have an additional focus to discern “marginal contribution” of attributes. The explicative methods usually assume independence among attributes, which is far from being the case in property value formation. When interaction exists, there are various solutions such as finding an appropriate functional form of the attributes, using reliable data, removing/combining specific attributes, operating a spatio-temporal a structural segmentation of the market (Fox, 1991; Wickens, 1995).

Along with the issue of interactions, research has shown there are several other questions related, for instance, to the inefficiency of the housing market (Anenberg, 2016), the information disparities among the negotiating subjective agents (Genesove and Mayer, 2001), spatio-temporal autocorrelation of prices (Basu and Thibodeau, 1998), the optimal weighting of comparables, and number of attributes (Vandell, 1991). In dealing with the reliable estimations, authors usually resort to more robust statistical models like neural networks (Riley et al., 2001), case-based reasoning (Bonissone and Cheetham, 1997), data mining (Soibelman and Gonzalez, 2002), spatial autoregressive models (Wyatt, 1997) or some hybrid approaches (McCluskey and Anand, 1999).

These modeling approaches with advanced statistics provide alternative solutions to a particular difficulty, however they are generally affected by the interaction among attributes, the complexity of techniques, the generalization of subtle valuation details, and the derivation of artificial results (Kummerow and Galfalvy, 2002; Epley, 1997). In comparison to these models, Shapley method as well can suffer from the imperfections of the market and the quality of information. However, it is flexible to take into account these issues, in particular the interactions among the attributes (Su et al., 2019). It also benefits from the capacity of fuzzy measures with a Gaussian function, under several logical constraints and optimization, tested with detailed information from the market.

Shapley in game theory and real estate. The issue of value separation is not only relevant in real estate, but also elsewhere, especially in fields related to value creation, portfolio optimization, compensation, or cost allocation (Moriarty, 1975;

Goetzmann and Rouwenhorst, 2005; Greyserman et al., 2006). In these areas, there are promising value separation analyses with reference to the basis of game theory (Neumann and Morgenstern, 1944; Nash, 1951). In the setting of a cooperative game, a fair allocation of value to entities jointly involved in its creation is examined. The Shapley value is a solution in this framework of game theory formally presented by Shapley (1953). Mainly welcomed in economic, financial, and socio-political fields, the Shapley method estimates separate values comparable to a payoff (or bargaining power) for an agent involved in the cooperative game of “value creation”. It is well adapted to the problems of resource sharing/partition with separate portions of value creating attributes, sometimes designated as *Shapley value*, *influencing factor* or *power index* (Young, 1988). This scheme of value developed by Shapley thus offers a convenient and practical basis for exploration in property price separation.

The application of the Shapley method is not direct and may involve several preparation steps in the context of real estate analysis and valuation. The method in fact requires the importance of attributes considered individually and in combination to be delineated. Originally developed by Zadeh (1965), and generally applied in control systems/decision making, fuzzy set theory offers a flexible way to take into account the importance, logical constraints, and vagueness of interactions between the attributes (Zimmermann, 2001). The integration of fuzzy measures over integrals in the Shapley method is extended by Sugeno (1974) and others (Atanasov, 1986; Grabisch, 1996) in dealing with specific interactions among attributes using non-additive measures (also known as fuzzy integrals or capacity function). The fuzzy measures of relations can refer to the experience (expert opinion) or any other particular method, such as the optimization and the behavior learning in databases (Grabisch, 1997).

The specialists in real estate analysis in general and to a much lesser degree in valuation (total and separate values) are, for the most part, unaware of the Shapley method. There have however been several attempts to use the fuzzy logic process in real estate management and investment risks (Sun et al., 2003; Barranco et al., 2004). Dilmore (1993) presented an interesting paper at a conference about the relevance of fuzzy logic in real estate as agents inaccurately handle and acknowledge utility levels from attributes. Bagnoli and Smith (1998) also restated similar observations about the relevance of fuzzy logic in price formation dynamics. The Shapley method presented in this work thus proposes a new perspective in real estate evaluation, involving at least two important challenges described in the following sections: conceptualizing attributes as players participating in a game and empirically verifying its consistency in practice.

Shapley separate values

The following is a conceptual description of the Shapley scheme of value in price comparison method; similar steps of separation can also be followed in cost and income approaches, respectively, using the data on building cost and stabilized future incomes. We first present Shapley’s original method, integrating the fuzzy set of measures. Building these fuzzy measures varies depending on the context and the type of subject, but a convenient process is introduced and explained in the property valuation context. Following the original steps of the Shapley method allows us to globally separate the values of the land and the building.

Considering value attributes as players. In the application of the Shapley method, we primarily need to explain on why the property parameters can be seen as players in a game of negotiation. We believe that the basis of game theory welcomes such

an assumption. There are ample examples, as mentioned above, with other types of capital and contexts in value creation (i.e., portfolio evaluation or cost allocation). It might seem strange to assume, for instance, that the “living area” attribute is an agent acting in the negotiation game of value creation (or price formation). The agent is not the attribute, but instead the deciders, such as the buyers/sellers, who feel and simulate the exercise of negotiation themselves as representatives in the market.

Having a living area more or less bigger than, for instance, 100 square meters expresses the power of that attribute in the same market of different size properties. We can assume this variation from the *same attribute* as its power of “vertical influence” on the realized prices (with changing living areas of buildings). When introducing the “bargaining power” of another attribute in the game, like the “presence of a garage” in the property, the economic agents will do the same exercise of comparing the different sizes, at the same time being under the “horizontal influence” of comparing *different types of attributes* on the price. Should they include the garage in their basket of utilities rather than other types of possible attributes? The vertical and horizontal types of a power relationship modulate the hedonic decisions in holding specific attributes and dropping others in the value creation game. Considering the property attributes as players makes even more sense when realizing the “invisible hand” of the market that objectively retains value determining attributes and strongly clearing those individually subjective during the negotiation.

Moreover, depending on the type of property, there are plenty of attributes no longer attractive in the game of negotiation, like the old-style external brick facing which is being progressively replaced by vitrified facades (usually in the case of commercial properties) or the outmoded laundry chutes in old buildings which were popular in the past. Some other parameters might reappear or be newly introduced in the market like sustainable building characteristics in comparison to those in conventional ones.

On a special note, familiarity with building/land parameters (about their potential values) leads to an informational advantage for economic agents during the game of negotiation. The market is less aware on the amount and the type of impacts (positive/negative) from certain attributes, for instance, of electric wind turbines newly built in the district, in comparison to the size of the living area attribute. In fact, *familiarity* with price differential comparisons reflects better the economic impacts (the internalization of benefits) of living area as this attribute is usually negotiated in the previous sales that informs and re-informs new buyers/sellers. However, there is less information on price differential regarding wind turbines, newly introduced on the market (with the disadvantage of having few comparables, if any).

Global separate values

The utility space U of properties is a subspace of the standard space $\mathcal{R}^N$, for a finite or infinite $N > 0$. As a subspace, U has a number of observable utility attributes u_i . We suppose that i belong to a finite set $\{1, \dots, M\}$, and that u_i s are those characteristics of properties like their size, number of rooms, the presence of a swimming pool, the proximity to urban parks, etc. The utility space U has to be appropriate with a sufficient number M of observable functions u_i , allowing one to distinguish prices P_i, P_j for different properties i, j . Property price P is a nonnegative function $P: U \rightarrow \mathbb{R}^+$ and it is a Euclidean measure on utility space U , so that (U, P) is a measure of that space in an appropriate sense. $P(A)$ expresses the total price of properties, identified with all points of the set $A \subset U$.

Considering a set of finite attributes, let coordinates u_i of the utility space U be split into two groups: attributes related to the building (we will call them x_i s) and attributes related to the land (respectively y_j s) so that μ is a capacity function which corresponds to a degree of importance for each of the building and land attributes. It follows that $\mu(x_i) = u_{x_i}$ and $\mu(y_j) = u_{y_j}$, so $\mu(x_i, y_j) = u_{ij}$. If the μ function is additive, then $u_{ij} = u_i + u_j$ (this is a case of linearly weighted aggregate function).

Let also L_y be the sum of all the normalized utility parameters of land (with a weight of importance ω_1) and B_x the sum of all the normalized utility parameters of the building (with a weight of importance ω_2). The total price P of a property will be the sum of all the weighted utilities belonging to its land and building components, with respectively $X = \{x_1, x_2, \dots, x_m\}$ and $Y = \{y_1, y_2, \dots, y_n\}$.

$$P(X, Y) = \sum_{i=1}^m \omega_{x_i} x_i + \sum_{j=1}^n \omega_{y_j} y_j \quad (1A)$$

where ω_{x_i} is the importance weights of x_i and ω_{y_j} the importance weights of y_j . This specification expresses the estimation of the total price based on the weighted linearity in the respective importance of attributes. In aggregate and for more clarity, we can approximate $p(X, Y)$ as

$$P(X, Y) = \omega_1 L(y) + \omega_2 B(x) \quad (1B)$$

where ω_1 is the average weight of the land, and ω_2 the average weight of the building.

This additive and weighted form of land and building utility aggregation provides an estimation of the total price. The reality of price formation is more complex mainly because of possible interactions among land and building attributes. The economic agents express their multi-criteria utility decisions based on the availability of some attributes taken individually, but also in combination. Elsewhere, adding an attribute of a property in the model does not necessarily change the price. Taken in isolation, it might provide a nil or weak utility, but it sometimes might turn out to be very significant when combined with another attribute. For example, the attribute “altitude” of a piece of land that is far from a river (another spatial attribute) would probably not affect the price, but it will significantly change it when closer to the river (mostly for a scenic view).

In general, the price explanation and prediction models do not consider interactions caused by the combination of attributes. Considered as a problem to solve, interactions are either kept constant or an attempt is made to exclude them from the models. But, attributes are naturally in interaction and coalition in forming the total price (or the worth of the “game”) for properties:

$$P(L, B) = \omega_1 L + \omega_2 B + \mu(LB) \quad (2)$$

Before the application of the Shapley method, the analyst needs to specify a set of attributes, with their values, such as the living area, the relevant weights and constraints, if there are any, based on experience or in any other ideally objective ways. The normalized and weighted values of property attributes in a way reflect their capacity in determining a property’s total price. These capacities of attributes, considered individually and in combination, are non-additive fuzzy measures, i.e., with probabilities varying between 0 and 1. These measures can be assumed to increase randomly or in following a given type of function.

There is in fact no standard method for building fuzzy indicators, which depend on the context and the type of phenomenon in question. The initial fuzzy measures (with abstract relations) become “defuzzified” providing the analyst with more flexibility to specify some conditions, with others being less obvious, but

logically and generally following the predictable behavior of attributes and their combinations. Some parts, such as the importance of certain attributes, may benefit from expert opinion or initial results based on the use of hedonic models. The minimization of the squared error criterion can be used in identifying more probable fuzzy measures (Torra, 1999).

Shapley's basic method. In specifying Shapley's method, consider $E = \{x_1, x_2, x_3, \dots, x_N\}$ and $\mathcal{P}(E)$ being the set of subsets of E . This is represented by $\mathcal{P}(E) = \{X\}$, with $X \subset E$ and $\text{Card}[\mathcal{P}(E)] = 2^N$. The normalized and weighted original values of attributes $X \in \mathcal{P}(E)$ are obtained from the sample. Let μ be a partition capacity function and a monotone non-additive fuzzy measure on $\mu(X)$. Initially, we assume that $\mu(\emptyset) = 0$ and $\mu(E) = 1$ and, where E is the global set of $\mathcal{P}(E)$. The non-additive capacities between these intervals are generated, in the present research, by a Gaussian function in conformity to the capacity criteria such that $\forall X$ and $T \in \mathcal{P}(E)$, if $X \subset T$ implies $\mu(X) \leq \mu(T)$.

The value of Shapley (1953) is a concept of solution verifying certain properties (axioms). The concept of the solution is adapted to the problems of sharing resources (or cost allocation) with the characteristic function $\mu: \mathcal{P}(E) \xrightarrow{\mu} R^+$, with $X \in \mathcal{P}(E) \rightarrow \mu(X) \in R^+$, $\emptyset \in \mathcal{P}(E) \rightarrow \mu(\emptyset) = 0$. We are looking for a solution: $\phi(\mu) = (\phi_i(\mu))_{i \in N}$, $\phi_i(\mu)$ being a capacity index of player i (i.e., the value of the game for player i), and satisfying the following Shapley axioms (Winter, 2002; Hu and Yang, 2010):

- Efficiency:** $\sum_{i=1}^N \phi_i(\mu) = \mu(E)$, with $\mu(E) = 1$.
- Symmetry:** If i and j are symmetric (substitutes), i.e., $\mu(X \cup \{i\}) = \mu(X \cup \{j\}) \forall i, j \notin X$, then $\phi_i(\mu) = \phi_j(\mu)$.
- Nullity** (null player): If i is zero, i.e., $\mu(X \cup \{i\}) = \mu(X) \forall i \notin X$, then $\phi_i(\mu) = 0$.
- Additivity:** If μ_1 and μ_2 are two coalitional games, then $(\mu_1 + \mu_2)(X) = \mu_1(X) + \mu_2(X)$, then $\phi(\mu_1 + \mu_2) = \phi(\mu_1) + \phi(\mu_2)$.

The separation (sharing) of total value among property attributes is realized under these conditions (axioms). The *efficiency* condition indicates that the total value available to the property attributes (players) in the grand coalition is distributed among them, following their relative capacity. In the *symmetry* condition, equivalent attributes will have equal contributions and an attribute with zero contribution to all coalitions receive zero payoff according to the *nullity* condition. *Additivity* condition simply settles that values of two games (markets) sum up to the value computed for the sum of both games.

We use the following alternative formula to estimate Shapley value, which satisfies these axioms (Osborne and Rubinstein, 1994; Branzei et al., 2008):

$$\varphi_i(\mu) = \frac{1}{|\Theta|!} \sum_{\theta \in \Theta} [\mu(P_i^\theta \cup \{i\}) - \mu(P_i^\theta)] \quad (3)$$

where $\varphi_i(\mu)$ is a weighted sum of the marginal contributions of the attribute i , Θ is the set of all orders of the set of players, and P_i^θ is the set of players which precede i in the order θ . This expression captures the separate value of each property attribute over all the different sequences in the grand coalition. Based on these separate estimations (individually informative to compare an attribute's marginal production cost to its selling price), it becomes possible to estimate property's total and separate values for land and building components.

The Shapley separate value is a global contribution (capacity) of an attribute considered alone and in interaction within the same game of total value creation. The interaction indices between each pair of attributes are sufficient in multi-criteria decision-making (MCDM) problems even though it is possible to define them further (Magoc et al., 2011). Several authors tried

Shapley's index extensions of degree two which takes into account the interactions between pairs of attributes (two attributes considered simultaneously) and all other attributes in the set of combinations (Grabisch, 1997; Roth, 1988). In this work, for the demonstration, Shapley's basic method suffices to obtain marginal values of attributes that are needed in total price apportionment.

Gaussian fuzzy measures

MCDM problem is extensively investigated for the evaluation and the choice of alternatives, offering multiple attributes (Wallenius et al., 2008). Agents take decisions under uncertainty and cognitive bias as they vary in their motivations, use of information and capacity of evaluation (Kahneman et al., 1982). Uncertainties in alternatives involve imprecise decisions in various contexts that can be reliably studied based on fuzzy measures (Zadeh, 1965). To date, various fuzzy measures are proposed in control systems, cooperative game theory, and combinatorial optimization, varying in relation to factors like the application domain, the availability of data, the mathematics involved and particular needs (Atanassov, 1986). When attributes explaining a phenomenon of interest are nonadditive and in interaction, fuzzy measures can address them adequately (Ishii and Sugeno, 1985). Sugeno (1974) initially proposed a nonadditive monotone fuzzy measure that can take into account the interactions between attributes (Weber, 1988). In that perspective, the Shapley index has been proposed to redistribute (or separate) fuzzy measures (Murofushi and Soneda, 1993).

Expert opinions or hedonic measures on the contribution of parameters might help in building the weights of fuzzy measures containing all the possible combinations of criteria 2^n in the whole set $\mathcal{P}(E)$, but they are essentially biased (Grabisch, 1997). The main reason is the imperfect price levels given the subjective decisions of economic agents, but there is a coherence when related to the quality/quantity of attributes, allowing to derive some objective measures. In our case, we use weighted and normalized correlations of attributes with observed prices per district in Montreal.

These correlations are used as initial values in a Gaussian function that generates capacities of attributes under logical conditions, Shapley's axioms and Sugeno's criteria of nonlinear monotonic increase of contributions (Grabisch and Raufaste, 2008). The choice of that function rests on several arguments; notably, it is known that real estate price formation usually obeys a Gaussian law of behavior (Wyman et al., 2011), responding adequately to the criteria of Sugeno (1974) in building the fuzzy measures of capacities. The total capacity in the Gauss integral is assumed to cover well the functional behavior of decisions on the probable value of combinations, contributing to a more decrease in bias factor. The general Gaussian function is the following:

$$P(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\bar{x}}{\sigma}\right)^2} \quad (4A)$$

where $P(x)$ is the probability of the Gaussian function, $\bar{x}$ the average value, and σ the standard deviation. Without getting into the details of the proofs, we can easily show that the probability is equal to unit 1.

The fuzzy measure can be a mathematical function of probability, for instance, as a Gaussian function, under the conditions to be satisfied such as $\mu(\emptyset) = 0$, $\mu(E) = 1$ and $X \subset T$ if $\mu(X) \leq \mu(T)$. The condition $\mu(E) = 1$ is not automatically reached in the field of probability and the highest combined value "E" will instead have a maximum value of capacity in the Gaussian function that we use. Mathematically speaking, the maximum

weight is to be given to “ E ” in such a way that $\mu(E) = \text{Max}$ in the Gaussian function.

For the fuzzy measures of capacities, we prefer to use an adjusted version of the Gaussian function. In the equation above, $\bar{x}$ is the average value of the sample and it is a characteristic which changes from sample to sample and determines the location of the Gaussian function on the x -axis. If $\bar{x} = 0$, then the expression above becomes:

$$P = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} \quad (4B)$$

The value of the standard deviation σ flattens the Gaussian curve, in a more clearly distinguishable way, and independently from the small size of the samples. In other terms, $\bar{x}$ is only a translation of the $\bar{x}$ values in the capacity measures while σ directly influences the results of $\mu(X)$ containing all the combinations $\mathcal{P}(E)$. Ultimately, what is important here is keeping the Gaussian form, which reflects the image of the general tendency of all the samples and not one particular sample. By considering $\bar{x} = 0$ in the equation, we allow that the system is not depending on diverging behaviors from the small size of the samples. The behavior should rather depend on the system. It is the main reason why we kept σ from the sample and omitted $\bar{x}$ from the expression above.

The treatment of the rough data and its normalization to build the probabilistic values of x (small values) allow them to be kept between $0 < x < 1$. The probabilistic P values in the Gaussian function will, however, be concentrated around a maximum value, very close to each other with tiny differences that are difficult to detect using the original Gaussian function. The sum of probabilities $\sum_{i=0}^{\infty} P_i$ continuously increases and disrupts unnecessarily the capacities in the fuzzy table. To remedy that situation, as a first reason, we operate a change of variable x by $X = \frac{1}{x}$. The second reason for that modification is to keep the monotonic increase in the capacity of combinations. The normalization of the rough data provides individual values of x and the computation of values for their combinations $\mathcal{P}(E)$ is additively and linearly operated. Regarding this choice, one may then ask why we do not take their product or other types of combinations. The following mathematical demonstration responds to that question.

The sum of two criteria x_1 and x_2 has a probability higher than $P_{x_1} \cdot P_{x_2}$. For the demonstration of this theorem, consider the simplest function of a Gaussian distribution $P = e^{-x^2}$, with $P_{x_1} = e^{-x_1^2}$, $P_{x_2} = e^{-x_2^2}$, and $P(x_1 + x_2) = e^{-(x_1 + x_2)^2}$:

$$P(x_1 + x_2) = e^{-(x_1 + x_2)^2} = e^{-[x_1^2 + x_2^2 + 2x_1x_2]} = P_{x_1} \cdot P_{x_2} \cdot e^{-2x_1x_2} \quad (5)$$

we deduce that $P(x_1 + x_2) \geq P_{x_1} \cdot P_{x_2}$.

Error minimization process. In the following practical example, we considered the most important parameters of land and building in Shapley’s method that approaches a realistic explanation of price differentials. However, individuals taking part in the market subjectively express their utility in deciding on the number, type and importance of these parameters, therefore requiring an objective support of results. A consistent way in the field is to compare separate values with those of other approaches such as city and OLS evaluations. The performance of comparable models can rely on a common behavior of the market, i.e. a normal distribution of separate prices (and errors). A direct basis of comparison is clearly through the recently sold vacant lands within the same usage and market. Such a data is rare (and sales are usually biased) in almost entirely built-up cities.

In our approach, we compare separate results of Shapley method with those of OLS and city evaluations. Also, we use a

statistical error minimization test, integrated within the entire process of Shapley’s separate value computation. This test is theoretically and practically appropriate as the weights of attributes are objectively estimated using a key assumption in valuation. The adjusted prices of the comparables (here three) and the observed price (or the estimated market value) of the subject have to converge at the end of the process. In fact, following the process of price adjustments, properties become very similar. Weights of attributes are iteratively changed under the constraint of error minimization that converges adjusted prices of similar properties.

Observed prices can be adjusted using either a ratio or marginal contribution of each attribute (provided by an expert or derived from, for instance, an OLS model). To simplify the computations, we use a ratio derived from a Euclidean norm representing a degree of proximity (similarity/dissimilarity) of each comparable to the subject property. Considered in an n -dimensional Euclidean space, the Euclidean norm of a vector is expressed by

$$|y| = \sqrt{y_1^2 + y_2^2 + \dots + y_n^2} \quad (6)$$

Although there are various types of distance estimation, this expression measures the size of a vector and reduces to an absolute value for $n = 1$. It assigns to each attribute vector the length (or magnitude) from the origin to point y , a consequence of the Pythagorean theorem. This weighable distance helps to measure the degree of similarity between comparables and the subject property. For each comparable and subject property j , we form the weighable distance $|y_j| = \sqrt{\omega_1^2 x_1^2 + \omega_2^2 x_2^2 + \dots + \omega_n^2 x_n^2}$, where it has assumed that $x_i \omega_j \equiv 0$, $i \neq j$. We normalize the weighable distance for each property with the summation of all weighable distances; we call this parameter, normalized weighable distance $||y_j||$. The adjusted price of property j is calculated by multiplying the observed property j price P_{oj} and the subject normalized weighable distance $||y_s||$, divided by the normalized weighable distance of property j , $||y_j||$. Therefore, the adjusted price of property j can be represented by $P_{Adj} = \frac{P_{oj} ||y_s||}{||y_j||}$.

In deriving the weights of the attributes, we use an optimization process that aims at minimizing the distances between the adjusted and the observed prices. The following root-mean-square error (RMSE) formula computes the standard distance between these prices:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (\Delta P_{Adj})^2}{n^2}} \quad (7)$$

where ΔP_{Adj} represents all the relative differences between P_{oj} and the other three adjusted prices P_{Adj} $j = 1, 2, 3$, with $n = 6$. Once an optimal weight of attributes is found, Eq. (7) should converge to zero, in theory, as the normalized attributes of the three comparables and the subject are the same.

A practical example

The practicality of the Shapley method is verified here based on the observed prices of 14,715 single-family properties sold during the year 2007 over all the 27 districts in Montreal. This data is convenient and already operational as the technical details of gathering, validation, codification, and computation have been realized by the city. It already contains two comparable bases of separate values obtained from the OLS method and the city evaluations for the same properties.

Data specification. In this quite heterogeneous market, Table 2 shows that the total prices range between \$51,000 and \$617,829,

Table 2 Definitions and statistics of attributes.

Codes	Definition	Average	Min.	Max.	Std. dev.
Price	Selling price of properties in 2007 (in can. \$)	151,272	51,000	617,829	63,153
L-city	City evaluation of the land	50,033	10,200	430,000	30,246
B-city	City evaluation of the building	100,254	30,000	434,900	46,181
Total-city	City evaluation of the property	150,286	43,800	662,500	65,714
L-OLS	OLS evaluation of the land	48,328	-21,862	408,335	38,181
B-OLS	OLS evaluation of the building	95,684	11,296	385,592	39,760
Total-OLS	OLS evaluation of the property	144,012	32,624	675,828	60,759
<i>Building attributes</i>					
B1	Living area in sq. feet	0.384	0.098	1.000	0.113
B2	Structural quality of the building	0.569	0.000	1.000	0.167
<i>Land attributes</i>					
L1	Land area in sq. feet	0.413	0.105	1.000	0.008
L2	Location quality of the land	0.422	0.000	1.000	0.108

with an average of \$151,272. The database comes with detailed information about four essential attributes of the properties, distinguished between their land and building components. The X–Y geographical coordinates of each of the properties enabled us to locate them precisely on the map and appreciate the spatial distribution of separate estimations.

Among the four important attributes, the structural quality of the buildings has been qualified in different classes by the city appraisers who have visited each property in the past. We build an attribute on the quality of the land with the aid of the geographical information systems (GIS), using the multiple thematic maps that show the distances of each property to important location points by district (city parks, schools, subway stations, river, commercial centers, etc.).

It is possible to consider a higher number of attributes related to the land, building and financial aspects, assuming that the time effect is kept constant (adjusted prices or sales within the same year). For a residential building, the model might include, for instance, the type of exterior covering, the floors, the size of the garage, the presence of a fireplace or a swimming pool. These attributes might provide more explanations for price variations. In this practical example, we expect that each pair of size and quality attributes about the land and building will address most of the explanations in price variations. In practice, professionals often manually operate a classical grid method in which the comparable properties found around the subject usually require an average of 3–4 adjustments. The price method works better when comparables are very similar, thus requiring fewer adjustments. The Shapley method can also work with more attributes, but this needs an automatized algorithm of computation due to the higher number of possible combinations in the fuzzy table.

The Shapley method we use here will directly separate the observed total price of the subject property (randomly taken from the data) between its land and building. A global estimate of the market value for the same subject can also be divided between the two components, assuming that the evaluation correctly integrated interactions between the attributes. The Shapley method has also the potential to predict the separate values of the subject, based on the information from its neighbors. Using the GIS tool, we were able to identify the three comparables in the data that are the closest to the subject.

Fuzzy measures. Before proceeding to direct separation and prediction of the separate prices, we can build the relative values (or bargaining powers or “X values”) for the four attributes in the fuzzy table. There are several options to consider in building these values, with an elaborated discussion around the possible

conditions or specific constraints for each attribute, taken in isolation and in combination. For example, it is easy to guess the minimum and maximum values of attributes for an empty set (property without any attributes) and a full set (all the attributes, with their combinations). A land attribute can participate in the game, without any quality or building (polluted and vacant land). In this case, a building’s negotiation power of the two attributes is absent. Other different interpretations and situations might exist, but the challenge is between the extreme values, especially with the combined bargaining power.

Results

Thus far, we have prepared the data and the cumulative capacity of $\mu(X)$ values in Table 3, needed for the application of the Shapley method. Considering that all the lands are built (improved) in the data, normalized and weighted correlations of land and building attributes are used as initial values. The Gaussian cumulative distribution of capacities is computed under the scheme of Shapley’s axioms and Sugeno’s criteria of non-linear monotonic increase of contributions. Instead of considering the expert’s subjective opinion on these contributions, the Gaussian function works with the correlations between property attributes and observed prices.

Shapley’s separate values are computed based on a key assumption we made. Accordingly, when the observed prices of comparables are correctly adjusted by their attributes of various importance, in comparison to those of the subject, they are theoretically supposed to be the same or, at least, very close. To objectively conduct this process, we allowed weights of attributes to change under the optimization process of price convergence that automatically connects all the steps of Shapley’s separate value estimations. When these weights change, prices of comparables are adjusted with error minimization. Weighting optimization stops when the total distance between the adjusted prices is minimized.

As shown in Table 4, the optimal weights of L1, L2, B1, and B2 attributes are, respectively, 0.16, 0.24, 0.40, and 0.20. Based on these weights, Shapley method computed an indication of capacity for each attribute of the subject property (S) in the example, allowing to estimate separate values. Accordingly, for the subject property, the estimated separate value for the building is \$259,304 (57%) and \$195,616 (43%) for the land.

For the test, we used correlations without optimization which resulted in a RMSE of 0.039, in comparison to a RMSE of 0.011 for the optimized weights. The adjusted prices with optimization are very close to the observed price of the subject property and, consequently, suggest closer separate values for land and building components. The RMSE optimization process allowed to tune

Table 3 Fuzzy measures derived from the Gauss function.

Set of combinations	Partition function $\mu(X)$	$\mu1(X)$ normalized	$\mu1(X)$ from Gauss	$X = (1/X1)$	X1 original values
{ $\emptyset$ }	0.0000	0.0000	0.0000		
L1	0.0007	0.0007	0.0011	5.20	0.192
L2	0.0007	0.0000	0.0001	6.34	0.158
B1	0.0467	0.0460	0.0772	2.43	0.411
B2	0.0512	0.0045	0.0076	4.18	0.239
L1L2	0.0806	0.0293	0.0493	2.86	0.350
L1B1	0.1673	0.0867	0.1457	1.66	0.603
L1B2	0.2185	0.0513	0.0861	2.32	0.431
L2B1	0.2995	0.0809	0.1359	1.76	0.569
L2B2	0.3416	0.0421	0.0708	2.52	0.397
B1B2	0.4352	0.0936	0.1572	1.54	0.650
L1L2B1	0.5415	0.1063	0.1786	1.31	0.761
L1L2B2	0.6259	0.0844	0.1418	1.70	0.589
L2B1B2	0.7365	0.1106	0.1857	1.24	0.808
B1B2L1	0.8498	0.1133	0.1903	1.19	0.842
L1L2B1B2	1.0000	0.1502	0.2523		
Sum		1.0000	1.6797		

Table 4 Shapley's separate values.

Cases	B1	B2	L1	L2			
A	1428	3048	3234	3959			
B	2140	46,738	6097	3887			
C	2469	36,590	6540	3931			
S	3320	60,686	11,265	3455			
Sum:	9357	147,063	27,136	15,232			
Weights-corrélacion: B1 = 0.76; B2 = 0.27; L1 = 0.21; L2 = 0.53					Module	Norm.	Subject/comparable
Weights - optimization: B1 = 0.40; B2 = 0.20; L1 = 0.16; L2 = 0.24							
A	0.061	0.004	0.019	0.062	0.089	0.164	2.072
B	0.091	0.064	0.036	0.061	0.132	0.243	1.402
C	0.106	0.050	0.039	0.062	0.138	0.253	1.346
S	0.142	0.083	0.066	0.054	0.185	0.340	4.821
S (norm.)	0.411	0.239	0.192	0.158	0.544	1.000	
Shapley's separate values (with optimization)					Prices	Building	Land
A	0.46	0.13	0.25	0.16	446,151	263,229	182,922
B	0.34	0.15	0.28	0.23	434,659	212,983	221,676
C	0.39	0.20	0.24	0.17	440,043	259,626	180,418
S	0.38	0.19	0.27	0.16	454,920	259,304	195,616
Observed prices	Adjusted prices with correlation	Adjusted prices with optimization	Relative differences	Results	RMSE - optimization	RMSE - correlation	
A 215,342	388,903	446,151	(A-B)/A	0.026	0.0007	0.007	
B 309,920	421,397	434,659	(A-C)/A	0.014	0.0002	0.005	
C 326,842	416,942	440,043	(A-S)/A	-0.020	0.0004	0.029	
S 454,920	454,920	454,920	(B-C)/B	-0.012	0.0002	0.000	
			(B-S)/B	-0.047	0.0022	0.006	
			(C-S)/C	-0.034	0.0011	0.008	
					0.011	0.039	

Shapley's separate values based on the essential argument of price convergence. Other statistical approaches like the OLS derive marginal contributions of attributes based on price differentials and use these contributions for price adjustments in later steps from which an opinion on the market value of land and building is reached. In Shapley's method, integrating the RMSE process, any slight adjustment in the weight of a given attribute simultaneously changes all the interconnected measures in the algorithm of computation.

Do these separate values make sense or are they in line with the practical expectations? Even though they are in line with practice, direct response cannot be provided with this single example; we

thus need to consider a higher number of cases as in the data per district. Using the same algorithm, we computed separate estimations for all the other 14,714 properties in the data, considering each time their own three comparables. Also, a comparison with the separate results of the same properties in the data obtained by the two other independent methods (i.e., city and OLS) are more enlightening.

Table 5 displays the average results of the method per district, sorted in increasing order of the observed prices. The results are interesting not only regarding the separation but also in several other related aspects. For instance, the land values are close to or even higher than the building values in certain districts; the

Table 5 Average results of separate estimations per district.										
Districts	No. of sales	Observed prices	Land value	Building value	Land ratio	Size land (L1)	Quality land (L2)	Size Building (B1)	Quality Building (B2)	
1. Villaray-StMichel-Ext.	208	107,829	37,871	69,958	0.351	0.316	0.035	0.408	0.241	
2. Sud-Ouest	233	113,241	30,348	82,893	0.268	0.165	0.103	0.284	0.448	
3. RivPrair-PAT-MtlEst	1708	113,441	31,939	81,503	0.282	0.231	0.051	0.375	0.344	
4. Montréal-Nord	292	117,299	39,642	77,657	0.338	0.332	0.006	0.329	0.333	
5.Pierrefonds	1787	120,385	25,094	95,291	0.208	0.153	0.055	0.461	0.331	
6. Mercier-Hochel.-Mais.	441	120,830	36,577	84,253	0.303	0.290	0.013	0.359	0.338	
7. Dorval	418	128,742	29,002	99,740	0.225	0.149	0.076	0.456	0.318	
8. Pointe-Claire	1190	132,926	43,542	89,384	0.328	0.186	0.142	0.456	0.257	
9. Verdun	177	133,794	66,399	67,395	0.496	0.232	0.264	0.291	0.213	
10. Lasalle	294	134,479	29,174	105,305	0.217	0.137	0.080	0.358	0.425	
11. Rosement-P.Patrie	176	136,840	46,014	90,826	0.336	0.335	0.002	0.335	0.328	
12. Anjou	286	137,346	47,572	89,774	0.346	0.327	0.019	0.346	0.308	
13. Doldorm-Roxbor	1618	142,677	30,653	112,025	0.215	0.116	0.099	0.457	0.328	
14. Ilebiz-Stanne-Stgeun	676	146,982	52,031	94,952	0.354	0.189	0.165	0.408	0.238	
15. Ahuntsic-Cartierville	668	156,773	52,452	104,320	0.335	0.220	0.114	0.421	0.245	
16. St-Laurent	525	158,311	31,605	126,706	0.200	0.142	0.058	0.432	0.368	
17. Plateau-Mont-Royal	120	167,253	56,235	111,017	0.336	0.328	0.008	0.367	0.296	
18. St-Léonard	240	168,701	54,697	114,004	0.324	0.323	0.001	0.341	0.335	
19. Kirkland	843	176,188	56,146	120,042	0.319	0.296	0.023	0.393	0.288	
20. Beacnflid-BaiUrf	1004	183,642	32,627	151,015	0.178	0.066	0.112	0.454	0.369	
21. CDN-NDG	697	189,622	72,066	117,557	0.380	0.121	0.259	0.458	0.162	
22. StLuc-Hampstd	431	214,879	85,540	129,339	0.398	0.253	0.145	0.386	0.216	
23. Ville-Marie	71	239,362	76,533	162,829	0.320	0.203	0.117	0.415	0.266	
24. Ile-Des-Sœurs	151	245,103	73,963	171,140	0.302	0.199	0.103	0.442	0.256	
25. Mont-Royal	236	314,294	98,941	215,353	0.315	0.275	0.040	0.365	0.320	
26. Westmount	135	343,691	149,542	194,149	0.435	0.244	0.191	0.324	0.241	
27. Outremont	90	347,600	116,727	230,874	0.336	0.332	0.004	0.337	0.327	
Average	545	173,786	55,664	0.297	0.313	0.228	0.085	0.386	0.301	
Min.	71	107,829	25,094	0.105	0.178	0.066	0.001	0.284	0.162	
Max.	1787	347,600	149,542	0.644	0.496	0.335	0.264	0.461	0.448	

inverse is usually the case. This shows a phenomenon of building depreciation and land value appreciation in the districts close to downtown areas such as Verdun and Westmount. Interestingly, those two districts are on opposite sides of the market. In Westmount, like in Verdun, the buildings are mostly old and this probably reflects the reticence of the market due to the exponential increase of residential property values (and taxes) on lands that have a huge potential in rent. When the districts are far apart, the land values usually decrease, with a better quality of buildings.

The distribution of the bargaining power of the four attributes demands particular attention per district. We can observe where clearly the size and the quality of the land and building dominate. These results imply that we cannot consider and interpret a uniform distribution of the so-called “hedonic contribution” on the utility of each attribute. It varies considerably in space, depending on the district in which the attributes are in competition. Some districts seem to favor the quality of the building (with the maximum bargaining power of 0.664 in Anjou), others being influenced by the size of the land (like 0.385 in Ile-des-Soeurs). And most interestingly, in contrast to what is believed in practice, a good quality parcel of land does not necessarily have a good quality building according to these average results per district.

The bargaining powers of the L1, L2, B1, and B2 are respectively 0.219, 0.155, 0.294, and 0.332. These average values are in coherence with the observations in practice. The ratio of land to total value oscillates between 0.195 and 0.631, with an average of 0.374 for all the districts. In comparison, the average ratio of the land is 0.348 in the case of the city evaluations (values between 0.187 and 0.511) and 0.399 according to the OLS model (with values between 0.194 and 0.626).

The following figure gathers the behavior of the results of the land and the building in four different districts selected from the data (the price of the land is sorted in increasing order) (Fig. 1). By observing Fig. 1a–c, we can discuss several striking

interpretations. First, the building values in general closely follow the same trend as the land values, considering districts of Dorval and St-Léonard. Another observation in these districts is the apparent fluctuations of building values in Dorval, but a more stable trend in the values of both components in St-Léonard. This might signify that the increasing values of lands are almost systematically improved with the increasing quality of the buildings. The situation is completely different in the case of Anjou where the building values fluctuate indistinctly in comparison to the slightly increasing values of lands, with a large gap between the values of the two components.

In the bottom part of the same figure, the separate values of the land steadily increase and even overtake the values of the building for most of the properties in Westmount (Fig. 1d). Building values in this district are almost insensitive to the increases in land values, meaning that the capacity of the land is far from being optimized. We can extend this interpretation to the concept of “sleeping rents” on which a political debate can be elaborated. In fact, rents being captured “under the land” might signify a loss of value and consequently less income in property taxes for the municipalities in the long run. Recognition of that situation might motivate politicians to take some incentive measures, such as property tax relief, to encourage a better quality of buildings. Figure 1e shows the separate values of the same lands and the buildings, estimated almost by a rule of thumb, based on a typical ratio that professionals use in the city. This hides the reality of the practical dialectic between the two components and shows at the same time the importance of reliably estimating the separate values.

The separate results of the mass valuation on the map show these patterns, especially interesting from a spatial economics point of view (Fig. 2). They clearly follow what Alonso (1964) described in his bid rent theory, hypothesizing a decreasing pattern of land values with distance from the center of a city. According to separate estimations of Shapley method, the building values are higher around the old high-status markets

Fig. 1 Averages of land and building values in four different approaches. In a–c, land value are clearly lower than the building. It is interesting to notice the building quality fluctuations, which are very similar in St-Léonard. Elsewhere, in Westmount in d, land value is significantly higher according to the Shapley's improved method, in comparison to city valuations e.

Fig. 2 Land and building values in Montreal Island. The map interestingly shows a negative correlation between land and building values: in general, lands of higher value accommodating less expensive buildings, and vice versa (verifying here the trade-off between the location and building attributes under income constraints, as Alonso (1964) emphasized).

encircling Mount Royal, while also being present far from downtown areas, especially in the west part of the city (due to the demand for direct access and greater visibility for lands along the border of the St. Lawrence River).

Discussion

Value clearly is a fundamental concept and knowledge about the separate contributions of its attributes is important in multiple contexts, for various economic agents. Ignoring separate values of property attributes and avoiding their interactions not only perpetuates the belief that they are meaningless, but it may also eliminate precious information. The issue of value separation is however important and challenging both in theory and in practice (Özdilek, 2012). There are mainly three types of difficulties that are related to the identification/definition of the land and building attributes, the measure of their separate values and the application of a reliable method. Arguments on the separability/inseparability thesis regarding the contributive values have been quite extensive in previous research. There are also several value separation methods in existence, each having some degree of consistency in practice.

This work contributes to the issue of separation by exploring and applying the Shapley scheme of value as a new method, involving both the conceptual basis of game theory and the use of fuzzy logic in the settlement of the initial conditions. This method provides a robust solution, with a relatively simple and accommodating basis, to the issue of separate and total value

estimations. It allows professionals to tangibly (and even manually) break down a total value by clearly understanding the computational steps, otherwise uncertainly applied in practice.

Reference to game theory is quite new in a field where the analysts are used to thinking in terms of the marginal contributions of property attributes. The role of the player is given to the attributes in a strategically cooperative (or not) game of value creation/destruction. This parallel is helpful in understanding the attributes of utility that are in a weaker or stronger position to influence the economic deciders who actually play the game. These attributes are in the position to “claim”, as we defend here, their separate part of the contribution in the game (individually or in cooperation). In settling the “rules” and the most probable constraints, the fuzzy measures also provide an interesting solution in taking into account the interactions among the value creating attributes in the game. In that context, we have introduced a Gaussian function that “defuzzifies” the fuzzy measures of probabilities in the set of combinations.

The Shapley method globally provides separate contributions (or payoffs) of property attributes, as shown in this work. The empirical validation of this method rests on the use of an important database from the market of 14,715 sold single-family properties. The results are very coherent and without ambiguity for all the properties within the data. They are in line with practical expectations of values, either taken case-by-case or in aggregation per district. They are indeed encouraging when compared to the results of the city evaluations and the OLS predictions for the same properties and attributes. The results are

derived from an objective solution to the difficulty of the separation issue, offering there the occasion to appreciate other related issues from the land-building *dialectic*, which are usually overlooked or simply ignored.

The Shapley method depends on the reliability of weights of attributes processed as fuzzy measures within the Gaussian function. The initial values based on the correlations of attributes per district are iteratively changed and stabilized under the constraint of the minimization of distances (errors) between the adjusted prices of comparables and the observed price of the subject. Professionals will benefit from the reliability of the method, but most importantly, they can confidently defend and demonstrate that those weights are objectively in line with the theoretical and practical expectance in valuation.

Further work. Based on the general framework of the Shapley method that we presented, future work might be interested in several technical aspects to enhance separate values. For instance, we processed direct separation of observed prices based on the Shapley method. Because of the possibility of residual interactions among attributes that can resist after the application of the Shapley's global approach, an extended version can be elaborated (for instance, by the integration of simultaneous interactions of two or more attributes on other ones in the same framework).

Even though we presented only briefly, the Shapley method also has a coherent framework for the prediction of the separate values. The process of prediction can be submitted to a robust modeling algorithm with feed forwarded and back forwarded learning steps of separation under the constraints to decrease the errors of estimation. This sounds advantageous especially in the context of mass evaluation, but the tangible benefits of the Shapley method might become complex for professionals. There is also the possibility of improving several other steps in the method, such as using another probability distribution function and changing the number of attributes (and their weights). In this perspective, spatial autocorrelation methods can be joined to the Shapley method, especially in improving the explanation portion of the land value. Also, in future work, more discussion on the political issue of the separate system of tax rates (and its consequences) would be beneficial for public finance management. Finally, using Shapley axioms, more sensitive approaches can focus, for instance, on the optimal timing of improving vacant lands and demolishing the depreciated buildings for new constructions.

Data availability

The author confirms that the data supporting the findings of the example in this study are available within the article. The raw dataset is not publicly available due to legal restrictions (they contain individual information that could be accessed and compromise the privacy of properties and their owners).

Received: 9 October 2019; Accepted: 13 March 2020;

Published online: 24 April 2020

References

Alonso W (1964) Location and land use: toward a general theory of land rent. Harvard University Press, Cambridge

Andelson RV (2000) Land-value taxation around the world. Blackwell Publishers

Anenberg E (2016) Information frictions and housing market dynamics. *Int Econ Rev* 57:1449–1479. <https://doi.org/10.1111/iere.12204>

Appraisal Institute (2013) The appraisal of real estate, 14th edn. Appraisal Institute, Chicago

Atanassov K (1986) Intuitionistic fuzzy set. *Fuzzy Sets Syst* 20:87–96. [https://doi.org/10.1016/S0165-0114\(86\)80034-3](https://doi.org/10.1016/S0165-0114(86)80034-3)

Bagnoli C, Smith CH (1998) The theory of fuzz logic and its application to real estate valuation. *J Real Est Res* 16(2):169–199

Barranco CD, Campana J, Medina JM, Pons O (2004) ImmoSoftWeb: a web based fuzzy application for real estate management. *Lect Notes Artif Intell* 23034:196–206

Basu S, Thibodeau TG (1998) Analysis of spatial autocorrelation in house prices. *J Real Est Fin Econ* 17(1):61–85

Bonissone PP, Cheetham W (1997) Financial applications of fuzzy case-based reasoning to residential property valuation. In: *The IEEE Neural Networks Council (ed) Proceedings of the 6th International Fuzzy Systems Conference*, vol. 1. Barcelona, Spain, pp. 37–44

Bourassa SC, Hoesli M, Scognamiglio D, Zhang S (2011) Land leverage and house prices. *Regional Sci Urban Econ* 41(2):134–144. <https://doi.org/10.1016/j.regsciurbeco.2010.11.002>

Branzei R, Dimitrov D, Tijs S (2008) Models in cooperative game theory, vol. 556. Springer Science & Business Media

Brueckner JA (1986) Modern analysis of the effects of site value taxation. *Natl Tax J* 39(1):49–58

Clapp JM, Salavei K (2010) Hedonic pricing with redevelopment options: a new approach to estimating depreciation effects. *J Urb Econ* 67(2):362–377

Connellan O (2004) Land Value Taxation in Britain: Experience and Opportunities, Lincoln Institute of Land Policy, Cambridge, Massachusetts

Dancasescu N (2000) Assessment value to market value ratios in Alachua County, FL: an examination using GIS. Thesis, University of Florida

Davis M, Palumbo MG (2008) The price of residential land in large US cities. *J Urb Econ* 63:352–384

Dilmore G (1993) Fuzzy set theory: an introduction to its application for real estate analysts. Paper presented at the annual conference of the American Real Estate Society in Key West

Dorau HB, Hinman AG (1969) Urban land economics. McGrath Publishing Company, College Park

Dye RF, McMillen DP (2007) Teardowns and land values in the Chicago metropolitan area. *J Urban Econ* 61:45–63. <https://doi.org/10.1016/j.jue.2006.06.003>

Ely TR (1922) Research in land and public utility economics. *Land Econ* 1(1):1–5

Epley DR (1997) A note on the optimal selection and weighting of comparable properties. *J Real Est Res* 14(1–2):175–182

Fischel WA (2000) Municipal corporations, homeowners and the benefit view of the property tax. Research paper, Lincoln Institute for Land Policy

Fisher EM (1958) Economic aspects of urban land use patterns. *J Ind Econ* 6(3):379–386

Fox J (1991) Regression diagnostics. Sage Publications, Newbury Park

Gaffney M (1994) Land as a distinct factor of production. In: Tideman N (ed), *Land and taxation*. Shephard-Walwyn, London, pp. 74–82

Genesove D, Mayer C (2001) Loss aversion and seller behavior: evidence from the housing market. *Quart J Econ* 116(4):1233–1260. <https://doi.org/10.1162/003353001753265561>

George H (1881) Progress and poverty, 4th edn. Schalkenbach Foundation, New York

Gihring TA (1999) Incentive property taxation: a potential tool for urban growth management. *J Am Plan Assoc* 65(1):64–79. <https://doi.org/10.1080/0194369908976034>

Gloudemans RJ (2001) An empirical analysis of the incidence of location on land and building values. Working paper, Lincoln Institute of Land Policy

Goetzmann WN, Rouwenhorst, KG (2005) Origins of value: a document history of finance. Oxford University Press.

Grabisch M (1996) The representation of importance and interaction of features by fuzzy measure. *Pattern Recognit Lett* 17:567–575. [https://doi.org/10.1016/0167-8655\(96\)00020-7](https://doi.org/10.1016/0167-8655(96)00020-7)

Grabisch M (1997) *k*-order additive discrete fuzzy measures and their representation. *Fuzzy Sets Syst* 92:167–189. [https://doi.org/10.1016/S0165-0114\(97\)00168-1](https://doi.org/10.1016/S0165-0114(97)00168-1)

Grabisch M, Raufaste E (2008) An empirical study of statistical properties of Choquet and Sugeno integrals. *IEEE Trans Fuzzy Syst* 16(4):839–850

Greyserman A, Jones DH, Strawderman WE (2006) Portfolio selection using hierarchical Bayesian analysis and MCDM methods. *J Bank Fin* 30:669–678. <https://doi.org/10.1016/j.jbankfin.2005.04.008>

Guofang Z, Teruki F, Saburo I (2003) Effect of flooding on megalopolitan land prices: a case study of the 2000 Tokai flood in Japan. *J Natl Disaster Sci* 25(1):23–36

Hendriks D (2005) Apportionment in property valuation: should we separate the inseparable? *J Prop Invest Fin* 23(5):455–470. <https://doi.org/10.1108/14635780510616043>

Hu CC, Yang YY (2010) An axiomatic characterization of a value for games in partition function form. *SERIEs* 1:475–487. <https://doi.org/10.1007/s13209-009-0004-9>

- Ishii K, Sugeno M (1985) A model of human evaluation process using fuzzy integral. *Int J Man-Mach Stud* 22(1):19–38. [https://doi.org/10.1016/S0020-7373\(85\)80075-4](https://doi.org/10.1016/S0020-7373(85)80075-4)
- Kahneman D, Slovic P, Tversky A (1982) *Judgment under uncertainty: heuristics and biases*. Cambridge University Press, Cambridge
- Keligion DL (1994) Appraisal issues now require greater attention for tax planning to be effective. *J Tax* 80(2):98–103
- Kitchen H (2003) Property taxation: issues in implementation. Working paper. Trent University, Peterborough, Ontario
- Kummerow M, Galfalvy H (2002) Error trade-offs in regression appraisal methods. In: Wang K, Wolverton M Eds. *Real estate valuation theory*. Kluwer Academic Publishers, Norwell, pp. 105–132
- Magoc T, Wang X, Modave F, Ceberio M (2011) Applications of fuzzy measures and intervals in finance. *Reliab Comput* 15:300–311
- Marshall A (1922) *Principles of economics*. Macmillan, London
- McCluskey WJ, Anand S (1999) The application of intelligent hybrid techniques for the mass appraisal of residential properties. *J Prop Invest Fin* 17(3):218–238. <https://doi.org/10.1108/14635789910270495>
- Mills ES (1967) An aggregative model of resource allocation in a metropolitan area. *Am Econ Rev* 57(2):197–210
- Mills ES (1998) Is land taxation practical? University of Illinois. *Real Est Lett* 12(4):1–5
- Moriarty S (1975) Another approach to allocating joint costs. *Account Rev* 26(2):791–95
- Murofushi T, Soneda S (1993) Techniques for reading fuzzy measures (III): Interaction index. In: Japan Society for Fuzzy Sets and Systems (ed) *Proceedings of the 9th Fuzzy System Symposium*, Sapporo, Japan. pp. 693–696
- Muth RF (1969) *Cities and housing: the spatial pattern of urban residential land use, third series: studies in business and society*. University of Chicago Press, Chicago
- Nash J (1951) Non-cooperative games. *Ann Math* 54(2):286–295
- Neumann J, Morgenstern O (1944) *Theory of games and economic behavior*. Princeton University Press
- Ohno K (1985) Preparation and use of land value maps. *Appraisal J* 53(2):262–268
- Osborne MJ, Rubinstein A (1994) *A course in game theory*. MIT Press
- Özdilek Ü (2012) An overview of the enquiries on the issue of apportionment of value between land and improvements. *J Prop Res* 29(1):69–84. <https://doi.org/10.1080/09599916.2011.583670>
- Özdilek Ü (2016) Property price separation between land and building components. *J Real Est Res* 38(2):205–228
- Quigg L (1993) Empirical testing of real option pricing models. *J Fin* 48(2):621–640. <https://doi.org/10.2307/2328915>
- Ratcliff RU (1950) Net income can't be split. *Apprais J* 18(1):168–172
- Ricardo D (1817) *On the principles of political economy and taxation*, 1st edn. John Murray, London
- Rice H (1982) The value of developed land considered vacant and unimproved. *Real Est Rev* 37(1):45–50
- Riley NF, Waller BD, Greer TH (2001) An appraisal tool for the 21st century: automated valuation models. *Aust Prop J* 36(7):636–641
- Roth AE (1988) *The Shapley value*. Cambridge University Press, New York
- Russell B (1945) *History of western philosophy*. G. Allen and Unwin, London
- Shapley LS (1953) A value for n -person games. In: Kuhn HW, Tucker AW (eds) *Contributions to the theory of games, vol. II* (28) in *Annals of Mathematics Studies*. Princeton University Press, pp. 307–317
- Sirmans GS, Macpherson DA, Zietz EN (2005) The composition of hedonic pricing models. *J Real Est Lit* 13(1):3–43
- Smith A (1886) *An inquiry into the nature and causes of the wealth of nations*. Modern Library, New York
- Soibelman L, Gonzalez MA (2002) A knowledge discovery in databases framework for property valuation. *J Prop Tax Assess Admin* 7(2):77–106
- Su Z, et al. (2019) Distribution-based approaches to deriving weights from dual hesitant fuzzy information. *Symmetry* 11(85):1–20. <https://doi.org/10.3390/sym11010085>
- Sugeno M (1974) *Theory of fuzzy integrals and its applications*. PhD thesis, Tokyo Institute of Technology
- Sun Y, Huang R, Chen D, Li H (2003) Fuzzy set-based risk evaluation model for real estate projects. *Tsinghua Sci Technol* 13:158–164. [https://doi.org/10.1016/S1007-0214\(08\)70143-3](https://doi.org/10.1016/S1007-0214(08)70143-3)
- Torra V (1999) On hierarchically s -decomposable fuzzy measures. *Int J Intell Syst* 14(9):923–934. [https://doi.org/10.1002/\(SICI\)1098-111X\(199909\)14\(9\):923-934](https://doi.org/10.1002/(SICI)1098-111X(199909)14(9):923-934)
- Vandell KD (1991) Optimal comparable selection and weighting in real property valuation. *AREUEA J* 19(2):213–239
- Wallenius J, et al. (2008) Multiple criteria decision making, multiattribute utility theory: recent accomplishments and what lies ahead. *Manag Sci* 54(7):1336–1349. <https://doi.org/10.1287/mnsc.1070.0838>
- Weber RJ (1988) Probabilistic values for games. In: Roth AE Ed. *The Shapley value. Essays in honor of Lloyd S. Shapley*. Cambridge University Press, New York, pp. 101–119
- Wickens T (1995) *The geometry of multivariate statistics*. Lawrence Erlbaum Associates, inc., Hillsdale
- Winter E (2002) The Shapley value. In: Aumann RJ Aaron M (ed) *Handbook of Game Theory with Economic Applications*, 1st edn. vol. 3 Elsevier, North Holland, pp. 2025–2054
- Wyatt PJ (1997) The development of a GIS-based property information system for real estate valuation. *Int J Geogr Inf Sci* 11(16):435–450. <https://doi.org/10.1080/136588197242248>
- Wyman D, Seldin M, Worzala E (2011) A new paradigm for real estate valuation? *J Prop Invest Fin* 29(4/5):341–358
- Yavas A, Sirmans CF (2005) Real options: experimental evidence. *J Real Estate Financ Econ* 31(1):27–52. <https://doi.org/10.1007/s11146-005-0992-6>
- Young HP (1988) Individual contribution and just compensation. In: Roth AE Editor *The Shapley value: essays in honor of Lloyd S. Shapley*. Cambridge University Press, New York, pp. 267–878
- Zadeh L (1965) Fuzzy sets. *Inf Cont* 8(3):338–353
- Zimmermann H-J (2001) *Fuzzy set theory—and its applications*, 4th edn. Springer Science and Business Media, New York

Competing interests

The author declares no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Ü.Ö.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020