LUISS

Dipartimento
di Impresa e Management

Cattedra di Microeconomics

An agent-based mechanism for the contingent valuation of a shared good

Prof. L. Marengo
RELATORE

237751
MATRICOLA

Index

1. Introduction

This work aims at modelling a novel method to carry out contingent valuations. The method is based on a type of simulation technique by the name of "agent-based modelling". It essentially consists in deconstructing a complex system into smaller entities, called 'agents', and letting these entities act according to one or more algorithms. The purpose of this methodology is to simulate the emergence of complex patterns starting from simpler rules, and therefore provide "historically genetic explanations" (Stegmüller 1983).

The first chapter introduces agent-based models giving an operative definition and outlining the history of their developments, from the origins to future applications. A summary of the main characteristics of agent-based modelling, together with their advantages and disadvantages is provided.

In the second chapter, a brief dissertation on different approaches to valuation is provided together with a synthesis of the main implicit theoretical underpinnings that are relevant to understand the rationale of the model presented herein. In particular, the chapter succinctly outlines the differences between the substantialist and subjectivist approach to the theory of value. Moreover, a famous theoretical framework to classify goods is described, that is helpful in introducing the type of good on which this model is based: the public good.

Next, the third chapter introduces the model, formulated and formalised in a more rigorous way. Short and simple mathematical considerations are provided to frame what to expect from the simulation from a probabilistic point of view.

Finally, the code is presented, and its results are stored, analysed and discussed. Surprisingly, the simulations show that the model is promising, in that it is robust and consistent. But, along the positive aspects, also its limitations are explained, and suggestions for further improvements are proposed.

2. Agent-based models

In this chapter, the notion of agent-based model (thereafter 'ABM') is presented through a summarized history of its developments. In particular, the second section aims at framing the context for ABMs starting from their origins, to presenting well-known results, and by fostering what is still to come in the field. Some examples are presented in more detail in order to help the reader understand the subtle differences within the different simulation and modelling approaches, as well as to show some useful applications, especially in policymaking.

2.1 What ABMs are

Agent-based modelling is a theoretical and computational framework that allows researchers to run simulations and tests of problems that are usually difficult to solve by using a purely mathematical approach with paper and pencil. The 'agents' in question are a collection of autonomous entities that are capable of interacting or making decisions following a strategy that is dictated by an equation or a rule.

It is possible to summarize the main characteristics of an ABM as follows:

- There is a finite set of agents.
- An environment in which they can interact.
- Agent can perform a set of actions according to their decision rules.
- The interaction between an agent and its environment (other agents or external inputs) generates changes in some features over time.

Some authors (see Klein D., Marx J., and Fischbach K., 2018) also identifies the spatial dimension as a key characteristic of ABMs, although they include in the definition of 'space' also abstract measures of closeness such as the degree to which opinions diverge (as in Scheller 2018, Baumgaertner 2018).

The common features of these techniques, as well as a detailed account of their pros and cons will be given after having shed a light over the history and the surrounding context in which ABMs developed.

2.2 History of Agent-based modelling

2.2.1 The origins

Some scholars trace back the origins of agent-based computational modelling in the pioneering work of J.S. Coleman in the 1960s, although ABMs became to be known among scholars only after the publication of the first two volumes of *The Journal of Mathematical Sociology* in 1971 (Bianchi and Squazzoni 2015, 284). In particular, Schelling's segregation model became famous for its results on the emergence of spontaneous spatial segregation among agents with group preferences. Schelling's model essentially shows that segregation does not necessarily result from a discriminatory behaviour held by a segregating group towards a segregated one, but it rather emerges from the preference of each agent to be neighbour of an agent of the same group.

Among the other popular advances in agent-based modelling, albeit with a different name, there is the famous Conway's Game of Life, published in 1970, which has contributed towards the creation of an entirely new field of academic research called "cellular automata", at the edge of computer science, mathematics, and biology. Conway's Game of Life is an algorithm that generates unpredictable patterns of clustering 'cells' (gluing pixels on a grid) that evolve according to simple rules. It is worth noting that defining Conway's Game of Life as an ABM is legitimate if one identifies each cell as an agent. But given the deterministic nature of the Game, together with the fact that rules are the same for each cell, scholars conventionally assign Conway's Game of Life to a slightly different research line. It is can certainly be said that it is part of the numerous works that inspired the creation of agent-based simulations at the beginning of ABMs.

2.2.2 ABMs' growing popularity

With the growing popularity of personal computers in the 1980s, scholars took the opportunity to exploit the new technological trends to apply computer-based simulations on an increasing variety of topics. At the same time, almost every field of research had open problems that were considered intractable due to the large amount of computation required.

In Growing Artificial Societies (1996), the very pioneers of agent-based modelling, J. Epstein and R. Axtell, presented the famous Sugarscape model, a simulated environment where agents could exchange their wealth. The agents in the Sugarscape model were living on a 51x51 units' grid, and they were capable of doing many types of actions such as collecting sugar (a resource on the grid), trade, die, pollute, inherit and so on. Although it was difficult to replicate, the Sugarscape model is regarded as a milestone in ABM because it was one of the first times in history that a social simulation was not seen just as a thought experiment, but it was used to convey stylised facts about the aggregate from simple agent-level decision rules. In particular, it was evident that the agents on the grid formed clusters depending on the distribution of the sugar. Not only demand and supply curves at which agents would trade their goods were an accurate reflection of what the theory on microeconomic equilibriums predicted, but the simulation also showed the endogenous formation of shifts from equilibrium, resulting in continuous redefinitions of the curves. The Sugarscape model begun well-known, among the other things thus far discussed, because it also succeeded in simulating the right-skewness of the distribution of wealth, by setting simple exchange rules for the agents at the beginning of the simulation.

Later on, different disciplines started to adopt the same approach to search for emergence patterns. In epidemiology – to name a few – following the seminal work of Epstein and Axtell, who provided an important agent-based model for epidemics transmission in 1996, many scholars worked on ameliorating the predictive capacity of epidemic transmission simulations and developed one of the most applied techniques for the estimate of epidemiological parameters during the Covid-19 pandemic (see OpenABM-Covid19). In Nagel and Rasmussen (1994) the authors developed an ABMs for traffic congestion, and similarly ABMs have been applied to a large array of topics, including: stock market price series (Arthur et al., 1999), trade patterns (Tesfatsion, 1995), believes dynamics and the role of influencing opinions (Hegselmann and Krause, 2002), dynamics of scientific discoveries (Weisberg and Muldoon, 2009), dynamics of political parties (Schmitt and Franzmann, 2018), learning dynamics in the context of reliable information sharing (Boero et al., 2010).

Undoubtedly, one on the major repercussions of agent-based modelling literature can be seen in macroeconomics and policymaking. In the 1990s, a generalised sense of dissatisfaction towards the traditional macroeconomic tools, that always come short in predicting new crises, together with a need to reconcile the different contrasting theories of economics within a unique comprehensive framework, became the perfect occasion for ABMs to step in. Indeed, ABMs could possibly foster the search for micro-foundations so highly coveted by macro theorists. That is why different attempts were made to bring ABMs into policymaking, the most ambitious of them being EURACE.

2.2.3 ABM in policymaking: the EURACE

In 2006 the EURACE project was launched, a massive macroeconomic agent-based simulator with the ambition of depicting the whole European economy. The set up and the subsequent implementations took up to 10 years. It comprehends three types of learning agents: households (around 10^7), firms (around 10^5) and banks (around 10^2). Five types of market are modelled, namely: consumption goods, investment goods, labour, credit and financial assets. Governments, central bank, and Rest-of-the-world actions are changed to see the effects on the systems and help policymakers understand the impacts of their decisions. Each agent is equipped with a double-entry balance sheet, and agents can interact with each other by removing and adding entries in the balance sheet, consistently with the simulated market value of their assets.

The EURACE project was completed in 2016, requiring huge efforts both in financial terms and in human capital, since even the 2001 Nobel laureate J. Stiglitz was included in the pool of experts that worked at implementing the model. The EURACE infrastructure can count on a parallelized modelling environment, called FLAME (Flexible Large-scale Agent Modelling Environment), developed at University of Sheffield (see www.flame.ac.uk for further details) that enables the ABM to run in parallel in long series of supercomputers in order to speed up the computational work.

Nonetheless, it is important to stress that, albeit the huge resources invested in such projects, EURACE is still not able to compete in terms of explanatory power with traditional econometrical tools (see Deissenberg C., van der Hoog S., Dawid H., 2008). The EURACE is then mainly used as a secondary investigation tool to have more insights on the repercussions of particular non-canonical manoeuvres.

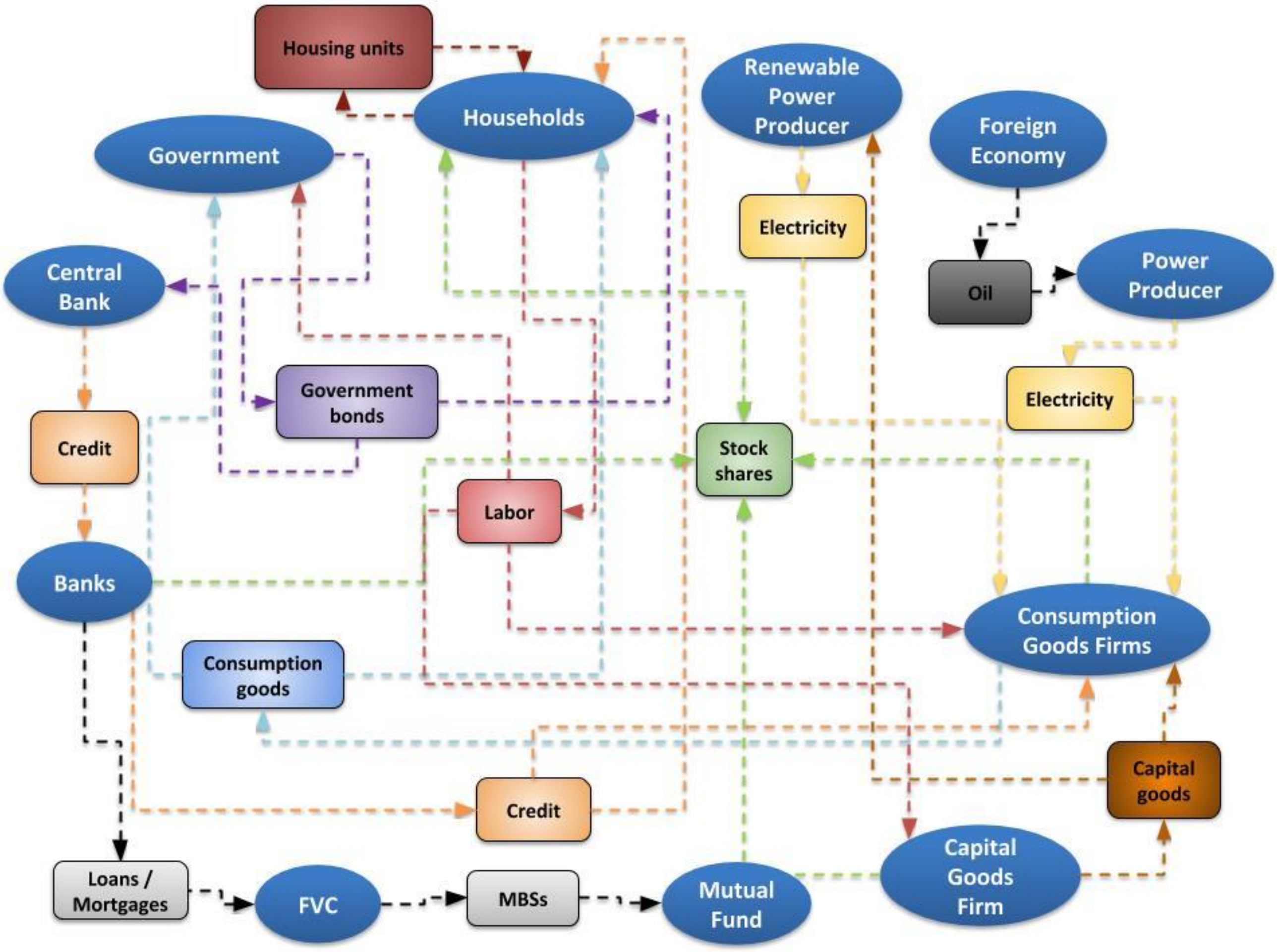


Figure 1: Official EURACE model diagram retrieved at www.eurace.org

2.2.4 Modern and future developments

Moreover, a growing body of literature in ABMs may nowadays also involve neural networks to simulate complex situations. Even though neural networks are not a necessary feature of ABMs, we stress how this new field will be probably addressed as the natural continuation of ABMs, since neural networks can help in creating agents that are more adaptive in their virtual environment. For example, in P. Dütting, Z. Feng et al. 2020, the authors make use of ABMs and neural network to confirm some important theoretical results from the field of auction theory, where agents learn how to optimally bid in a multi-item English auction. The study confirms that it is possible to achieve the same theoretical results already known for actions, such as optimal bidding price, but it is also possible to manipulate the rationality and the heterogeneity of the agents in order to get more insights on the multiple strategies that they can employ.

2.3 Characteristics, advantages, and disadvantages

2.3.1 Distinctive traits

Thus far, key characteristics of ABMs have been discussed, alongside their history. So now one might wonder: why should a modeller choose AB modelling over traditional analytical and statistical techniques? The answer is that ABMs make less assumptions regarding the representative agent, an equation that shows the behaviour of the agents at the aggregate level. But instead, they very much rely on flexible assumptions. In this way, it is possible to twist the features of each agent so that some necessary assumptions on the representative agent can be relaxed. In particular, in an ABM:

- Agents' rationality can be bounded
- Agents' information can be limited
- Agents can be heterogenous and they can follow different decision rules

All of this of course comes at the expense of robustness and generality: some ABMs may lead to very different results even if initial conditions are just slightly different.

2.3.2 Advantages over traditional modelling techniques

It is possible to argue that ABMs are useful especially for four reasons:

i. They can capture emergent systemic phenomena.
ii. They are more directly related to the observable entities.
iii. They are flexible in that they avoid making many assumptions on the nature of the phenomenon.
iv. They can be easily parallelized so that the time required to carry out the computation can be greatly reduced.

The first three are also argued by E. Bonabeau (see E. Bonabeau, 2002). These four characteristics make ABMs particularly useful in social sciences. In particular, the first one helps researchers to rethink about the underlying causal relationships of a social phenomenon. To know what causes a phenomenon, researchers must plug into their equations a quantity to describe a variable, but many times this epistemological tool does not work properly, because there is often an interdependence between endogenous and exogenous variables. Thanks to ABMs, scholars have realised that many patterns, that were previously described through convoluted differential equations, could be more easily posed and understood if allowed to evolve by themselves on a carefully chosen set of rules.

The second important consideration made by Bonabeau is especially true for models that aims at simulating complex interactions that would have been otherwise treated as set of equations describing the evolution of indirect quantities, such as the population density to describe pedestrian flows on a bridge. In this sense, it would be better to directly refer to the single agent's movements and then join each contribution, rather than deriving an abstract quantity that could describe the state of the whole system. The third consideration is also similar. Following the same example of a pedestrian flow on a bridge, if one was to describe the phenomenon on the aggregate, he would assume homogeneous mixing, but reality can easily deviate from this assumption. In his TEDxUMV talk in 2011, Axtell explained the advantages of ABMs using a table similar to the following, that can easily summarize the points discussed earlier:

Simple models	**Complex models**
Global information	Local information
Perfect rationality	Bounded rationality
Single decision-maker	Multi-agent institutions
Homogeneous mix	Heterogeneity, heavy-tails, infinite variance
Equilibrium, fixed points, static solutions	Perpetual adaptation and co-evolution
Markets: law of one price	Personal prices

Table 1: Axtell's simple/complex dichotomy

The fourth consideration instead is a practical advantage that comes natural with ABMs. Since agents can be easily represented in a program, and they need to operate independently, it is also easy to parallelize computations. This comes particularly helpful when hundred million of agents are included in the model, but it is also a feature that common computer architectures can achieve by multithreading.

2.3.3 Disadvantages of ABMs

It is important to stress also what kind of drawbacks the aforementioned advantages hide. Surely the most significant disadvantages can be summarized in the following points:

- Lack of robustness.
- Difficulty in accurately representing all the agents and their decision rules.

- It is not always possible to derive simple solutions.

The lack of robustness is essentially omnipresent in ABMs. It is indeed difficult to interpret the results of an ABM because a tiny change in initial conditions usually means a huge change in the output of the simulation. This becomes even more problematic once we add the second issue of ABMs: the difficulty in representing real agents. Because, if real agents are accurately depicted in the model, then the simulation may potentially yield extremely accurate results. Otherwise, results will be perfectly precise as perfectly wrong.

The third disadvantage, instead, is more of an epistemological issue. To what degree, if any, can ABMs prove something? What ABMs are indeed addressing are not really the *causes* of a problem, but rather how the problem get generated as a product of a convoluted set of mutually reinforcing variables. So, this change in perspective also implies that most ABMs cannot really tell what can be done to solve the issue in the first place. The explanations provided by ABMs, therefore, fall in another category of explanations for which Stegmüller coined the term “historically genetic explanations” (Stegmüller, 1983). In other words, many times the lack of generalization makes a model less interpretable and it loses explanatory power.

3. Valuations

This chapter deals with valuations, introducing the concept of value and the underlying principles of different valuation techniques, as well as the dichotomy between a substantialist approach to value versus a subjectivist approach. Also, a summary of a well-known conceptual framework for classifying goods according to the two dimensions of excludability and rivalry is provided. The fourth section deals with the problem of the free-rider, arising from public goods, and presents a famous theoretical solution to it, the Lindahl equilibrium, which is briefly described to introduce the more practical alternative of contingent valuation. Outlining its problems, the fifth section prepares the ground for the novel model presented in the next chapter.

3.1 Underlying principles

One of the central problems that almost all the fields of economic research must deal with, at least to some degree, is valuation. In contrast with the difficulties in carrying out a proper valuation, the underlying assumptions are rather simple: for each resource there is a value, whether it be relative to each person or to an aggregate, and a set of techniques that aim to predict it in the most accurate way.

This problem is so central in economics, that a whole branch of economic theory developed after it. Different and often contrasting theories of value, that aim to answer the questions: is there a correct way to price a good? If there is, what can be done to measure it? And if there isn't, how can one practically assess how convenient are day-to-day exchanges or how much should be paid for a public deed, an environmental tax or a university fee?

3.2 A brief history of different theories of value

3.2.1 Value and price

Since it is difficult to convey on what should be the right price of an item, many modern economists gave up in defining what value is. Boltanski et al. 2015 arrived to say "Value talk only happens in situations in which there is a problem with the price... So, what is the function of "value"? It is the justification of the price, plain and simple". But the problem of value, even in its simplest form of a justification of price, is indeed still relevant, especially when dealing with public goods, where a sense of fairness is a real and constant concern of the legislator.

This question created also the first debate in economic theory, that is, whether or not a good has an intrinsic value.

3.2.2 Substantialism in the form of labour theory of value

The well-known Labour Theory of Value (LTV) was one of the most discussed approaches to the concept of value. Indeed, many thinkers in history understood that the value is dependent to some degree on the labour needed to produce the good. In Wealth of Nations, Adam Smith was supposing that in a primitive society the hours of labour needed to produce a good were the most rudimental way of defining the value of a good. Later on, Ricardo tried to give more grounding to the idea that labour is the main determinant of relative prices, and Marx built his theory of value in a critique to Ricardo's one. For the purposes of our analysis, it would be inappropriate to present a complete summary of the development of LTV, it suffices to know that this approach was addressed to be a 'substantialist' approach: the value of a good is determined by some feature of the good itself (specifically the hours of labour required to produce it).

3.2.3 Subjectivism in the form of marginalism

The idea that goods have value per-se, however, has been at the centre of many debates throughout history. The very concept of subjectivism, with regards to the different theories of value, has been left somewhat ill-defined since scholars of different schools of thought have all defined subjective value in their own way or criticised other definitions (see Mises 1933, Menger 1871, Hayek 1995). Nonetheless, it is possible to say that a subjectivist approach to value is one that recognizes that value is defined by an agent rather than a rule. The classic paradox often cited to criticise a substantialist approach is the diamond-water paradox, commonly attributed to Adam Smith (although earlier thinkers already discussed the same issue in other forms). This thought-experiment is based on the common perception that a diamond is worth more than water, yet for a man starving in the desert surely a bottle of water would be much more valuable than a diamond and he may be willing to pay even more than the price of a diamond if his life was in danger. This is to say that, whatever objective ground value may have, it will always be contingent to specific circumstances and personal evaluations.

The paradox seemed to be solved by introducing the notion of marginal utility, which will be thereafter seen as one of the main pillars of the marginalist theory of value proposed by W.S. Jevons, L. Walras, and C. Menger. The assumptions of marginalism are those of subjectivism (in the sense stated above) plus the idea that there exists an ordinal or cardinal scale of preferences for each individual, and the idea that consumption should be decomposed into units of consumption. To these hypotheses, most neo-classical thinkers also add the law of diminishing returns, which states that the more units of a good one consumes, the less per-unit utility the agent receives.

3.2.4 Some valuations techniques implicitly relying on marginalism

The most iconic example in this regard is the ideal market of goods and services. In this case, one may define the equilibrium price (that is, the intersection of the demand with the supply curve of the good or service that is being traded) as one of the possible measures of the true value of that good. Here, the goodness of the fit could be measured by different indicators, such as liquidity, bid-ask spreads as well as the degree to which the price reflects other measures of value. Moreover, price in a free market is created in a way that is intrinsically subjective to the market participants, it already accounts for many exogenous factors, and the valuation technique in this case just consists in solving for the intersection of the demand and supply curve.

But there are many other more interesting valuation techniques. Auctions, for examples, are especially suitable in pricing goods that are unique. In fact, the illiquidity of the asset prevents to compare the prices of multiple items of the same kind. In the case of auctions, the true value of the item is determined following the rules of the type of auction, but the principle is to maximize the price at which the item is sold. Price that in the English auction is indeed the highest bid for the item.

In other words, valuation can differ widely depending on the assumptions not only about the ‘true’ value of the good, but also about other characteristics of the good.

3.3 Different valuations for different types of goods

3.3.1 The dimensions of excludability and rivalry

Following a well-known conceptual framework, it is possible to define different categories of goods based on two dimensions: excludability and rivalry.

Excludability is the degree to which the consumption of a good can be restricted to include only a specific set of individuals. Rivalry, instead, is the degree to which the consumption of a good by an agent impairs the consumption of the same good by another agent. In this sense, our analysis will be concerned with public goods, which are goods that are nonexcludable and non-rivalrous. Nonexcludable therefore means that it is not possible to exclude other agents from using it, even if they do not have paid for it, and non-rivalrous means that a good can be used by more agents simultaneously without lowering the private utility of any agent.

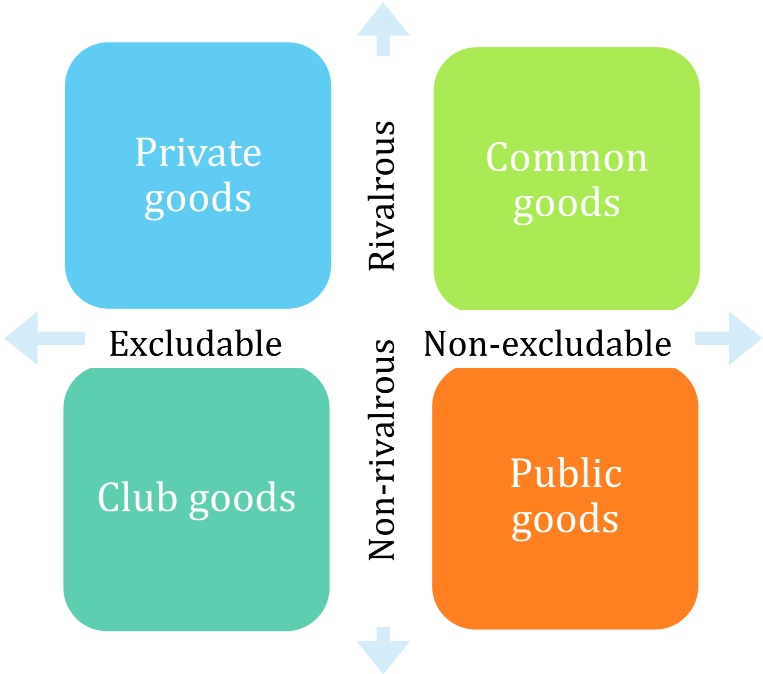


Figure 2: Excludability and rivalry

3.4 The free-rider problem and possible solutions

3.4.1 Definition and naïve approaches

These two characteristics lead to a notorious issue with public goods, the free-rider problem, where an agent that does not contribute for the good can still benefit from it like the ones who pay.

The solution to this problem is not straightforward. In fact, forcing everyone to pay the same amount easily leads to inefficient estimations. The provider of the good may indeed purposely overvalue the individual cost to seek profit, or it could underestimate how important the quality of the good is for the citizens, therefore lowering the individual contribution under the price that the agent is willing to pay. So, what can be done to improve the efficiency of this mechanism in the allocation of costs and resources?

3.4.2 Lindahl equilibrium is difficult to achieve

An important attempt to solve the issue of the free-rider problem was given by Lindahl, when presenting the so-called Lindahl equilibrium, a framework that assumes three important facts:

1. Each agent decides how much of a public good should be provided.
2. Each agent pays a Lindahl tax based on his consumption.

3. The sum of these contributions is enough to pay the cost of providing the public good.

In this way, everyone pays for what he benefits from. Although it is technically the best solution to the free-rider problem, Lindhal equilibrium seems difficult to achieve, and it should be considered more as a theoretical benchmark rather than a practical valuation technique. One of the major problems of this framework, following the reasoning applied in section 3.2, is that it implicitly assumes that it is possible to precisely marginalised the consumption of a public good. There are many examples in which such consumption is difficult to quantify. For example, a public green area at the centre of a city can be assumed to be a public good since it is both non-rivalrous (assuming it is big enough to host a seemingly infinite number of people) and non-excludable (the park is open to everyone). The benefits of having a green area in a city is underestimated if one only considers the value of the hours spent in the park as a measure for private consumption of the good. How about the fresh air, or the benefit in terms of reduced carbon emissions, or the placid view from the neighbouring buildings? All these factors cannot be reduced to only one measure of consumption, and this also implies that deriving demand curves for a public good is essentially infeasible.

It is therefore crucial to develop a valuation mechanism that takes into consideration the following facts:

- Values are relative to each citizen and the units of consumption are difficult to quantify.
- A public good is non-excludable and non-rival, therefore one should try to fix the arising free-rider problem .

3.5 Contingent valuations

3.5.1 What CVs are

To overcome these problems with a more flexible approach, in 1947 S.V. Ciriacy-Wantrup proposed a simple mechanism to value a non-marketable public good, a survey that could help inform about the stated preferences of the citizens and from there make a better estimate of the overall valuation for that good. This process took the name of “contingent valuation” (CV), and it has been widely used to estimate the value of environmental goods and in order to give grounding to environmental taxation.

In its simplest form, contingent valuation consists of these steps:

1. Design a survey in which respondents have to express a stated preferences on how much they would pay for a particular resource.
2. Explain the intricacies of the choice to the respondents: what resource/good/land the question is about.
3. Collect several answers and from them infer the value of the resource and the willingness-to-pay (WTP) of the participants.

It is essential to note that contingent valuation has been applied in a wide array of cases, especially in dealing with environmental goods, healthcare and real estate (see M. Sagoff. 1998, K.G.Willis and G.D.Garrod 1993, Sung-Yoon Huh et al. 2015).

3.5.2 Issues to be solved in CVs

It is easy to imagine some drawbacks in implementing this technique. One of the problems of contingent valuation is that results heavily depend on how the question is asked and on the way in which information is presented. Sometimes, even if the information is correct, the respondent may submit his answer based on other factors, since he may not be incentivized to express his honest opinion. Moreover, setting up and carrying out a survey may be very expensive if compared to all other valuation techniques such as auctions, or market valuations.

Naturally, some of these issues have been partially fixed by the adoption of stricter protocols in the way CVs are carried out. For example, some authors (Arrow et al. 1993) concluded that discrete-choice CV format is to be preferred to the open-ended format. In a famous article, R.T. Carson (Contingent Valuation: a User's Guide, 2000) outlined numerous rules-of-thumb to carry out a contingent valuation that can be meaningful to the maximum possible extent, from the type of information to be attached to the questions, to the statistical tools to be used in analysing the sample of respondents. But the need to systematize CVs and make them more reliable is still present, and the dissatisfaction with the approach is evident in literature.

3.5.3 ABM for contingent valuation

Given these drawbacks, it is possible to slightly change the way in which contingent valuation is carried out by introducing an ABM to predict its efficacy and simulate a discrete-choice CV to estimate the willingness-to-pay of each agent. In other words, it is better to introduce a type of contingent valuation which does not require extensive previous information and whose survey is made up of only closed questions, such as asking whether or not the agent is willing to pay a precise amount for a given tax.

Now, apart from the assumptions presented in chapter 4, there are also some implicit ideas on the nature of the valuation that is required for this type of problem. A brief summary of these ideas is presented below.

3.5.4 Some marginalist assumptions relaxed

One can certainly note that the investigation presented in this work will be relying on the assumptions of subjectivism, that is the assumption that the value of a good is subjective to the agent that benefits from it. It would be erroneous to assume that agents consume a portion of a public good, since the good is assumed to be non-rivalrous so the supply of the good left to the other individuals is not affected by previous consumption. This explains why this marginalist assumption is not essential for our model. This also implies that the model will not have to rely on the law of diminishing returns, since there is no need to introduce a unit of consumption. In short, the marginalist approach seeks to determine which are the demand curves for each individual, while in the model there is no need to actually derive these curves, but only the WTP of each individual, which is a single point, entirely contingent to the specific good. The properties of how these points behave in the aggregate for certain type of public goods for which it is possible to meaningfully determine a unit of consumption is an *ex-post* assessment that is not included in the main purposes of this valuation.

What the model will be assuming, instead, is another fundamental fact: the total value of the good shared among these agents will be defined to be the sum of all the individual private utilities, as if all the negatives or the drawbacks of sharing a good have been already accounted in the private utility by the agents themselves. There is no need though to consider agents as omniscient entities, because even with heterogenous information the result is still a meaningful estimate of the perceived value of a good. And, given our subjectivist framework, we believe that this simplification is therefore useful in many applications, where a great number of agents is involved. In fact, the focus of our investigation is exactly that of presenting a method to discover a measure of the perceived value of a public good and assessing the goodness of this fit by virtually simulating the method under other simplifying assumptions, as it will be shown in chapter 4.

4. The model

In this chapter, an ABM for the contingent valuation of a public good is presented. After having outlined all the model's assumptions, we mathematically derive some interesting properties of it, and we discuss specific examples to help the reader understand the generality of its formulation. Then, the work introduces the next chapter, where the Python code will be presented, and some important results discussed.

4.1 Representation of the model

Assume there are N_0 agents, a non-excludable and non-rivalrous good and a discrete time setting. Each of the agents has a private value v_i (that will not be revealed nor changed), which can be thought of as the private utility of the good according to agent i. Each private value is drawn independently from a distribution $\mathcal{V}$ and it is assigned at time 0. Then, starting from t=1 until the model stops, a cost c_i^t will be drawn independently from a distribution $\mathcal{C}$ and proposed to agent i at time t. At time t=1, the agents will check whether the cost is less than or equal to their value, and they will independently vote "Yes" or "No" based on this comparison. If more than $\frac{N_0}{2}$ agents voted 'Yes', the model stops, the ones who voted 'Yes' (that we may alternatively call 'winners') will pay the amount agreed, while the others will pay the arithmetic average of the costs of the other group. Otherwise, if a majority is not yet reached at time t, costs are redrawn from the same distribution at time t+1 and reproposed only to those who voted against, until a majority is reached; that is, until the total votes in favour from the start exceed $\frac{N_0}{2}$.

In general, the essential elements of the model are:

- A set of agents $\{A_1, \dots, A_{N_0}\}$.
- A non-rivalrous and non-excludable good.
- A discrete time schedule $t_0, \dots, t_s$.
- A distribution of costs $\mathcal{C}$.
- A distribution of private values $\mathcal{V}$.
- A decision rule $a_i(\pi_i): \mathbb{R} \rightarrow \{0,1\}$ for each agent, based on the individual payoff $\pi_i = v_i - \hat{c}_i$, where $\hat{c}_i$ is the cost or contribution that the agent pays at the end. The output of

the decision rule is a Boolean value representing the vote (1 corresponds to a vote in favour).

Then, the Public Value of the good is defined as $V = \sum_{i=1}^{N} v_i$, hence the estimated Public Value is $\hat{V} = \sum_{i=1}^{N} \hat{c}_i$. The ideal result would be proving that it is possible to minimize the difference$|V - \hat{V}|$, which means that the model leaves little room to private exploitation under the assumption that $\hat{V}$ represents the value of the good. Moreover, the sum of the individual payoffs of the agents will be called Social Payoff.

In this work, we will further assume that

- Costs and values are i.i.d. according to a continuous uniform distribution $U(0,1)$, denoting by $f_X(x)$ the pdf of the values, and by $f_Y(y)$ the pdf of the costs.
- The decision rule is the same for all the agents and it is

$$a(\pi_i) = \begin{cases} 1 & \pi_i \geq 0 \\ 0 & \pi_i < 0 \end{cases}$$

  In other words, the agents will vote 'Yes' for a non-negative payoff.

4.2 What to expect: probability of one vote in favour

Moreover, it is possible to derive the distribution of the payoffs. Let X be the r.v. associated to values and Y be the one associated to costs, $f_X(x)$ is the pdf of X, $f_Y(y)$ is the pdf of Y. So, $Z = X - Y$ is the r.v. representing the payoffs. Then, the cdf of Z will be

$$F_Z(z) = P(X - Y \leq z) = \iint\limits_{x-y\leq z} f_{XY}(x,y)\,dx\,dy =$$

$$= \int_0^1 \int_0^{z+y} (f_X(x)\,dx) f_Y(y)\,dy =$$

$$= \int_0^1 F_X(z+y) f_Y(y)\,dy$$

Where we used the independence of the two r.v. to write $f_{XY}(x,y) = f_X(x) f_Y(y)$ and the fact that $x \leq z + y$ to rewrite the upper limit of the integral with respect to x. The integrals range from 0 to 1, since $f_X(x)$ and $f_Y(y)$ are bounded in [0,1].

In particular, $f_Y(y) = 1$ for $0 < y < 1$ and $f_X(x) = 1$ for $0 < x < 1$. Note that it is possible to write the above equations because all the hypotheses of Fubini's theorem are satisfied (the preimages of both $f_X(x)$ and $f_Y(y)$ belong to a sigma-algebra).

Now, it is possible to apply the fundamental theorem of calculus, with pertinent convergency hypotheses, and get

$$f_Z(z) = \frac{d}{dz}\int_0^1 F_X(z+y)f_Y(y)\,dy = \int_0^1 f_X(z+y)\,f_Y(y)\,dy$$

In short, we can write:

$$f_Z(z) = \int_0^1 f_X(z+y)\,f_Y(y)\,dy$$

Since both f_X and f_Y are continuous uniform distributions ranging from 0 to 1, the integrand will be 1 only when both $f_X(z+y) = 1$ and $f_Y(y) = 1$, that is, when also $0 < z + y < 1 \implies -z < y < 1 - z$. Since z ranges from -1 to 1, it is useful to split the range in two cases, as shown in the graph

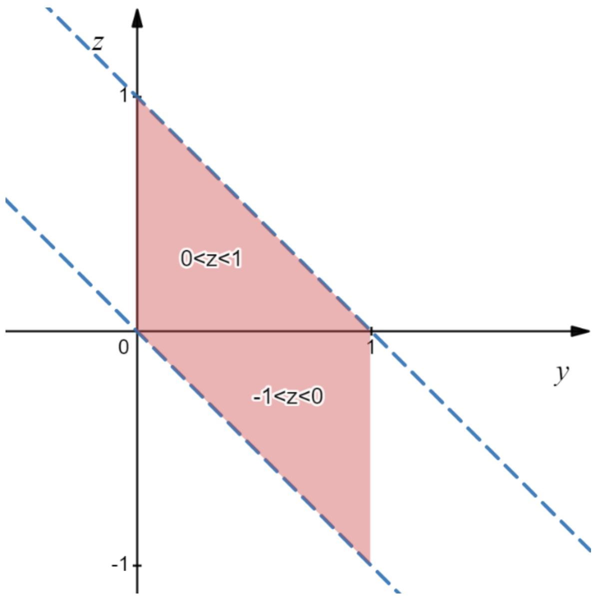


Figure 3: Plot of the integral

For -1<z<0, the upper limit is 1, and the lower limit is $-z$. For 0<z<1 the upper limit is $1 - z$, and the lower limit is 0. Therefore,

$$f_Z(z) = \begin{cases} \int_{-z}^{1} dy & -1 < z < 0 \\ \int_{0}^{1-z} dy & 0 < z < 1 \end{cases}$$

This leads to a triangular distribution centred in 0 with pdf:

$$f_Z(z) = \begin{cases} 1 + z & -1 < z < 0 \\ 1 - z & 0 < z < 1 \\ 0 & elsewhere \end{cases}$$

Therefore, the probability that one agent will be voting in favour is (not surprisingly) ½.

$$P(Z > 0) = F_Z(0 < z < 1) = \left[z - \frac{z^2}{2} \right]_0^1 = \frac{1}{2}$$

4.3 Convergence of the Markov process

Now that we know the probability that a single agent will or won't be voting in favour, since all the costs and values are drawn independently, it is possible to express the probability that there will be w=k positive votes (that we may also call 'winners') in a turn t with N citizens using the binomial distribution:

$$P_t(w = k) = \binom{N}{k} p^k (1-p)^{N-k} = \binom{N}{k} \left(\frac{1}{2}\right)^{k+N-k} = \binom{N}{k} \left(\frac{1}{2}\right)^{N}$$

Note that starting from a different distribution implies having a different p, therefore the first part of the above formula is valid even if the starting distributions of costs and values are different; it is only necessary to replace $p = \frac{1}{2}$ with a different probability of having positive payoffs.

In this specific case, from the equation above one can see that $p = 1 - p$. Moreover, the probability of having k winners in a turn only depends on N, thus, it is possible to express the number of winners in a turn as $k_t = N_t - N_{t+1}$. Given this property, the model can be represented as a Markov process, depending on the parameter N_t . Indeed,

$$P_t(k_t) = \binom{N_t}{N_t - N_{t+1}} \left(\frac{1}{2}\right)^{N_t} = \binom{N_t}{N_{t+1}} \left(\frac{1}{2}\right)^{N_t}$$

Thus, we can write the corresponding transition matrix *P*, in which the elements p_{ij} in each row represent the probability of getting from the state i to the state j, that is from N=i to N=j agents.

$$P(N_{t+1} = j | N_t = i) = p_{i,j} = \binom{i}{j} \left(\frac{1}{2}\right)^i$$

$$T \stackrel{\text{def}}{=} [p_{i,j}]_{i,j \in \left[1 \ldots \left(\frac{N_0}{2}+1\right)\right]}$$

The first row represents the end of the simulation (state thereafter denoted by 'A') that comes when the agents that replied 'Yes' are more than half of the total number of agents, in other words, when $0 \leq N_t < \frac{N_0}{2}$. This implies that the first row has a 1 in the first position and 0 in all the other entries, and that T has dimension $\left(\frac{N_0}{2} + 2\right) \times \left(\frac{N_0}{2} + 2\right)$, because in the original $(N_0 + 1) \times (N_0 + 1)$ matrix one has to replace $\frac{N_0}{2}$ states with 1 state, so

$$(N_0 + 1) - \frac{N_0}{2} + 1 = \frac{N_0}{2} + 2$$

Moreover, T is lower triangular. For example, for $N_0 = 10$ we have:

$$T = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ p_{5,A} & p_{5,5} & 0 & 0 & 0 & 0 & 0 \\ p_{6,A} & p_{6,5} & p_{6,6} & 0 & 0 & 0 & 0 \\ p_{7,A} & p_{7,5} & p_{7,6} & p_{7,7} & 0 & 0 & 0 \\ p_{8,A} & p_{8,5} & p_{8,6} & p_{8,7} & p_{8,8} & 0 & 0 \\ p_{9,A} & p_{9,5} & p_{9,6} & p_{9,7} & p_{9,8} & p_{9,9} & 0 \\ p_{10,A} & p_{10,5} & p_{10,6} & p_{10,7} & p_{10,8} & p_{10,9} & p_{10,10} \end{pmatrix}$$

Where $p_{i,A} = \sum_{j=0}^{4} p_{i,j}$ for the law of total probability. The state A represents an absorbing state, that is a state from which it is not possible to get to any other. This absorbing state is the state to which the process converges. To see why this is the case it is sufficient to say that:

1. A is the only absorbing state.
2. A is reachable from every state in a finite number of steps.

This means that there is no probability of getting out of A once it has been reached, no other state that share this property, and there is always a strictly positive probability of getting to A from every other state. These conditions are satisfied by construction; therefore, the process will be absorbed by A. In this way, one can be sure that the model do not take an infinite amount of time.

Given the convergency of the process, it is useful to study the expected time to absorption, that is, how long it takes on average for the simulation to come to an end.

4.4 Expected time to absorption

Now, to determine the expected time to absorption one needs to find the average path length from state N_0 to the absorbing state. Denoting by τ the time to absorption, and by $\mu_i = E[\tau|N_t = i]$ the expected time to absorption starting from state i, one can write:

$$\mu_i = E[\tau|N_t = i] = 1 + \sum_j P(N_{t+1} = j|N_t = i)E[\tau|N_{t+1} = j]$$

In fact, thanks to the total expectation theorem, it is possible to decompose the expected value into the conditional expected values under all possible scenarios weighted according to the conditional probabilities of these scenarios. Therefore, the average path length from state i to A is 1 plus the sum of the average path lengths to A from j (the other states reachable from i, which are all $\frac{N_0}{2} \leq j \leq i$), weighted according to the probabilities of getting from i to j ($p_{i,j}$). Note that it is necessary to add 1 because for any $i \geq \frac{N_0}{2}$ it is necessary to make at least 1 step to get to A, or – in other words – it is always necessary to make 1 step to get from N_t to N_{t+1} if one does not start in A. Consistently, if one does start in a state i, $\mu_i = 0$ for all i. Therefore,

$$\mu_i = 1 + \sum_{j=\frac{N_0}{2}}^{i} p_{i,j}\mu_j = 1 + \sum_{j=\frac{N_0}{2}}^{i} \binom{i}{j}\left(\frac{1}{2}\right)^i \mu_j = 1 + \left(\frac{1}{2}\right)^i \sum_{j=\frac{N_0}{2}}^{i} \binom{i}{j} \mu_j$$

So, by solving this linear system one can get the expected number of steps of our model, μ_{N_0}.

Note that in the above equation j starts from $\frac{N_0}{2}$, this means that $j \neq A$, therefore we are picking a submatrix of T, that we will denote by L, which is equivalent to T but without the first column and the first row. Thus, L has dimension $\left(\frac{N_0}{2} + 1\right) \times \left(\frac{N_0}{2} + 1\right)$. In the previous example where $N_0 = 10$, L is the following:

$$L = \begin{pmatrix} p_{5,5} & 0 & 0 & 0 & 0 & 0 \\ p_{6,5} & p_{6,6} & 0 & 0 & 0 & 0 \\ p_{7,5} & p_{7,6} & p_{7,7} & 0 & 0 & 0 \\ p_{8,5} & p_{8,6} & p_{8,7} & p_{8,8} & 0 & 0 \\ p_{9,5} & p_{9,6} & p_{9,7} & p_{9,8} & p_{9,9} & 0 \\ p_{10,5} & p_{10,6} & p_{10,7} & p_{10,8} & p_{10,9} & p_{10,10} \end{pmatrix}$$

If we denote by **m** the column vector whose elements are μ_i for $i \geq \frac{N_0}{2}$ in increasing order, we can rewrite the above linear system in another fashion:

$$\boldsymbol{m} = \mathbf{1}_{\left(\frac{N_0}{2}+1\right)\times 1} + L\boldsymbol{m}$$

$$(\boldsymbol{I} - L)\boldsymbol{m} = \mathbf{1}_{\left(\frac{N_0}{2}+1\right)\times 1}$$

$$\boldsymbol{m} = (\boldsymbol{I} - L)^{-1}\mathbf{1}_{\left(\frac{N_0}{2}+1\right)\times 1}$$

Where $\mathbf{1}_{m\times n}$ denotes a matrix made of 1s with m rows and n columns, while $\boldsymbol{I}$ is the identity matrix. In this way, one has to pick the last entry of the vector **m** to get μ_{N_0}. Note that $(\boldsymbol{I} - L)$ is always invertible. In fact, a triangular matrix is invertible iff all its diagonal entries (which are also its eigenvalues) are non-zero, which is true by construction since for all i, $p_{ii} > 0$.

Given these properties, let us now dig into the code.

5. The code

5.1 Python libraries used and code introduction

There exist many languages to code ABMs, but here we decided to use Python 3.7 because Python is used in many other applications, and it would be interesting in the future to integrate AB modelling with other modelling techniques such as machine learning or AI. In particular, we used mesa, NumPy, pandas, SciPy and SymPy libraries at their latest version.

Mesa is a special library for agent-based models developed by researchers from George Mason's University. NumPy, SciPy and SymPy are useful libraries to deal with mathematics, and pandas is useful to handle datasets. Documentations are presented in the bibliography.

Mesa allows to create the agents' class, in this case Citizen(), and model class AgentBasedCV(). Both these class have their own step() function. After having initialised the model, agents are created, costs and values are assigned and the model starts through the SimultaneousActivation scheduler, which allows each citizen to make their move simultaneously, and then it confirms these moves through the advance() function. Each citizen's vote is set to False by default. When their step() function is called, the agents change it to True if the payoff is non-negative. Once citizens' choices are confirmed and collected by the model, it separates the positive voters from the negative voter and checks through the check_votes() function whether a majority has been reached. If not, the model continues by proposing new costs to the negative voters. When a majority is reached, the function get_mod_params() is called and a list of local parameters are stored in a pandas dataframe.

The built-in batchrunner, FixedBatchRunner, allows to run the model a fixed number of times with a specific value of N. In the case of the results shown in Table 2, the code iterates the model 50 times for each value of N in RANGE. It is possible to manually set RANGE, as well as the "iterations" argument in the FixedBatchRunner in order to test the simulation with a varying number of starting agents. After each iteration, the results from get_mod_params() are stored in a dataframe and means are computed over all iterations. Note that both Average Payoff and Average Cost parameters are actually averaged twice: the first time among the agents of the same iteration (it is the average parameter among the agents), and the second time among the iterations.

Providing a summary of the results produced:

- Table 2 contains the results produced from the model iterated 50 times for each value of N. These results correspond to Python Code #1, and the RANGE used is the one specified therein.
- Figure 7 contains the results produced from the model iterated 500 times. These results correspond to Python Code #1 and Python Code #2. Both have been run with same RANGE value specified in Python Code #2.

At the end, we store the results into an Excel table, one shown in the Appendix, the other under Figure 7 in section 4 of this chapter.

5.2 Python code #1

```
from mesa import Agent, Model
from mesa.time import SimultaneousActivation
from mesa.datacollection import DataCollector
from mesa.batchrunner import FixedBatchRunner
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt

A = 0 # Lower bound for C.U. distribution
B = 1 # Upper bound for C.U. distribution
RANGE = range(10, 100010+1, 1000) # Range of n, n is the starting number of citizens

# Here are some useful functions
def get_stepadvance(model):
    """ this is a counter for steps"""
    return model.stepadvance

def pick_from_unif(a,b):
    """ continuous uniform distribution """
    return ((b-a) * np.random.uniform(0,1) + a)

def get_votes(model):
    """ return a list with all the boolean votes """
    l = [a.vote for a in model.schedule.agents]
    return l

def print_steps(model):
    """ collect and print model-level data in markdown """
    model.datacollector.collect(model)
    d = model.datacollector.get_model_vars_dataframe()
    print('\n' + d.to_markdown(index=False) + '\n')
    print('Values:    ' + str(get_values(model)))
    print('Costs:     ' + str(get_costs(model)))
    print('Payoffs: ' + str(get_payoffs(model)))
    print('Votes in favor: ' + str(model.votes_in_favor) + '\n')
    print('Stepadvance: '+ str(model.stepadvance))


# Here we have get_params functions
def get_sum_votes(model):
    """ return the number of True values in votes' list """
    return sum(get_votes(model))

def get_costs(model):
    """ return a list with all the costs """
    l = [round(a.cost, 2) for a in model.schedule.agents]
    return l

def get_values(model):
    """ return a list with all the private values """
    l = [round(a.private_value, 2) for a in model.schedule.agents]
    return l

def get_payoffs(model):
    """ return a list with all the payoffs """
    l = [round(a.payoff, 2) for a in model.schedule.agents]
    return l
```

```
# Check votes and produce lists
def check_votes(model):
    """ Check if there is a majority, if so, the model stops running and parameters
         are collected thorugh the calling of the function get_mod_params() """
    if model.votes_in_favor > model.num_citizens/2:
        model.running = False
        get_mod_params(model)
    return

def get_mod_params(model):
    """ Get called at the end. Returns the list:
         [Social Payoff, Avg Payoff, Social Value, Avg Cost, Steps].
         Two temporary lists are created for costs and payoffs. First, winners'
         payoffs and costs are appended, then the average cost of winners is
         computed (this will be the cost paid by those who voted against in the
         last step). Then, payoffs and costs of the other citizens are appended and
         the five quantities are rounded and returned."""
    l = [] #payoffs
    g = [] #costs

    """ Append winners' payoffs and costs in list l and g respectively """
    for w in model.winners:
        l.append(round(w.payoff,2))
        g.append(w.cost)
    """ Get the mean cost of these winners """
    avg_wcost =  round(np.mean(g),2)
    """ Append the losers' payoffs which are equal to (value - avg_wcost) """
    for a in model.schedule.agents:
        l.append(round(a.private_value - avg_wcost,2))
        g.append(avg_wcost)
    return [round(sum(l),2), round(np.mean(l),2), round(sum(g),2), round(np.mean(g),2),
model.stepadvance]

# This instead is the Citizen class

class Citizen(Agent):

    def __init__(self, unique_id, model):
        """ Here the Citizen class is initialized. A private value and a
             cost are generated from a continous uniform distribution
             ranging from A to B. The vote is set to False initially. """
        super().__init__(unique_id, model)
        self.private_value = pick_from_unif(A,B)
        self.cost = pick_from_unif(A,B)
        self.vote = False

    def advance(self):
        """ Confirms what has been done in step(), after every citizen has voted """
        pass

    def step(self):
        """ Each citizen check if its payoff is non-negative and vote consequently
             If the payoff is negative, another cost is picked to be used in the step after"""
        self.payoff = self.private_value - self.cost
        if self.payoff >= 0:
            self.vote = True
        else:
            self.cost = pick_from_unif(A, B)
```

```python
# The model

class AgentBasedCV(Model):

    def __init__(self, n):
        """ A list of winners is initialized so that it is possible to track them.
             The schedule is the way in which the model proceeds forward in time, it
             is set as SimultaneousActivation, which means that all the agents will perform
             their step() simultaneously, and then they will simultaneously confirm through
             the advance() function the changes that they have applied.

             At the beginning, we get the votes in favor because it is possible that as
             soon as the agents are instantiated, they already have positive payoffs.
             A model reporter is also present to track of the dynamics of the model."""
        self.num_citizens = n
        self.winners = [] # winners are those who vote 'yes'
        self.schedule = SimultaneousActivation(self)
        self.stepadvance = 0

        for i in range(self.num_citizens): # create citizens
            a = Citizen(i, self)
            self.schedule.add(a)
        self.votes_in_favor = get_sum_votes(self)
        self.running = True
        self.datacollector = DataCollector(
            model_reporters={"Step": get_stepadvance, "Votes": get_votes})



    def step(self):
        """ Here, a step in the schedule means that the schedule is activated and
             runs every agent's step(). The stepsdvance tracks the step of the model
             instead of the step of the agents. Votes in favor are collected after the
             agents have done their steps, and variables are collected.
             Later, votes are checked, and winners are added to the winners' list."""
        self.schedule.step()
        self.stepadvance += 1
        self.votes_in_favor += get_sum_votes(self)

        """ exclude those who voted 'yes' """
        for a in self.schedule.agents:
            if a.vote:
                self.winners.append(a)
                self.schedule.remove(a)
        check_votes(self)
        return
```

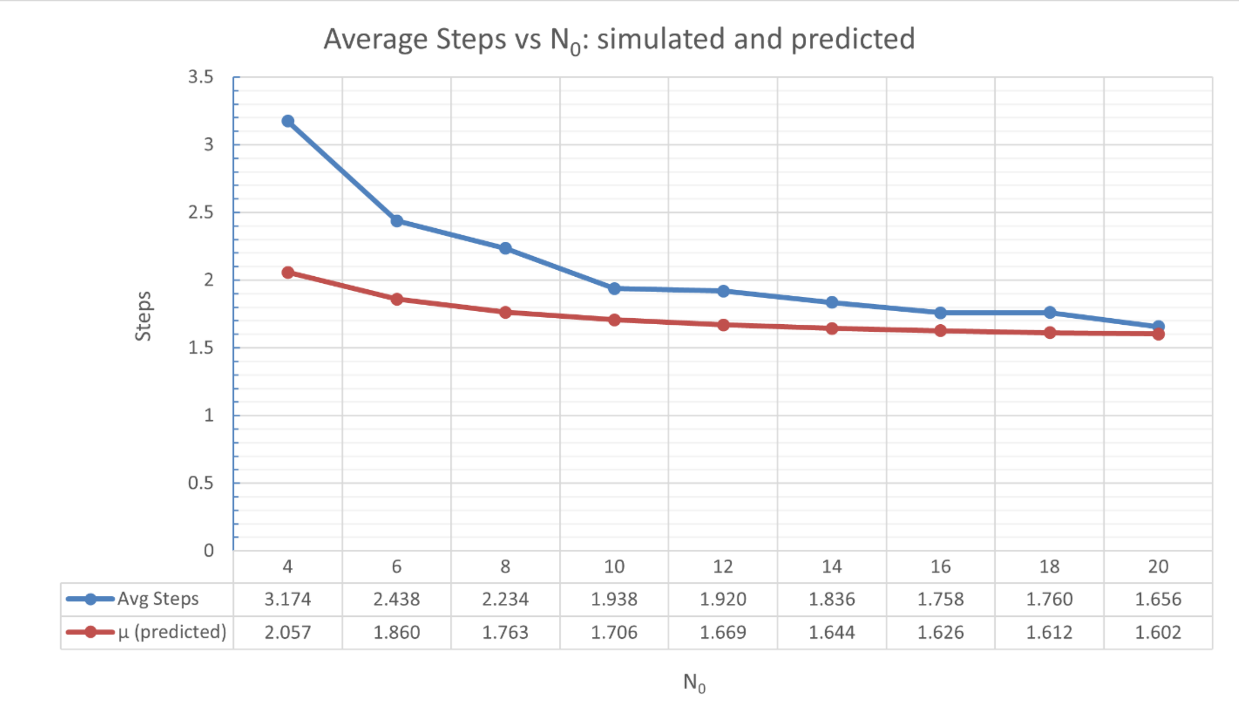
```
# Here many simulations are carried out so, while changing the initial number of agents. Data
are collected in pandas
# dataframes, and averages are computed. The results are stored into an .xlsx file.

means_collected = []

for n in RANGE:
    fixed_params = {"n": n}
    batch_run = FixedBatchRunner(AgentBasedCV, fixed_parameters=fixed_params, iterations=50,
max_steps=1000000, model_reporters={"All": get_mod_params})

    batch_run.run_all()

    data_collected = batch_run.get_model_vars_dataframe()
    data_collected = pd.DataFrame(a for a in data_collected['All'])
    data_collected.columns = ['Social Payoff', 'Avg Payoff', 'Public Value', 'Avg
Cost','Steps']

    mean0 = round(data_collected['Social Payoff'].mean(),3)
    mean1 = round(data_collected['Avg Payoff'].mean(),3)
    mean2 = round(data_collected['Public Value'].mean(),3)
    mean3 = round(data_collected['Avg Cost'].mean(),3)
    mean4 = round(data_collected['Steps'].mean(),3)

    means_collected.append([mean0, mean1, mean2, mean3, mean4])

means_collected = pd.DataFrame(means_collected)
means_collected.insert(0, 'N', pd.Series(RANGE))
means_collected.columns = ['N', 'Social Payoff', 'Avg Payoff', 'Public Value', 'Avg
Cost','Steps']
writer = pd.ExcelWriter('Final_Data.xlsx')
means_collected.to_excel(writer, index = False, header=True)
writer.save()
print(means_collected)
```

5.4 Python code #2

```
import numpy as np
from sympy import *
import scipy.special as sp
from fractions import Fraction
import pandas as pd

        """ The table shown in Figure 7 contains the results of
             both Python Code #1 and Python Code #2
             run with the same RANGE shown here."""

RANGE = range(4, 20+1, 2)

def entry(n, k):
    """ This funcion computes p(i,j), the elements of the matrix L. It returns a fraction."""
    return Fraction(sp.comb(n, k, exact=True), 2**(n))



def evaluate_mu(N):
    """ This is a function that solves the linear system presented in chapter 4.4.
         The only parameter of interest will be the last entry of the vector m.
         Here thanks to Fraction module and SymPy it is possible to compute matrices
         with fractions, they are approximated as floats only at the end.
         It is not advisable to run this function with N>30, given the computation required."""

    dim = int(N/2) + 1
    L = np.full((dim, dim), 0)
    L = L.astype('object')

    for j in range(dim-1, N+1 , 1):
        for i in range(j, N+1 , 1):
            r = i - (dim - 1)
            c = j - (dim - 1)
            L[r][c] = entry(i,j)

    L = Matrix(L)
    m = ((eye(dim) - L).inv()) * ones(dim,1)
    for cell in m:
        cell = float(cell)
    return float(m[-1])


""" Here the expected time to absorption is computed according to the linear system shown
    in chapter 4.4."""

mu = []
for N in RANGE:
    mu.append(evaluate_mu(N))

mu_data = pd.DataFrame(mu)
mu_data.insert(0, 'N', pd.Series(RANGE))
mu_data.columns = ['N', 'mu']
writer = pd.ExcelWriter('mu_data.xlsx')
mu_data.to_excel(writer, index = False, header=True)
writer.save()
```

5.4 Explanation of the results

As expected from the linear system presented in chapter 4 section 4, Python Code #2 produces number of steps required to end the model follows the expected number time to absorption (μ_{N_0}), showing more variance when citizens are few.



N_0	4	6	8	10	12	14	16	18	20
Avg Steps	3.174	2.438	2.234	1.938	1.920	1.836	1.758	1.760	1.656
μ (predicted)	2.057	1.860	1.763	1.706	1.669	1.644	1.626	1.612	1.602

Figure 4: Data is the output of Python Code #1 and Python Code #2. The RANGE used for both codes is the one indicated in Python Code #2. Avg Steps values are the average over 500 iterations for each value of N shown above.

As one can see, the average number of steps do not increase with N; on the contrary, it decreases until reaching a stable value at 1.5 average steps. This is because at each step half of the agents on average will vote positively. In other words, the ratio between negative and positive votes will be increasingly closer to 1 as the number of agents increases. This essentially confirms that the model scales efficiently relative to an increase in the initial number of agents.

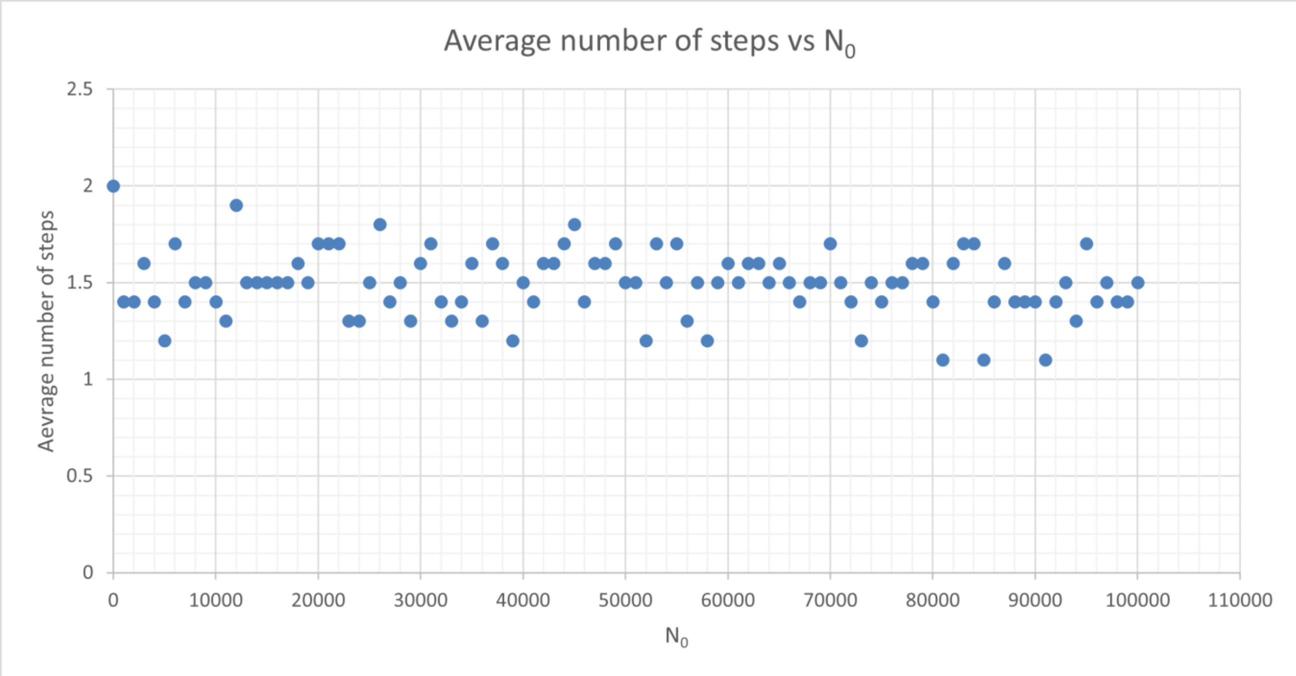


Figure 5: Average number of steps vs N_0. Data comes from Table 2.

Plus, also the individual payoff (the average difference between costs and values after all citizens have paid) shows the same promising robustness as plotted below, oscillating around 0.178. Its dactylographic appearance is due to the rounding up to 3 decimal digits.

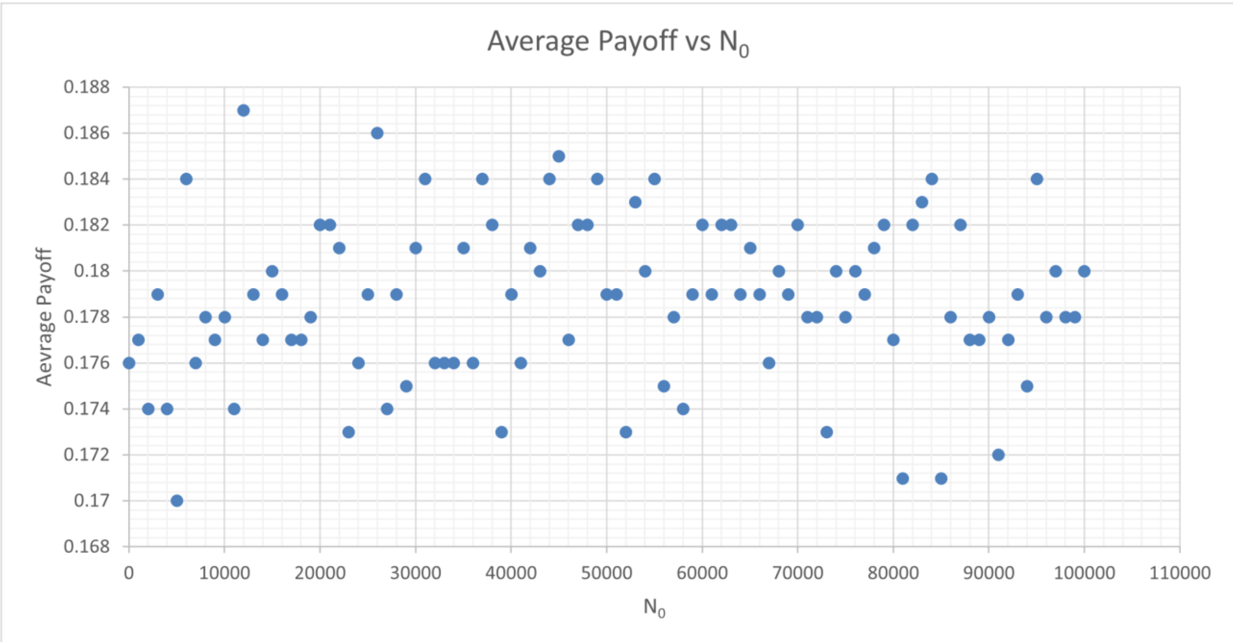


Figure 6: Average Payoff vs N_0. Data comes from Table 2.

It is therefore evident that the average payoff is a robust result with respect to the initial number of agents. What about costs then? Indeed, also the average cost per capita shows a similar robustness, converging to approximately 0.32. This robustness also implies that the Public Value increases linearly with N, confirming the intuitive idea that the more people make use of a good, the more it should be valued.

Note also that the data used in these plots of Figures 5, 6 and 7 come from the same simulation. All the figures in this section have been produced in Microsoft Excel for a better readability.

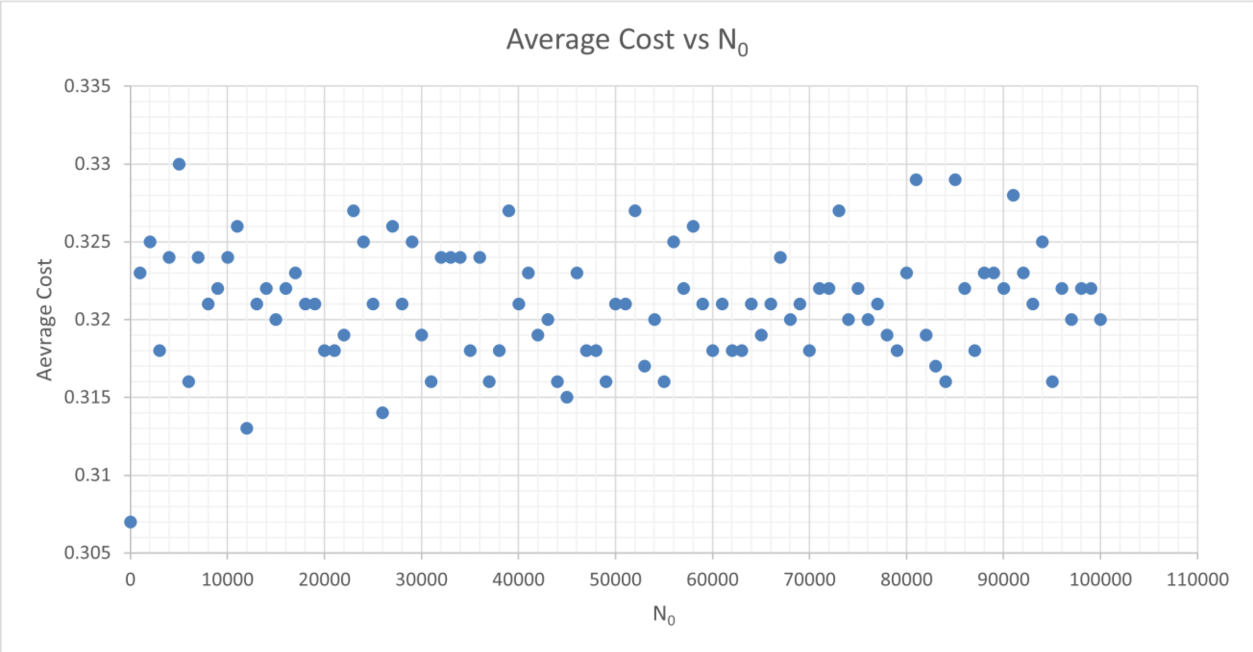


Figure 7: Average Cost vs N_0. Data comes from Table 2.

5. Concluding remarks

5.1 Robustness of the results

The model presented in this work shows robust results. It is possible to state that the average payoff of an agent is approximately 0.178, which is almost a half of the average cost paid by the individual, 0.32. This means that the model, under its assumptions, could be a good compromise between public and private, since it does leave room to private exploitation but at the same time it does represent the value of the good (defined to be the sum of all private values) quite accurately.

In fact, if the assumptions of the model hold, it could be used as a quantitative alternative to the more traditional approaches to contingent valuation that have been used since the present days, especially given the promising results of the simulations. More empirical testing is therefore needed in order to assess how a real sample will behave when involved into the survey.

5.2 Limitations and suggestions for further developments

5.2.1 From theory to practice

Although the model shows consistent and robust results, it may suffer of two structural limitations that could make it problematic if applied to real circumstances.

The first of such limitations is the impossibility for the agents not to take part to the valuation. In fact, it is implicitly assumed that all the agents will pay a certain amount. If not, agents have no personal incentive in being honest, in the same way respondents to traditional open-question contingent valuations could just make up a number if they are not interested. So, the assumption that agents should pay to make results more consistent poses concerns about when it is legitimate to impose a cost. Naturally, if the cost is to be imposed anyway (like in the case of taxes or social contributions) it is presumably better to let people express their preferences on the cost to pay. Otherwise, one has to accept the reduced reliability of the outcome and proceed with purely hypothetical valuations in the standard way.

The second limitation of the model has to do with the proportionality of the costs imposed. This is especially crucial in policy applications, where income inequality is an important ground on which social contributions should be evaluated. Indeed, this limitation is not difficult to overcome, as it is possible to change the upper and lower bound of the cost distribution as functions of personal income. But the problem then shifts to solving the question "how much should the individual income matter?". The more room is left to randomness, the more potentially unequal is the outcome. While the more the income conditions restrict the boundaries of the cost distribution, the less decisional power is left to the citizen.

In summary, this model is to be considered as a thought-experiment that still needs to be studied under many other aspects.

5.2.2 Suggestions for further developments

In particular, we encourage a further exploration in the following directions:

- It would be interesting to set up a comparison between this method and traditional contingent valuations in the outcome produced, in order to see the differences in how these two techniques value the same public good.
- It would be important to give more grounding to the behavioural assumptions of the model, that is, how far-fetched is the idea that agents reply according to their true private value. It is in fact plausible that the simulation over-estimates the number of votes in favour for each turn, maybe because in reality people could be more cost averse.
- Moreover, in order to make the model useful in public policy, it should be tested with different assumptions on the distributions of costs. For example, one might introduce lower bounds or upper bounds that are proportional to individual income. Perhaps, it will be possible to discover that private values are influenced to some degree by a change in the boundaries of the cost distribution.

In conclusion, there are always new ways to improve existing methodologies and we hope to have provided in this work a potential solution to a common problem. We likewise hope to provide in the future other potential solutions to the great number of challenges that are still waiting to be overcome.

6. Appendix

Table 2. Results corresponding to Final_Data.xlsx

N_0	Social Payoff	Avg Payoff	Public Value	Avg Cost	Avg Steps
10	1.741	0.176	3.076	0.307	2
1010	179.971	0.177	325.716	0.323	1.4
2010	351.161	0.174	653.293	0.325	1.4
3010	542.123	0.179	956.54	0.318	1.6
4010	697.417	0.174	1299.283	0.324	1.4
5010	858.513	0.17	1651.081	0.33	1.2
6010	1101.567	0.184	1898.746	0.316	1.7
7010	1235.084	0.176	2274.532	0.324	1.4
8010	1413.311	0.178	2574.611	0.321	1.5
9010	1583.151	0.177	2905.47	0.322	1.5
10010	1765.526	0.178	3244.652	0.324	1.4
11010	1912.896	0.174	3595.572	0.326	1.3
12010	2219.676	0.187	3770.719	0.313	1.9
13010	2301.051	0.179	4194.496	0.321	1.5
14010	2483.223	0.177	4512.102	0.322	1.5
15010	2679.928	0.18	4816.563	0.32	1.5
16010	2837.723	0.179	5161.976	0.322	1.5
17010	3020.618	0.177	5496.569	0.323	1.5
18010	3216.068	0.177	5786.094	0.321	1.6
19010	3373.523	0.178	6121.393	0.321	1.5
20010	3640.362	0.182	6370.459	0.318	1.7
21010	3816.345	0.182	6697.196	0.318	1.7
22010	3974.039	0.181	7030.629	0.319	1.7
23010	3974.733	0.173	7540.055	0.327	1.3
24010	4199.316	0.176	7810.31	0.325	1.3
25010	4466.598	0.179	8054.195	0.321	1.5
26010	4799.109	0.186	8199.168	0.314	1.8
27010	4699.118	0.174	8805.383	0.326	1.4
28010	4978.866	0.179	9021.481	0.321	1.5
29010	5060.676	0.175	9447.594	0.325	1.3
30010	5396.973	0.181	9602.001	0.319	1.6
31010	5665.975	0.184	9836.647	0.316	1.7
32010	5641.544	0.176	10390.456	0.324	1.4
33010	5756.948	0.176	10738.81	0.324	1.3
34010	5966.074	0.176	11035.841	0.324	1.4
35010	6311.581	0.181	11192.675	0.318	1.6
36010	6275.196	0.176	11710.616	0.324	1.3
37010	6743.11	0.184	11759.316	0.316	1.7
38010	6848.79	0.182	12153.022	0.318	1.6
39010	6715.282	0.173	12801.096	0.327	1.2
40010	7100.437	0.179	12889.887	0.321	1.5
41010	7194.51	0.176	13296.424	0.323	1.4
42010	7594.035	0.181	13454.442	0.319	1.6
43010	7698.716	0.18	13785.486	0.32	1.6
44010	8028.213	0.184	13967.417	0.316	1.7
45010	8272.323	0.185	14208.882	0.315	1.8
46010	8087.512	0.177	14909.149	0.323	1.4
47010	8489.337	0.182	15004.463	0.318	1.6
48010	8664.79	0.182	15336.382	0.318	1.6
49010	8930.171	0.184	15547.275	0.316	1.7
50010	8881.462	0.179	16115.477	0.321	1.5
51010	9068.901	0.179	16435.465	0.321	1.5
52010	8972.356	0.173	17060.331	0.327	1.2
53010	9630.952	0.183	16872.107	0.317	1.7

N_0	Social Payoff	Avg Payoff	Public Value	Avg Cost	Avg Steps
54010	9627.985	0.18	17359.701	0.32	1.5
55010	10027.058	0.184	17468.847	0.316	1.7
56010	9777.764	0.175	18275.454	0.325	1.3
57010	10105.923	0.178	18396.144	0.322	1.5
58010	10025.183	0.174	18991.333	0.326	1.2
59010	10500.642	0.179	18978.974	0.321	1.5
60010	10816.87	0.182	19166.258	0.318	1.6
61010	10872.306	0.179	19634.918	0.321	1.5
62010	11190.187	0.182	19830.959	0.318	1.6
63010	11377.911	0.182	20140.62	0.318	1.6
64010	11371.333	0.179	20604.903	0.321	1.5
65010	11681.013	0.181	20797.784	0.319	1.6
66010	11700.377	0.179	21277.509	0.321	1.5
67010	11724.699	0.176	21782.876	0.324	1.4
68010	12111.718	0.18	21872.228	0.32	1.5
69010	12297.012	0.179	22237.076	0.321	1.5
70010	12648.687	0.182	22307.871	0.318	1.7
71010	12608.333	0.178	22903.726	0.322	1.5
72010	12709.736	0.178	23296.249	0.322	1.4
73010	12564.152	0.173	23952.561	0.327	1.2
74010	13173.759	0.18	23808.206	0.32	1.5
75010	13217.087	0.178	24292.18	0.322	1.4
76010	13518.099	0.18	24430.025	0.32	1.5
77010	13666.353	0.179	24807.875	0.321	1.5
78010	13998.736	0.181	25000.409	0.319	1.6
79010	14232.146	0.182	25231.122	0.318	1.6
80010	14068.238	0.177	25949.224	0.323	1.4
81010	13800.131	0.171	26713.616	0.329	1.1
82010	14774.792	0.182	26253.716	0.319	1.6
83010	15078.11	0.183	26395.31	0.317	1.7
84010	15324.386	0.184	26659.794	0.316	1.7
85010	14443.234	0.171	28069.579	0.329	1.1
86010	15163.48	0.178	27806.855	0.322	1.4
87010	15665.558	0.182	27806.666	0.318	1.6
88010	15509.684	0.177	28525.623	0.323	1.4
89010	15652.374	0.177	28890.2	0.323	1.4
90010	15862.768	0.178	29136.327	0.322	1.4
91010	15591.792	0.172	29986.33	0.328	1.1
92010	16184.412	0.177	29850.024	0.323	1.4
93010	16503.801	0.179	29953.472	0.321	1.5
94010	16393.403	0.175	30670.022	0.325	1.3
95010	17284.413	0.184	30178.447	0.316	1.7
96010	16970.649	0.178	31042.069	0.322	1.4
97010	17307.735	0.18	31193.453	0.32	1.5
98010	17293.122	0.178	31751.24	0.322	1.4
99010	17469.323	0.178	32046.937	0.322	1.4
100010	17790.542	0.18	32202.932	0.32	1.5

Table 2: Results from model simulations with varying number of initial agents. Each of these values, except N_0, is an average over 50 iterations.

7. Bibliography

1. Arrow K, Solow R, et al. (1993). "Report of the NOAA panel on contingent valuation". https://edisciplinas.usp.br/pluginfile.php/4473366/mod_folder/intro/Arow_WTP.pdf. Retrieved Sep 2021.
2. Axelrod R, Ford GR, Riolo RL, Cohen MD. (2002). "Beyond geography: cooperation with persistent links in the absence of clustered neighborhoods". Pers Soc Psychol Rev, 6:341–346.
3. Axelrod R. (1997). "The Complexity of Cooperation: Agent-based Models of Competition and Collaboration". Princeton, NJ: Princeton University Press.
4. Axtell R. (2011). "Modeling the Economy with 150 Million Agents". TEDxUVM. https://www.youtube.com/watch?v=V_sGz4FcuPA&ab_channel=TEDxTalks. Retrieved Sep 2021.
5. Baumgaertner B. (2018). "Models of Opinion Dynamics and Mill-Style Arguments for Opinion Diversity". Historical Social Research 43 (1): 210-33.
6. Boero R, Bravo G, Castellani M, Squazzoni F. (2010). "Why bother with what others tell you? An experimental data-driven agent-based model". J Artif Soc Soc Simul, 13(3):6.
7. Boltanski L, Esquerre A, and Muniesa F. (2015). "Grappling with the economy of enrichment". Valuation Studies, vol. 3, no. 1, 75–83.
8. Bonabeau E. (2002). "Agent-based modeling: Methods and techniques for simulating human systems". Proceedings of the National Academy of Sciences 99 (suppl 3): 7280-7.
9. Carson R T, Mitchell R C, et al. (1992). "A Contingent Valuation Study of Lost Passive Use Values Resulting from the Exxon Valdez Oil Spill;" A Report for the Attorney General of the State of Alaska.
10. Carson R T. (2000). "Contingent valuation: a User's Guide". Environmental Science & Technology 34 (8), 1413-1418.
11. Ciriacy-Wantrup, S. V. (1947). "Capital Returns from Soil-Conservation Practices. Journal of Farm Economics". 29(4), 1181–1196.
12. Coleman J S. (1964). "Introduction to Mathematical Sociology". New York: MacMillan.
13. Coleman J S. (1964). "Mathematical models and computer simulation". *Handbook of Modern Sociology*. Chicago IL: Rand McNally & Co.
14. Deissenberg C, van der Hoog S, Dawid H. (2008). "EURACE: A massively parallel agent-based model of the European economy". Applied Mathematics and Computation. 2008;204(2):541-552.
15. Dütting P, Feng Z, et al. (2020). "Optimal Auctions through Deep Learning". arXiv e-prints: arXiv:1706.03459v5. Retrieved Sep 2021.
16. Epstein J M, Axtell R. (1996). "Growing Artificial Societies: Social Science from the Bottom Up". Washington, DC: Brookings Institution Press.
17. Epstein J M. (2006). "Generative Social Science: Studies in Agent-Based Computational Modeling". Princeton, NJ: Princeton University Press.
18. Games, M. (1970). "The fantastic combinations of John Conway's new solitaire game "life" by Martin Gardner". Scientific American, 223, 120–123.

19. Halvorsen B, Sœlensminde K. (1998). “Differences between Willingness-to-Pay Estimates from Open-Ended and Discrete-Choice Contingent Valuation Methods: The Effects of Heteroscedasticity”. Land Economics, 74(2), 262–282.
20. Hammond R A, Axelrod R. (2006). “Evolution of contingent altruism when cooperation is expensive”. *Theor Popul Biol*, 69:333–338.
21. Hayek F A. (1945). “The Use of Knowledge in Society.” American Economic Review 35, no. 4: 519–530.
22. Hegselmann R, Krause U. (2002). “Opinion dynamics and bounded confidence: models, analysis and simulation.”. Journal of Artificial Societies and Social Simulation vol. 5, no. 3. https://www.jasss.org/5/3/2.html. Retrieved Sep 2021.
23. Hinch R, Probert W J M, Nurtay A, Kendall M, et al. (2020). “OpenABM-Covid19 - an agent-based model for non-pharmaceutical interventions against COVID-19 including contact tracing”. medRxiv: 2020.09.16.20195925.
24. Jevons W S. (ed. 1888). The Theory of Political Economy. 3rd ed., London: Macmillan.
25. Klein D, Marx J, & Fischbach K. (2018). “Agent-Based Modeling in Social Science, History, and Philosophy: An Introduction”. Historical Social Research, 43(1), 7-27.
26. Lindahl E. (ed. 1958). “Die Gerechtigkeit der Besteuering. Lund: Gleerup”. reprinted as “Just taxation – a positive solution”, in Classics in the Theory of Public Finance, ed. R.A. Musgrave and A.T. Peacock, London: Macmillan.
27. Menger C. (ed. 1994). “Principles of Economics”. Grove City, Penn.: Libertarian Press.
28. Mesa documentation: https://mesa.readthedocs.io/en/master/overview.html. Retrieved Sep 2021.
29. Mises L. (ed. 2003). “Epistemological Problems of Economics”, 3rd ed. Ludwig von Mises Institute.
30. Nagel K, Rasmussen S. (1995). “Traffic at the Edge of Chaos”. arXiv e-prints: arXiv:adap-org/9502005v1. Retrieved Sep 2021.
31. NumPy documentation: https://numpy.org/doc/. Retrieved Sep 2021.
32. Palmer R G, Arthur WB, et al. (1999). “An artificial stock market”. Artificial Life Robotics. 3:27-31.
33. Pandas documentation: https://pandas.pydata.org/docs/. Retrieved Sep 2021.
34. Randall A, Ives B, Eastman C. (1974). “Bidding games for valuation of aesthetic environmental improvements". Journal of Environmental Economics and Management. Elsevier. vol. 1(2), 132-149.
35. Sagoff M. (1998). “Aggregation and deliberation in valuing environmental public goods:: A look beyond contingent pricing”. Ecological Economics, vol. 24. Issue 2-3: 213-230. Elsevier.
36. Scheller, Simon. (2018). “When Do Groups Get It Right?”. Historical Social Research 43 (1): 89-109.
37. Schelling T C. (1971). “Dynamic Models of Segregation”. *J Math Sociol*. ch. 1.
38. Schmitt J, Franzmann ST. (2018). “A Polarizing Dynamic by Center Cabinets? The Mechanism of Limited Contestation”. Historical Social Research 43 (1): 168-209.
39. Smith, A. (2002). “The Wealth of Nations”. Oxford, England: Bibliomania.com Ltd. [Web.] Retrieved from the Library of Congress, https://lccn.loc.gov/2002564559, as of Sep 2021.

40. Sraffa P. (1960). “Production of Commodities by Means of Commodities - Prelude to a Critic of Economic Theory”. Cambridge: Cambridge University Press.
41. Stegmüller W. (1983). Probleme und Resultate der Wissenschaftstheorie und analytischen Philosophie. 1, C: Erklärung, Begründung, Kausalität: Historische, psychologische und rationale Erklärung. Verstehendes Erklären. Berlin, Heidelberg, New York: Springer.
42. Sung-Yoon Huh, et al. (2015). “The economic value of South Korea's renewable energy policies (RPS, RFS, and RHO): A contingent valuation study”. Renewable and Sustainable Energy Reviews, vol. 50: 64-72.
43. SymPy documentation: https://docs.sympy.org/latest/index.html. Retrieved Sep 2021.
44. Tesfatsion L. (2017). “Modeling Economic Systems as Locally-Constructive Sequential Games”. Economics Working Paper 17022. *Journal of Economic Methodology* 24 (4): 384-409.
45. Tsoulfidis L. (1998). “Ricardo’s theory of value and Marx’s critique”. History of Economic Ideas, 6(2), 69–88.
46. Walras, Léon. (ed. 1952). “Eléments d’Economie Politique Pure ou Théorie de la Richesse Sociale”. Paris: Librairie Générale du Droit et de Jurisprudence.
47. Weisberg M, Muldoon R. (2009). “Epistemic Landscapes and the Division of Cognitive Labor”. Philosophy of Science 76:2, 225-252.
48. Willis K G, Garrod G D, Harvey D R. (1998). “A review of cost–benefit analysis as applied to the evaluation of new road proposals in the U.K.”. Transportation Research Part D: Transport and Environment, vol. 3. Issue 3: 141-156.
49. Willis K G, Garrod GD. (1993). “Valuing Landscape: a Contingent Valuation Approach”. Journal of Environmental Management. vol. 1 (37). Issue 1, 1-22.