Chapter 23
Un Drôle De Type: The Schelling Model, Calibration, Specification, Validation and Using Relevant Data

Edmund Chattoe-Brown

Abstract This paper identifies a potential blind spot in ABM, linking aspects of methodology and data use. The relative neglect of "specification" (empirical justification of model components like particular agent decision processes) combined with a relative paucity of qualitative data in ABM draws attention away from the possibility that agents may make decisions in heterogeneous ways with uncertain implications for macroscopic system properties. Using the Schelling model as a simple and well known example, this paper considers the role of specification as complementary to calibration and validation, the way that different kinds of data (qualitative and quantitative) map on to different aspects of ABM methodology to justify a model empirically and the possible implications of systematically heterogeneous decision making. Some preliminary simulation results are presented and discussed for mixtures of heterogeneous decision "types" grounded in existing secondary data. The paper also considers exactly what the Schelling model can (and cannot) be taken to show and how the legitimacy of various claims made for it relate to its empirical justification.

Introduction

Famously, according to Maslow [1], to the man [sic] who has only a hammer, everything begins to look like a nail (though it is less well known that the earliest citable statement of this definite idea originated with Abraham Kaplan in 1964). One of the general advantages of ABM is that it does *not* require assumptions to be made for "technical" (i.e. not empirically justified or perhaps even justifiable) reasons as some other formal methods do. However, as I hope to show in this paper, using variants of the well known Schelling model [2], it is still unfortunately possible for ABM to mistake something else for a nail, although rather through inadvertence and lack of

E. Chattoe-Brown (✉)
School of Media, Communication and Sociology, University of Leicester, Leicester LE1 7JA, UK
e-mail: ecb18@le.ac.uk

P. Ahrweiler and M. Neumann (eds.), *Advances in Social Simulation*,
Springer Proceedings in Complexity,
https://doi.org/10.1007/978-3-030-61503-1_23

methodological care than the technical limitations of the approach itself. The argument proceeds as follows. The next section considers the relative neglect of model *specification* (empirical justification for claims about how particular elements of a social process come to be represented as they do in an ABM) in contrast to the more widely recognised concepts of *calibration* and *validation*. The third section considers how various data types relate to particular aspects of specifying the Schelling model (chosen as an example), namely how it is assumed that agents take moving decisions. The fourth section presents and analyses some preliminary simulation results on heterogeneous decision making. The fifth and final section concludes and discusses the possibilities for further research on heterogeneity involving more effective use of qualitative data.

The Assumptions You Don't Realise You Are Making

The first problem with discussing this issue is clearly one one of concepts and associated terminology. Leaving aside the contentious issue of what the range of legitimate aims for ABM are (and how we establish whether they have been successfully realised in a particular piece of research)—see for example [3, 4], let us suppose that we are interested in building an empirical ABM, one that is constructed using (and assessed in terms of) data. Although their names sometimes vary, two concepts (which I shall call calibration and validation following widespread but not universal usage) involve assigning numerical values to parameters arising in the construction of an ABM (calibration) and comparing simulated output with corresponding real data (validation). The Schelling model has been chosen for the argument here because it is widely known in the ABM community and simple enough to demonstrate relevant issues clearly without "bogging down" in potentially divisive modelling detail. (Another disadvantage of non-empirical ABM is that arguments about the relative plausibility of variant assumptions are typically unproductive.) The argument presented here does not hinge on whether the Schelling model is "legitimate" or "valid" or exactly what it "proves" (though it does potentially cast additional light on such matters as I shall try to show shortly).

In this context, then, calibration involves assigning numerical values (a single value for all agents, different constant values for different agent types, different ranges for different agent types and so on) to the value of PP (the Preferred Proportion of one's own kind in the immediate neighbourhood) and validation involves choosing a measure of clustering according to which corresponding real and simulated data might be compared. This measure can be very qualitative ("there are some clusters") or rigorously quantitative (taking account of the advantages and pitfalls of different formal measures of clustering [5]) without affecting the core argument of this paper.

However, for my purposes, what is important here is the related concept which (while obvious *ex post* and probably "unavoidable" to effective empirical modelling) is *not* commonly given a name (and certainly not a standard name). Leaving any other assumptions of the Schelling model aside for a moment (although in fact we will

rapidly return to them), we have to ask on what *grounds* Schelling decided that agents would make the decision in the way he assumes they do and, in particular, why they would all make that decision in the "same" way—leaving aside any variation in PP which has been added to subsequent variant models. (We could ask on what *empirical* grounds but that would be unfair to Schelling's stated intention—despite being a reasonable question generally—but even making the arguments Schelling wished to make, there may be more or less legitimate grounds for specifying a particular decision process. For example, it would not be ideal in scientific terms if he had discarded decision processes without acknowledgement until he found one that would serve the theoretical or policy point he wished to make! There seems to be a belief in the ABM community that if you decide not to build an empirical model you are not obliged to provide *any* objective grounds for saying that you have achieved your modelling goals. In extreme cases, modellers do not even make clear what these goals are which, of course, makes it very hard to fail in them!) Once this aspect of the model specification (what kind of decision process or processes agents should have) has been decided, the task of calibration is relatively banal in that it is obvious what parameters need to be assigned values and (broadly) how those parameters should be interpreted. (Other interesting matters can arise here however. If specification is not carefully considered in an ABM then calibration may fail because the parameters of an arbitrarily conceived decision process turn out to have no empirical referents.) Without some decision process involving movement, an ABM would really no longer be an example in the class of Schelling models at all. But the question legitimately remains even though there may be a range of appropriate responses (depending on the purposes of the model for example). Why *that* decision process? (And why that decision process *for everyone?*)

Here, however, we divert briefly into the "meaning" and "value" of the Schelling model to support arguments that follow. Clearly, regardless of the empirical correspondence of either assigned PP values or the decision process, the Schelling model *does* provide a simple and accessible introduction to concepts like *non-linearity* (the "collapse" of increasingly rigorous clustering above a certain PP value), *emergence* (the properties of clusters themselves which are not "reducible" to the properties of agents) and *complexity* (the non intuitive relationship between individual agents and their interactions and the macroscopic properties of clustering)—see [6] for more detail on these ideas in the context of the Schelling model. This outcome does not require the Schelling model to be "true" in any empirical sense, only that the social story told about the model is not *so* implausible that the reader cannot engage effectively with it at all. The "conceptual" value of the model can also be translated into legitimate advice about real styles of social science reasoning (though not without some additional cautions). Under the set of assumptions that Schelling makes (including the specification of a homogeneous agent decision process), the model shows that clustering should not "automatically" be assumed to arise from the "obvious" causes of individual "xenophobia" (or perhaps less morally disagreeable self-preference). There will still be clusters even if people are quite tolerant given Schelling's assumptions (and some but perhaps not all other sets of assumptions he might have chosen instead). There can also be an *absence* of clusters if people are very

intolerant so that demonstrates that the naïve relationship assumed between levels of analysis is neither necessary *nor* sufficient. However, this finding probably has no bearing on what would happen in the real world if people made decisions differently (and in particular with more diversity) than Schelling asserts that they do. It also has no bearing on whether (in fact) intolerance or self preference *are* contributors to observed pattern of ethnic residential segregation (which we have solid empirical evidence to think that they may be—see, for example [7, 8]). Finally, we can be fairly clear that (except under a rather special and unlikely set of circumstances) the Schelling model does not "explain" the actual pattern of ethnic residential segregation in, say, a particular area of Chicago in 1965. The reason for this is that it is unlikely that, "from his armchair" as it were, Schelling has hit on the correct (or "processually dominant") decision process and been correct in his (implied) judgements that things like demography (specifically changes in family composition) or the housing market (for example) do not have a significant impact on real patterns of segregation. In fact, we have independent reasons [9] for thinking such things *do* matter. (It is an interesting question whether it would be "fair" to test the Schelling model empirically and how much it would really prove if it "failed" such a test. Schelling himself did not claim it was an empirical model, though he did use it to draw certain classes of conclusions which may or may not then be legitimate based on an abstract model of this kind. However, subsequently, data collection methods have been created to give at least a reasonable proxy for the PP parameter [10] thus making it testable at least "in principle". There have also been a fairly large number of ABM "extending" the Schelling model [11] without really addressing the issue of specification—and largely without validation either—thus implying that more complicated models are somehow more plausible—or perhaps just less implausible. This reminds me of the old joke about the farmer giving advice to the lost tourist: "If I were you I wouldn't start from here at all." If the problem is not the specific features of any variant model but their arbitrary nature, then simply having more arbitrary features is not likely to produce scientific progress but rather models that are increasingly difficult to interpret to no advantage).

Arguments about definitions can often bog down effective debate but it is clear that, whatever we choose to call it, the empirical justification for a particular specification of decision making in the Schelling model is an important matter, likely to affect both the behaviour of the model (and thus even its "abstract" conclusions about complexity and legitimate inference between micro and macro levels of analysis) and its ability to empirically "track" corresponding elements of the real world. This being so, I shall refer to the empirical justification of model elements (like the decision process) as *specification* to form a trinity with the better known and more widely recognised processes of calibration and validation. (I would be happy to change this term for a widely used pre-existing term but I can't find any evidence of one. Obviously, there is no particular reason for changing the arbitrary term I am using to an arbitrary term used by someone else unless it is already generally accepted.) Taken together, these three processes thus have the opportunity to offer empirical justification for all aspects of an ABM (barring one complication discussed in the next section).

But What About the Data?

On the whole, ABM seems far more interested in data in principle than in practice [12]. Given the general preponderance of non empirical models, there is a small but thriving community interested in the distinctive role of *qualitative* data (semi structured interviews, ethnography and so on) in ABM (see, for example, the set of articles discussed in [13]). From the more general social science perspective, however, the relations between different data types and ABM methodology actually seem relatively straightforward. The validation of the Schelling model involves large scale surveys of the ethnic composition of households in neighbourhoods that can be compared with the output of the model. (However, the lack of interest in data can still cause ABM to overlook important issues. Taking the Schelling model or some variant as potentially "true" for a moment, moving decisions would be determined by how neighbours *appear* in terms of the subjective ethnic categories of each decision maker. This is not the same—necessarily—as the self declared ethnicity that would be established by a large scale household survey.) By contrast, broadly speaking, how people *actually* make moving decisions (and how diverse these decisions are as a matter of fact) i.e. specification, seem likely to be established predominantly by *qualitative* methods [14]. This claim is clearly something of a simplification for the purposes of argument. Firstly, some motivations may be inaccessible to self report and/or extremely socially unacceptable (for example because they are illegal) and may thus require covert and unobtrusive methods [7, 8] or perhaps experiments (provided the broad shape of the decision process is already known so these can be designed effectively). Secondly, data about moving decisions is not *irreducibly* qualitative. Once one has a justified idea about *broadly* how people decide to move which has been established robustly from interviews, surveys can then be used to obtain more representative information about, for example, the relative importance of different aspects in the decision. (In fact, given the unsuitability of existing qualitative research on moving decisions, I am, in a manner of speaking, using this approach in reverse—suitable secondary survey data temporarily "standing in for" qualitative data—to justify the ABM presented here.) Finally, it should be noted that qualitative research for use in ABM may have particular requirements which pre-existing research on the same topic may not happen to satisfy [15]. For example, it may need to be more process based and more focused on the exact detail of decision mechanisms [14].

There simply seems to be very little qualitative research on house moving decisions generally and what there is does not seem very suitable to the specification of ABM (having been designed without reference to the strengths and requirements of the method). Pending the collection of more effective data in this area, however, it is possible to make the necessary arguments for this paper provisionally using secondary survey data. In a highly regarded study, which would have been available to Schelling, Rossi [9] asked people to report reasons for wanting to move. The answers given are not necessarily of direct help in formulating a well grounded ABM but they do give a more nuanced sense of what *might* matter in moving decisions and

thus whether these might be systematically heterogeneous. For example, the largest single response category (33%) concerns amount of closet (cupboard) space. This is a property of a *dwelling*, not a "neighbourhood" or wider geographical context and several other large response categories fall into this broad class: Open space around the house (28%), amount of room (22%), heating equipment (16%), amount of air and sunlight (14%) and amount of privacy (12%). Other reported reasons deal with the *neighbourhood* but in a physical, rather than social sense: Street noise (23%). The economic dimension is also relevant but not, contra economics as an academic discipline, dominant: Rent (15%). Only two relatively small categories potentially address Schelling's interest in "neighbour preference" and even then we must be cautious in our exact interpretation. Kind of people around here (13%) could be interpreted in this way but need not be exclusively about ethnic composition. It could also be about a neighbourhood that is noisy, untidy, unfriendly, "snobby" and so on. Thus this 13% puts an *upper bound* on this kind of Schelling style motivation (assuming accurate reporting but if we do *not* assume this, at least in the first instance, we are dooming ourselves to a competition between entirely speculative models). The same applies to the category kind of schools round here (6%). This *could* be read as a coded reference to ethnic composition but might also refer to other issues such as school quality. (It is actually an element of good questionnaire design that questions should not be open to more than one interpretation by respondents.) Finally, there is a moderately large class of responses which refer to the *locations* of places and persons as part of the moving decision: nearness to friends or relatives (15%), nearness to church (9%), travel to work (8%), shopping facilities (6%). We can thus see that the self-reported motivations on which the Schelling model is based are quite a small proportion of the overall reasons given. (This data is not perfect and should not be treated as quantitatively definitive. In several cases, it could be argued that the data is open to more than one reasonable interpretation. For example should open space around the house be treated entirely as a property of a dwelling or could there also be an element of neighbourhood properties there? Fortunately, for the argument presented here, based on qualitative not quantitative properties of the data, this does not matter).

This analysis gives rise to the caveat about the exhaustiveness of calibration, validation and specification for empirical ABM raised in the last section. One relatively neglected implication of the Schelling model is that it needs no "real" geography. Because (by assumption) people only care about neighbouring types, there is no virtue in modelling school location, shopping centres and so on. But if people *do* report that they mind about such things (as in the Rossi data for example) then the model needs to include them and how schools come to be located where they are is not a fact that we can understand effectively just by talking to individual households. We may thus need to access other kinds of data to justify our model assumptions, creating the kind of simulated environment in which we would expect the empirically observed decision making processes to generate at least broadly the right kind of simulated segregation. (It is an interesting question whether, to do this, we just need a plausible *distributio*n of schools or their exact locations for example.) Thus, in order to be exhaustive, specification cannot be limited to agents but must also

apply to the environment (and the properties of the environment cannot legitimately be simplified just by making arbitrary assumptions about how moving decisions are made.) In the next section I explore the patterns of segregation that arise under the assumption of heterogeneous decision making in the Schelling model based broadly on the Rossi data.

Some Preliminary Results

In the abstract I referred to this issue as a possible blind spot in ABM and it is interesting that existing research (as far as I have yet been able to determine) seems to interpret "heterogeneous decision making" in the Schelling model as meaning *not* that people make decisions in systematically different ways (for example based on different aspects of the environment following Rossi) but that they make decisions in the same way but based on different *preferences* [16, 17]. In fact, ABM with agents that are heterogeneous in decision *process* rather than parameters are still relatively rare. (One reason for this may be that, without specification data, such models could be argued to be avoidably complicated with no compensating virtues.) The same might be said to apply to any model which simply gave different weights to different features of the dwelling, the immediate environment and the wider city like churches and schools. (It is an interesting conceptual question whether permitting "zero weights" would allow these decision models to be considered systematically heterogeneous. It would partly depend on the distribution of weights to different reasons across the population. In such a case, one person could decide entirely on dwelling properties and another entirely on target locations but, arguably, they would still be choosing "in the same way" but based on different weights.) The apparent interpretation of "heterogeneous" in the literature conjecturally supports the view that lack of attention to qualitative data (how do people actually make decisions?) and to specification (why this decision rule rather than that?) has "shaped" the approach to modelling in an implicit way that (once made explicit) cannot really be justified in empirical terms. (This is fine, of course, if the modeller says they are not attempting to build an empirical model as long as they also clearly state what other objective contribution their model makes.) Once the problem has been set up this way, however, it is relatively easy to build a relevant model and investigate the results. The ABM (written in NetLogo version 6.0.4 and available from the author) starts with the standard "Schelling type" found in the relevant Models Library. However, it then adds the possibility of several other types (in proportions the modeller can specify) based on the broad classes of responses found in the Rossi data. For example, one type of agent is unhappy unless they are within a certain target distance (which can again be specified by the modeller) of a particular (arbitrary) location. This is a deliberate abstraction from the idea that various physical locations like schools, churches and offices may have some bearing on whether an agent wants to move. But how will this reason for moving "interact" with the standard Schelling motivation in various different proportions of the population? Can the presence of one type impact on the

convergence of another to an equilibrium state? How and under what circumstances? Another type of agent (again deliberately abstracted) is one that becomes unhappy for "random"—which could actually mean non-modelled—reasons and becomes happy again once they have moved. (It should be noted that this is a moderately plausible way of introducing "noise" into the system—though it would be more plausible if every agent sometimes moved for non modelled reasons rather than one type *always* doing so. It is not really plausible that agents should do something as important as moving house genuinely "at random"—moving costs are another issue missing from the Schelling model quite apart from the housing market—so we might be claiming instead that they did it for reasons without a process yet explicitly specified in the model such as a death in the family.) Such agents are again fundamentally different from both the "location targeting" agents described above and the standard Schelling ones. This is how I use the term heterogeneous in this paper (which is not always how it is used in other research which claims to deal with heterogeneous decision making): That an agent makes decisions according to a different process/specification and not just based on different parameters for a homogenous process. Again, how will different combinations of these various types impact on the resulting patterns of ethnic segregation? The final type presented in this analysis is the "person targeting" type, who judges their happiness not in terms of a physical location but in terms of the position of a designated "significant other". Again, this is a deliberate abstraction from all the possible "person based" reasons (children, grown up parents, friends, partners in business) given by Rossi that might impact on the ranking of different locations. (I have only begun sketching out this approach and one detail still to be clarified is exactly when we should call two decision processes different. For example, the location seeking and person seeking types both make the decision based on distance between them and the target and aiming to reduce it by random movement but the kind of target differs in each case: A place or a person. Further, the person seeking agent may find that the person they "seek" moves while the location seeker will not. Thus, for example, the effect of random movers may be "amplified" to the extent that person seekers happen to nominate them as targets. Does it make more sense, then, to call these two decision processes homogenous or heterogeneous? I merely remark on this issue for now.) For reasons of space and simplicity, one broad type identifiable from the Rossi data has not yet been implemented (though it should be simple to do so). This is the "dwelling targeting" type that is happy if the property they live in meets various criteria of rent, adequate storage, privacy and so on. It is hoped that this type may be implemented by the time of the conference but it could be argued that in terms of the logic of the Schelling model, such types are, on their own, socially trivial. (More generally, however, types that have trivial behaviours in isolation may have non trivial ones in combination with other types. This is one motivation for exploring the behaviour of such systems).

The form of the simulation results is actually very simple to report. Using a population of standard Schelling types as a baseline, a number of runs (ten per experimental condition) were performed with various combinations of types (to be described in detail below). These runs were characterised (as with the standard Schelling model) using % of similar neighbours and % of agents "happy" at the end of the run.

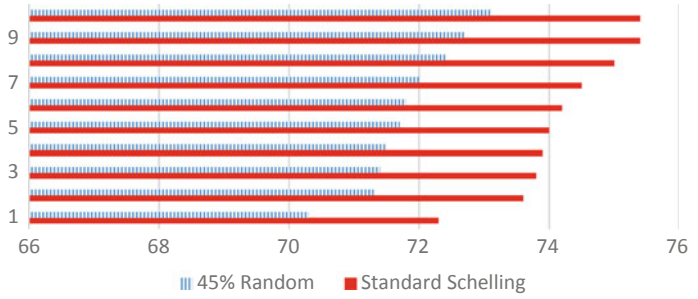

Fig. 23.1 Comparison of % similar neighbours for ten runs where all agents are standard Schelling types (dashed) and when 45% are random movement types (solid)

However, because with some combinations of types, complete satisfaction is impossible (random movers are potentially *never* satisfied), the simulation was run until complete satisfaction or for 100 ticks whichever was shorter. This time period was chosen because the standard Schelling model converges easily within this length of time for the parameters used and there was no sign that any run had failed to converge as far as it was going to within this time period. (The initial population density in all runs was 90% and the % of similar neighbours wanted—where relevant—was 30%.) Figure 23.1 shows a typical result from this comparison between sets of runs.

Several points should be noted from this Figure supporting the rest of the analysis. Firstly, by presenting each set of ten runs ordered by value, it is clear that there is an effect of a certain size between the different conditions. In all runs, the % similar is higher in the standard Schelling condition by about the same amount. Not all such results can be reported here (for reasons of space) but are available from the author. However, such an analysis justifies reporting average values of conditions (and differences between them) as meaningful (and demonstrates that ten runs per condition is "sufficient" for the results to be relied upon.) Secondly, there is the result itself. Clustering is higher when all agents are the standard Schelling type rather than 45% being random movers. (This % was chosen arbitrarily because it divided easily into the three other types being analysed but also because it was a large though not dominant proportion. The majority of agents in all runs—55%—are still standard Schelling types.) Thirdly, although the result is robust, it is small, with a difference between averages of only 2.39%. If we choose to think of the Schelling model in empirical terms, it might be quite hard to distinguish between these two possibilities assuming, for example, that there was noise (or non response) in the household survey data or if the modelled types were only an approximation to the whole population of behaviours discovered by qualitative interviews. This shows that the homogeneity assumption "matters" to the behavior of the model but that the size of the effects may make them quite hard to distinguish without careful attention to data and research design.

It would be wasteful to report all the bar charts for different combinations of types but all bar one (which shows a single anomalous pair of runs out of ten) display the

Table 23.1 % similar and % unhappy for different proportions of four decision making types (standard Schelling, random mover, ideal location seeker, ideal person seeker)

Proportions of non Schelling types (random, location seeking, person seeking)	% similar (average of 10 runs)	% unhappy (average of 10 runs)
0, 0, 0	74.21	0
45, 0, 0	71.82	2.46
0, 45, 0	73.99	7.21
0, 0, 45	64.21	0
15, 30, 0	73.6	5.61
15, 0, 30	69.05	0.84
0, 30, 15	71.58	5.15
15, 15, 15	71.56	3.11

same pattern of consistent effects sizes between pairs of conditions. Table 23.1 is a summary of results from a number of conditions which will form the basis for the rest of the analysis.

A number of points can be noted from this table. Firstly, only the person seeking type on its own is compatible with the complete satisfaction found in the pure Schelling case but leads to significantly lower % similarity. Apart from this, even when there are no random types (which are, by design, perennially unsatisfied), there is always residual dis-satisfaction but the differences between the % similarity in all these conditions are small relative to pure Schelling and are therefore likely to be difficult to detect in data. (It should be noted that this provides another example of the advantages of taking the Schelling model "seriously" in empirical terms. It would be very straightforward to collect data about levels of housing disaffection and "aims" for desired new dwellings as part of the household surveys potentially used to validate the Schelling model. To my knowledge this has not been done however).

Secondly, apart from the condition mixing Schelling types and person seekers, all levels of similarity are broadly comparable across conditions and below the value for the pure Schelling population. There is more proportionate variation in the % unhappy across conditions and we can consider whether this phenomenon can be decomposed into separate effects for each type. Table 23.2 shows the results of this analysis, which

Table 23.2 Calculated and actual % unhappy for different combinations of non Schelling types

Proportions of non schelling types (random, location seeking, person seeking)	Calculated % unhappy	Actual % unhappy
15, 30, 0	5.63	5.61
15, 0, 30	0.82	0.84
0, 30, 15	4.81	5.15
15, 15, 15	3.22	3.11

involves combining the effects of runs with only one other type in proportion to their fractions in the population for runs with more than one other type. For example, 45% random gives an unhappiness of 2.46 and 45% person seeking gives an unhappiness of 0%. This predicts that (assuming the type effects are separable), that a combination of 1/3 * 2.46 + 2/3 * 0 = 0.82%. The corresponding actual value with three types is 0.84% suggesting that these effects are very close to decomposable.

Finally, we can think about what these results might actually signify. Because of the way they are specified, random movers will never be permanently satisfied and thus, in principle, it is easy to distinguish populations containing them from those that do not. In practice, however, this is only true because of another arbitrary simplification in the standard Schelling model, namely that agents do not arrive from outside or leave the area. If they did this (which in reality they do of course), then it might be significantly harder to distinguish dis-satisfaction of this kind from that arising via random mover types. Because of the simplifying assumptions made to illustrate points about decision heterogeneity, there is a plausible explanation for the decomposability of different type effects observed here. Getting near to a single target person (who is statistically unlikely to be the target for many other people) is relatively unchallenging. The same is true of getting close to a particular location because, again, there is little "competition" for any particular spot. The same might not be true if, for example, some target locations were much more "in demand" than others (which we would actually expect, for good schools for example) and your target individual happened to be located close to one of these. Interestingly, the least convincingly decomposable combination involves a mixture of location seekers, person seekers and *no* random movers. Perhaps there is at least some "aggregate interference" by each type in the goal satisfaction of the other which random movers in some way meliorate. (It should be noted that, arbitrarily, agents have to be really rather close to their target to count as satisfied, no more than four "coordinate points"—sum of differences in x and y coordinates.)

However, it is important to draw the right conclusions from this analysis. What is important is that we can explore the effects of decision homogeneity in the Schelling model (and by extension in ABM more generally) but also, to avoid this being a purely abstract exercise introducing complication to no empirical purpose, we can ground this heterogeneity in data (in this case secondary survey data but ideally effective qualitative data) by being explicitly aware of the concept of specification. In the final section I turn to the wider implications of this approach.

Conclusions and Further Work

The argument can usefully be recapped at this point. There seems to be a tendency to minimise the importance of *specification* in ABM (empirical justification of, to take the Schelling model as an example, why agents are assumed to make decisions the way they do and why they are all assumed to make decisions in the same way).

Qualitative data allows us to think effectively about this issue by acting as a potential justification for decision processes *and* their observed diversity. (The standard method of systematic qualitative analysis—so called Grounded Theory—argues that one should focus on the most recurring patterns in data which could be for example, if the right questions were asked, the commonest ways of making decisions [18].) Since this is a work in progress and existing qualitative research on relocation has proved unhelpful regarding moving decisions, I have made the argument "provisionally" using secondary survey data.

Nonetheless, this analysis allows us to think more carefully about exactly what we mean by heterogeneous decision making, to explore the dynamic implications of such heterogeneity (and consider how challenging it might be to confirm or reject such diversity using real data) and to extend our thinking about exactly why (and when) different types of decision making might have (or fail to have) effects that are decomposable in a simplifying way. At this stage, the broad answer seems to be that effects are decomposable when type decision processes do not create "lumpiness" in the structure of residence which might distort the outcomes of other types of decision making. It so happens that the types specified here satisfy this requirement but others that can easily be envisaged need not.

This in turn suggests a number of obvious areas for further the research. The first is a much more detailed exploration of the parameter space with 45% of non Schelling types being pretty much arbitrary. Another obvious way to look at the decomposability of effects is to see whether these "scale". (Are 15% of location seekers 3 times as unhappy as 5%?) Do these results depend on the size of the Schelling type proportion (which has been left constant across all runs?) Are there other combinations of types behaving in unexpected ways? Are type effects less decomposable when satisfaction is more challenging to achieve? A second set of explorations involves contextualising the observed behaviour. Is it still so easy to distinguish between levels of unhappiness when agents arrive and leave in the simulation? What would happen if "random" (i.e. non-modelled) behaviour wasn't a type in itself but applied to other types with a low probability? (It is generally useful to think about what is means for the Schelling model to be non deterministic in empirically plausible ways.) Does it matter that these results abstracted out from agents seeking dwellings with particular attributes? Would such agents "break" the significant amount of decomposability observed here? (More generally, what can we learn from the fact that some type effects are plainly decomposable and others apparently less so?).

Finally, a combination of having a definite model in mind (albeit a stylised one) and seeing what is wrong with existing qualitative research on relocation from the perspective of ABM will make it much more feasible to collect qualitative data that can actually justify a specification for an ABM of ethnic residential segregation that is likely to "survive" validation.

References

1. A.H. Maslow, *The Psychology of Science: A Reconnaissance* (Harper and Row, New York, NY, 1966)
2. T.C. Schelling, Dynamic models of segregation. J. Math. Sociol. **1**, 143–186 (1971). https://doi.org/10.1080/0022250X.1971.9989794
3. E. Chattoe-Brown, Agent based modelling, in *SAGE Research Methods Foundations*, ed. by P. Atkinson, A. Cernat, S. Delamont, J. Sakshaug, R. Williams (Sage, Thousand Oaks, CA, 2019)
4. B. Edmonds, Different modelling purposes, in *Simulating Social Complexity*, ed. by B. Edmonds, R. Meyer (Springer, Cham, 2017)
5. S.F. Reardon, D. O'Sullivan, Measures of spatial segregation. Sociol. Methodol. **34**, 121–162 (2004). https://doi.org/10.1111/j.0081-1750.2004.00150.x
6. Chattoe-Brown, E: Why sociology should use Agent Based Modelling. Sociological Research Online, 18. (2013) https://doi.org/10.5153/sro.3055
7. M. Firth, Racial discrimination in the British labor market. ILR Rev. **34**, 265–272 (1981)
8. G.C. Galster, Racial discrimination in housing markets during the 1980s: a review of the audit evidence. J. Plan. Educ. Res. **9**, 165–175 (1990). https://doi.org/10.1177/0739456X9000900302
9. P.H. Rossi, *Why Families Move: A Study in the Social Psychology of Urban Residential Mobility* (Free Press, Glencoe, IL, 1955)
10. W.A.V. Clark, Residential preferences and neighborhood racial segregation: a test of the schelling segregation model. Demography **28**, 1–19 (1991)
11. J. Goering, Shelling [sic] redux: how sociology fails to make progress in building and empirically testing complex causal models regarding race. J. Math. Sociol. **30**, 299–317 (2006). https://doi.org/10.1080/00222500500544144
12. S.D. Angus, B. Hassani-Mahmooei, "Anarchy" reigns: a quantitative analysis of Agent-Based Modelling publication practices in JASSS, 2001–2012. J. Artif. Soc. Soc. Simul. **18** (2015). https://doi.org/10.18564/jasss.2952
13. B. Edmonds, Using qualitative evidence to inform the specification of agent-based models. J. Artif. Soc. Soc. Simul. **18** (2015). https://doi.org/10.18564/jasss.2762
14. C.H. Gladwin, *Ethnographic Decision Tree Modelling* (Sage, Thousand Oaks, CA, 1989)
15. A. Metcalfe, It was the right time to do it: moving house, the life-course and kairos. Mobilities **1**, 243–260 (2006). https://doi.org/10.1080/17450100600726621
16. E. Hatna, I. Benenson, Combining segregation and integration: Schelling model dynamics for heterogeneous population. J. Artif. Soc. Soc. Simul. **18** (2015). https://doi.org/10.18564/jasss.2824
17. Y. Xie, X. Zhou, Modeling individual-level heterogeneity in racial residential segregation. Proc. Natl. Acad. Sci. **109**, 11646–11651 (2012). https://doi.org/10.1073/pnas.1202218109
18. P. White, *Education and Career Choice: A New Model of Decision Making* (Palgrave Macmillan, Basingstoke, 2007)