Article

On the emergence of descriptive norms

Ryan Muldoon
University of Pennsylvania, USA

Chiara Lisciandra
Munich Center for Mathematical Philosophy, Germany and Tilburg Center for Logic and Philosophy of Science, The Netherlands

Cristina Bicchieri
University of Pennsylvania, USA

Stephan Hartmann
Ludwig-Maximillians University, Germany

Jan Sprenger
Tilburg University, The Netherlands

Abstract
A descriptive norm is a behavioral rule that individuals follow when their empirical expectations of others following the same rule are met. We aim to provide an account of the emergence of descriptive norms by first looking at a simple case, that of the standing ovation. We examine the structure of a standing ovation, and show it can be generalized to describe the emergence of a wide range of descriptive norms.

Keywords
norm emergence, social norms, norm dynamics, agent-based modeling, simulation, descriptive norms

Corresponding author:
Ryan Muldoon, Philosophy, Politics and Economics Program, University of Pennsylvania, 311 Claudia Cohen Hall, 249 S. 36th St, Philadelphia, PA 19104, USA.
Email: ryan.muldoon@gmail.com

Descriptive norms hide in plain sight. While we may not always think of them, they govern many of our day-to-day interactions: they help guide our fashion choices, our table manners, the colors we wear at weddings, and any number of other small features of our social interactions. This governance can become evident when we travel: many of our small-scale behaviors and interactions are culturally contingent. Americans typically greet each other with handshakes. Many continental Europeans greet each other with kisses to the cheek, but the number varies between countries. Standards for personal space vary across cultures. It would be difficult to argue that any one of these practices is 'right', descriptive norms do not carry the normative weight of social or moral norms, but we all follow the norms from our own cultural context and imitate the behavior of our peers. While this may provide a satisfactory account of how descriptive norms operate, it does not tell us about how they came to be. What is it about a given norm that caused everyone to start following it?

To begin to answer this question, we will turn to a simple case, that of the standing ovation. Standing ovations, like many other descriptive norms, are the result of spontaneous coordination of individual choices across many individuals. They have become a common practice after many live performances, even though there is no pre-arranged plan or even any formal coordination. All individuals can do is decide whether or not they wish to stand, based on their own preferences, and what they see others doing around them (Miller and Page, 2004). Put slightly more formally, most agents have preferences about whether or not they like to stand up, which depend on the quality of the concert. They know that standing to clap is a common option after a performance, and they have (conditional) preferences for standing up if the other agents stand up too. This exemplifies a *descriptive norm*: the agents know that there is a behavioral pattern (standing to clap) that applies to the situation they are in, and they prefer to conform with the group (Bicchieri, 2006, pp. 31–32). In other words, their behavior is not only determined by unconditional preferences for certain actions, but also by their desire to conditionally conform to the behavior of a sufficiently large group.

Standing ovations are a useful stand-in for describing societal transitions to a wide variety of descriptive norms. That is, there is a status quo behavior that can be upended by an alternative. For example, fashion often works in this way. Prior to mini skirts, women wore longer-length skirts, and upon their introduction, the population largely shifted to a mini-skirt norm. Movements in popular music also follow a similar pattern: teenagers largely coalesce around a few bands or a particular kind of music for a few years, before giving way to a new set of music. Calling etiquette has similarly shifted as email has become a more common form of communication. What we find in all of these cases is a status quo that, without any central coordination, loses out to a new behavioral rule. By focusing on a standing ovation model as an exemplar of this wider set of phenomena, we can avoid the problem of getting lost in small case-specific details, and instead try and identify the key features of individual decision-making that can affect the emergence of a norm. Thus, we aim to examine how ovations might arise, and in doing so, come to a more general account of the emergence of descriptive norms in a population.

To provide for this more general account, we investigate several features of individual decisions, such as a desire to conform, one's knowledge of what others are doing, and one's own preferences. These elements can affect the emergence of a descriptive norm

in a group and influence some aspects of the processes, such as whether the group converges on a single behavior, and if it does, how quickly this happens. Our model allows us to carefully explore the key aspects of individual decision making that drive these collective behaviors.

One additional fact that we wanted to take into consideration was that although descriptive norms can be built out of many small decisions, they do not always emerge. While many of our day-to-day activities are governed by norms, not all of them are. Plenty of our actions are different. We do not all walk in lock-step. Further, many of our seemingly coordinated actions can be simply described as behavioral regularities: agents act purely in accordance with their intrinsic preferences, which just happen to align with others'. Descriptive norms, on the other hand, arise when the desire to conform to the behavior of others overwhelms one's initial preferences. Our model helps us to explore the contingent nature of many descriptive norms. What this model suggests is that it is possible that some descriptive norms become descriptive norms for no particular reason other than the peculiarities of the individuals in the population.

In this paper we explore four main models:

1. In the first model, *the baseline*, we build a model for a standing ovation, which considers an individual's decision about whether she will stand to be a combination of her personal unconditional preference and her tendencies to choose to conform to the behavior of others.
2. In the second model, *the inertia model*, we introduce two new features of individual decision-making: first, a tendency of individuals to become increasingly set in their ways as time goes on, and second, a more nuanced model of social contagion, to better match how bandwagon effects occur.
3. In the third model, *the endogenous social sensitivity model*, we treat one's sensitivity to the behavior of others not as something separate from one's individual preferences, but as dependent on them.
4. In the fourth model, *the symmetric model*, we consider a reversible case, where each agent can decide whether to stand up or to sit down in any round. We assume that the population is made by two agent types with opposite preferences. This setting allows us to compare the emergence of descriptive norms, where all individuals set on the same action regardless of their own preferences, and behavioral regularities, where all individuals continue to follow their intrinsic preferences despite other peoples' actions.

Each model helps us learn more about the nature of social decision-making. The first three models explore the robustness of the standing ovation approach as an explanation of directed norm emergence. These three models together allow a more nuanced view of social effects on individual decision making than just the baseline standing ovation model would allow. By examining more decision procedures, we are able to better describe a larger class of directed norm transitions than we would otherwise be able to. Being able to look at information asymmetries between agents further allows us to examine the effects of social hierarchies on norm emergence. The fourth model that we consider widens the scope of our investigation into norm emergence by examining

a symmetric case that allows for norm emergence in either direction, or none at all. Here we are able to more fully investigate the conditions for when norms do not emerge, since the model allows us to compare behavioral regularities from descriptive norms, as we can directly inspect agent preferences. Unlike Lewis' approach to the analysis of conventions, according to which conventions are solutions to coordination games where 'each wants to conform conditionally upon conformity by others' (Lewis, 1969), we focus on the decisions of agents who do not reason about other agents' expectations, and on norms that themselves have no intrinsic coordination advantages. We are interested in studying the dynamics created by individuals who have both intrinsic preferences to act and some interest in conformity. This allows us to focus on a wider range of more common descriptive norms, many of which end up having large cultural variation. While much of this analysis could apply to Lewis conventions, we examine a less structured environment. In the family of models we consider, not only do we allow for agents to have interests beyond pure coordination, we allow for asymmetries between agents across several dimensions, and consider both complete and incomplete information conditions.

Together, these models provide a more general account of descriptive norm emergence than has been seen so far in the existing literature. First, by focusing on norms that have no inherent differences in utility, we can focus on the large class of under-studied everyday norms, such as norms of personal space, etiquette, dress, eye contact, and other small-bore issues. These norms when put together help explain a large portion of our social behavior, even if any individual norm has only a small effect. Second, by introducing information asymmetries between agents, we begin an analysis of the effects of social hierarchies on norm dynamics. Third, by splitting our study of norms into the directed and bidirectional cases, we can then study the differences between behavioral regularities and descriptive norms, which can otherwise get lumped together in other literature.

1 The baseline model

1.1 Model description

Let there be M people in the audience of a theater play. The variable $s_i^{(n)}$ with $i = 1, \ldots, M$ indicates whether person i is sitting ($s_i^{(n)} = 0$) or standing ($s_i^{(n)} = 1$) at time-step ("round") $n = 0, 1, \ldots$. Time is discretized and at $n = 0$, everybody is seated. Everybody who is not yet standing "updates" her position in each round. Our central idea is that whether or not a person stands up depends on her *effective propensity* to do so. The effective propensity of person i is the convex combination of two factors:

1. An *intrinsic preference* q_i to stand up. This represents an individual's preference to stand up or not, independently of what other people are doing.
2. An *extrinsic propensity* to stand up. This factor takes into account what other people in the audience are doing. So whether or not someone stands up in round n will depend on how many people $S^{(n-1)} := \sum_{i=1}^{M} s_i^{(n-1)}$ were standing up at round $n - 1$. Note that $S^{(0)} = 0$.

Combining these two factors, we arrive at the following expression for the *effective propensity* to stand up in round n:

$$P_i^{(n)} = \sigma_i \left(\frac{S^{(n-1)}}{M} \right) + (1 - \sigma_i) q_i \qquad (1)$$

We see that the variable $\sigma_i \in (0, 1)$ measures how much person i takes the extrinsic propensity, i.e. social considerations into account. It determines the relative weight of intrinsic and extrinsic propensity. Therefore, we say that σ_i measures the *social sensitivity* of person i. Let us call the model in Equation (1) the *baseline model*. In each round n, everybody who is still sitting considers her propensity $P_i^{(n)}$ and then decides, by a chance mechanism, whether or not to stand up.

Under the conditions of this model, the number of standing people will converge to 1, as n goes to infinity.[1] However, in practice, only a finite number of rounds will be played, and if the σ_i's are sufficiently small, it may well be the case that not everyone stands up. In order to better study this model, let us now turn to a numerical analysis of the model.

1.2 Numerical analysis

Our model (and its extensions, which we will discuss below) suggests a variety of numerical studies. In order to best investigate these cases, we turn to instantiating the models as agent-based computer simulations. The simulations were written in Netlogo 4.0.4.

In the agent-based simulations, 1089 agents are seated in a 33×33 grid, all facing the same direction, in order to represent individuals seated in a theater. As is standard in agent-based models, time is broken up into discrete steps. In each step of the simulation, seated agents independently assess whether or not they should stand. Their decision procedure is simply an instantiation of the equation described previously. As was noted before, agents who are standing remain standing in perpetuity. Each simulation is run until either every agent is standing, or 1000 steps have passed. If each step represents one second of actual time, 1000 steps represents nearly 17 minutes, which we consider to be the extreme upper end of how long a standing ovation might last. For the purpose of analysis, all simulations were repeated 100 times.

1.2.1 Complete information. In our first set of simulations, we examined agents who could see the entire audience. Their position in the theater had no effect on what information was available to them. As such, agents all worked from precisely the same information about what others in the audience were doing.

As we have noted previously, the baseline model guarantees convergence on a standing ovation. So instead of discussing whether or not a standing ovation occurs, we study the *speed of convergence*. In particular, we are interested in determining how the parameters specified in the baseline model affect convergence times.[2] To examine these effects, we must consider the agents' intrinsic propensity to stand up, and the agents' social sensitivity in turn. We will first examine the effects of the intrinsic propensity on ovation convergence.

In our studies of the intrinsic propensity we held social sensitivity σ_i fixed at various values (Figure 1) so we could examine how the increase of the intrinsic propensity by itself affected rates of norm convergence. In general we saw that, as the intrinsic
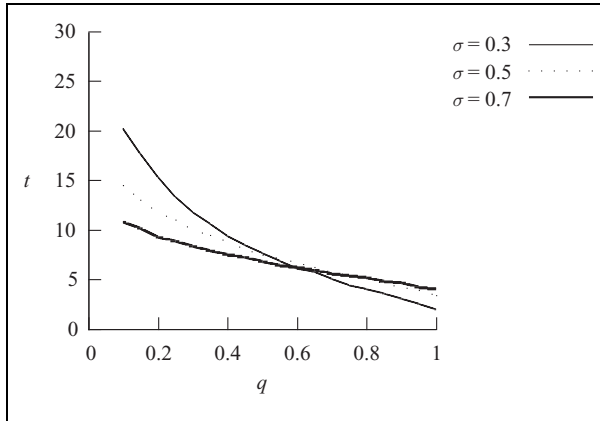
Figure 1. The effect of intrinsic preference on time to convergence.

propensity increases, the time of convergence decreases. This is, of course, exactly what we expect from the mathematics of the model (1).

We found a similar story with an examination of the social sensitivity (Figure 2). As before, what we saw is that as the social sensitivity increases, the time of convergence decreases. However, the speed of convergence diminishes at a fairly different rate when the social sensitivity is combined with a very low value of intrinsic propensity ($q_i = 0.1$). This makes perfect sense: when both the intrinsic propensity and the social sensitivity are very low, it takes a long time until *everyone* is standing. It is enough that the social sensitivity slightly increases for the initial deadlock to be resolved.

We found that both parameters make a notable difference: each can cause the convergence rate to be significantly faster. However, the manipulation of social sensitivity appears to diminish convergence time more pronouncedly. This also makes sense: sensitivity to one's peers will accelerate any bandwagon effect as the population moves towards convergence.

1.2.2 Incomplete information. In our second set of simulations, we examined agents who were limited in how much of the audience they could see. In particular, agents could only see the agents in front of them within their range of vision. In this model, agents could see all the way to the front of the theater, but only within a cone of 30°. Thus, agents could not see anyone behind them, nor anyone outside of this limited scope of vision in front of them. This extension assigns different degrees of influence to the agents: those in the front rows are highly influential as their choices are taken into account by agents seated towards the rear, whereas the latter's choices affect few other agents. Notably, we obtain an asymmetry between information and influence: agents at the front can only see the behavior of a narrow peer group (or they do not care about the rest), while those at the rear have complete insight. This extension is therefore a particularly intuitive way to model mutual impact in social hierarchies.

To describe this more formally, consider an audience of R rows with L seats in each row. Now everybody takes only a fraction of the whole audience into account when calculating
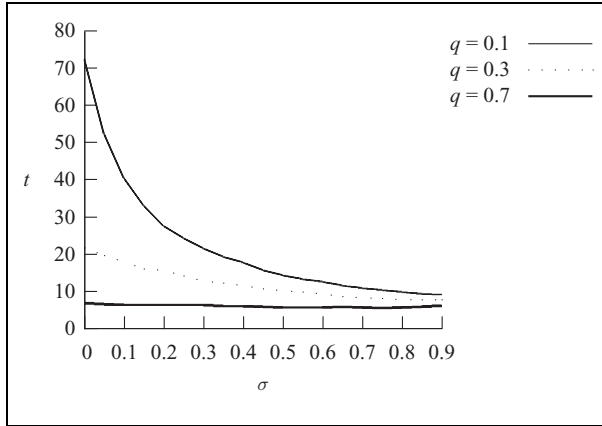
Figure 2. The effect of social sensitivity on time to convergence.

the extrinsic propensity, for example the cone of people in front of the person. Then, clearly, the behavior of the people in the first row will be more important than the behavior of those in the last row. After all, almost everybody will look at what the people in the first row are doing. In this case, the ratio $S^{(n-1)}/M$ is to be replaced by the expression

$$\frac{1}{|M_i|} \sum_{j \in M_i} s_j^{(n-1)} \qquad (2)$$

Here M_i is the group of people person i can observe.[3]

As with our study of model of complete information, we first held the social sensitivity parameter fixed at discrete values to study the effects of the intrinsic propensity's increase (Figure 3). We found that the model with incomplete information behaved very similarly to the complete information case, although convergence times were notably slower at low values of q_i. Whereas in the complete information case, when $q_i = 0.1$, the average time to convergence was 20.25, in the case of incomplete information the average time to convergence was 59.47. As q_i grew, however, these disparities disappeared.

This suggests that while limited information can have a notable effect in slowing down convergence times in cases of low levels of intrinsic propensities, this effect rapidly diminishes as agents' intrinsic propensity to stand increases.

As we turn our attention to the effect of the social sensitivity parameter, however, we find that limited vision has a strikingly large effect that reverses the trend seen in the model of complete information (Figure 4). Whereas before, moving σ_i from a low to a medium value induced significantly lower convergence times, in the model of incomplete information, as σ_i increases, convergence time also increases. Once $\sigma_i > 0.5$ we find particularly dramatic increases in both convergence times and variance. Limited vision of others has a dramatic effect on convergence, and for good reason. As the social sensitivity increases, the effect of the intrinsic propensity diminishes. When only very few agents are in a position to affect the behavior of others, it can easily happen that they remain seated. This can then lead those that look to them for guidance about standing also remain seated. This dynamic substantially
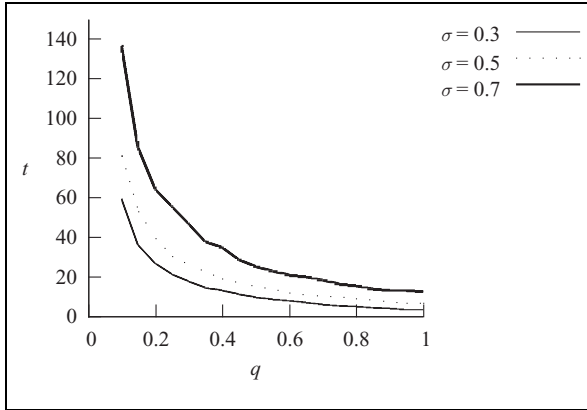
Figure 3. The effect of intrinsic propensity on time to convergence in the limited vision case.
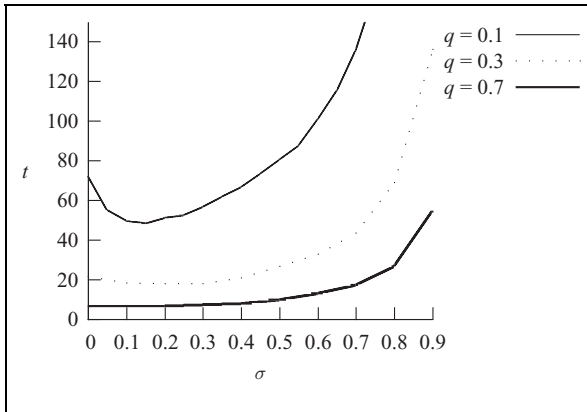

Figure 4. The effect of social sensitivity on time to convergence in the limited vision case.

dampens the bandwagon effect that is found in the baseline model with complete information.

2 Problems with the baseline model

While the baseline model helps to illuminate the basic structure of individual decisions that can result in the emergence of a descriptive norm such as a standing ovation, there are several reasons to suspect that the model is not yet adequate. Drawing upon philosophical and empirical literature on norm compliance (Bicchieri, 2006; Young, 2009), we can levy three major criticisms at the baseline model. We will look at each in turn.

The first way in which our model falls short is that it is not very sensitive to a more nuanced psychology of decision-making. One thing it fails to capture is the idea of entrenchment: people can often become set in their ways over time, and become less and

less willing to change their minds, even if social influences become significant. In addition, the baseline model lumps the notion of social sensitivity in with the notion of social contagion: it assumes that larger and smaller groups exert the same amount of social influence over a person's decision. However, it is likely that in some instances, small groups are sufficient for influencing individual choices, while in others, a much larger group is necessary to change an individual's decision. For example, one might not need too many people to join a small fashion trend, such as wearing scarves, to start following them, but we might expect that for other descriptive norms that are more drastic changes from the status quo that one would want to see more people following it before following suit.

The second way in which our model falls short is that it assumes that the amount others can influence us is always constant across different contexts. However, this is unlikely to be the case. In instances where one has strong preferences, it is likely that social pressure is less important. Whereas, when someone is fairly indifferent between two actions, social pressure might be the main determinant of that person's choice.

The third way our model falls short is that it makes significant structural assumptions about the nature of descriptive norms that may inhibit its ability to be a useful general model. This comes in two ways. Most obviously, the model always leads to a convergence on everyone standing. This is a highly suspect assumption: there are many concerts in which standing ovations fail to occur, just as there are many *potential* descriptive norms that never come to be. Even still, we can expect many situations in which *some*, but not all, agents take on a particular action, and for this to be stable. As it currently stands, our model cannot capture this fact. In addition, the model suffers from having a built-in implicit assumption about the directionality of norms. In our baseline model, people go from sitting to standing. It is impossible for sitting to become a norm. Likewise, it should be possible for no norm to emerge.

In the following sections, we will present extensions to the baseline model that will in turn seek to address these three deficiencies. What we will show is that the qualitative results from our baseline model continue to hold as we investigate the first two deficiencies. As we explore the structural assumptions, we will find additional constraints on the emergence of descriptive norms that further enrich our account.

3 The inertia model

In this extension, we seek to address the lack of nuance in the psychology of the baseline model's decision procedure. To do this, we make two changes. First, there is a scaling factor $e^{-\alpha n}$, $0 < \alpha < 1$, such that the greater the number of rounds that have passed, the less likely it is that someone stands up. We introduce a scaling factor $e^{-\alpha n}$, $0 < \alpha < 1$, such that the more rounds have passed (higher n) the less likely it is that someone stands up. That is, the effective propensity of an individual to follow the norm decreases as a function of the number of rounds. This allows us to more carefully investigate the notion of entrenchment. Second, the propensity to stand up as a function of the others' behavior $S^{(n-1)}/M$ is taken into the βth power, $\beta > 0$. All of this can be represented by the following equation:

$$P_i^{(n)} = e^{-\alpha n} \left(\sigma_i \left(\frac{S^{(n-1)}}{M} \right)^\beta + (1 - \sigma_i) q_i \right) \tag{3}$$

Here β can be thought of as a measure of *contagion* in the group: the smaller β ($0 < \beta < 1$), the higher the chance that a few isolated individuals who rise from their seats will affect the rest of the group. In other words, if we keep the number of agents following the norm fixed, the propensity of an individual to follow a norm is higher for smaller beta. This reflects the fact that there are circumstances in which it takes a few agents to trigger a conformity effect than others. So β determines the relative influence of the first agents adopting the norm vis-á-vis those agents that adopt it at a later stage. Thus, agents with a large β act on the basis of their propensity and the observed behavior of a crowd (as opposed to being responsive to the behavior of individuals and small groups). The break-even point is $\beta = 1$.

In this model, there is a non-trivial probability that not everyone stands up, even if infinitely many rounds are played.[4]

Finally, it should be stressed once more that contagion and social sensitivity play different roles: while social sensitivity balances an agent's individual preferences against the impact of the behavior of others, the contagion parameter determines the rate at which the influence of an additional person standing declines with the number of standing. If β is low the first few people standing will have a much larger influence than the final few.[5]

3.1 Numerical studies

The inertia model is meant to provide a mechanism for non-complete ovations, by providing a countervailing force on individual decision-making, encouraging some to remain seated. This is done with a time-dependent scaling function, which can be made more powerful by increasing the size of the inertia parameter α. Second, it dampens the effect of social influence, by taking the social sensitivity component of the base model and raising it to the βth power. We will investigate each of these modifications to the base model in turn, considering their effect on ovation size in equilibrium.

As the model description indicated, the inertia parameter α has by far the largest effect on ovation size. Here we will consider α with β fixed at 0.1. In general, as the inertia parameter grows, we find an exponential decay in equilibrium ovation size. We examined values of $0.01 \le \alpha \le 0.5$ in steps of 0.01. As represented in Figure 5, we find a rapid decay in equilibrium ovation size.

The inertia parameter controls the rate at which agents are willing to stand as time goes on, which heavily dampens their ability to respond to new information as it is revealed to them. As agents become increasingly stubborn as time goes on, this limits their interest in standing regardless of what anyone else is doing. We find a similar story with the contagion parameter (beta) in Figure 6: since individuals respond less to smaller groups if (beta) is large, it is more difficult to get a bandwagon effect initiated, even if they are increasingly sensitive to large groups. The large groups simply cannot form if smaller groups do not have sufficient attractive force. In this way, α and β work in concert to limit ovation size: β limits the power of an initial small group standing, and then α
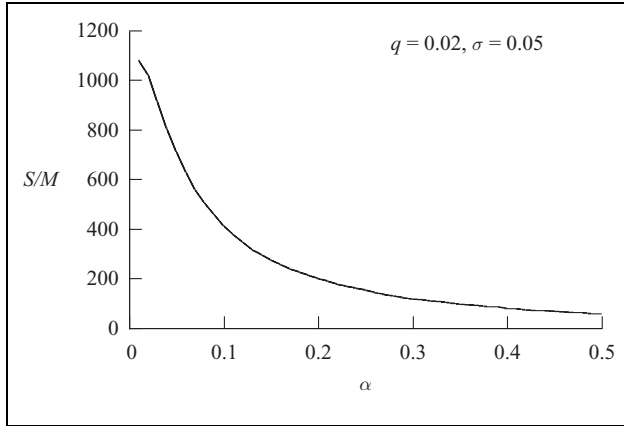
Figure 5. The effect of the inertia parameter on the number of people standing.

increases the stubbornness of agents sitting as the groups get slightly larger over time. This combined effect can neuter a group's ability to create a social bandwagon.

In the inertia model, we find that the imposition of incomplete information has very little effect on how the active variables in the model affect ovation size. In the case of α, we find no discernible difference between the complete information model and the model of incomplete information. In the case of β, we find few differences where $0 < \beta < 1$.

4 The endogenous social sensitivity model

The baseline model and its initial extension, the inertia model, consider social sensitivity as an exogenous parameter: it is a parameter that balances one's intrinsic propensity to comply with the behavioral rule with the impact of group behavior. In other words, for an agent with high social sensitivity, the impact of group behavior will dominate the impact of one's individual judgment on the quality of the concert, and vice versa. Social sensitivity does not, on that account, depend on one's intrinsic propensity or the number of people already following the behavioral rule.

This view can, however, be challenged. In her book "The Grammar of Society", Bicchieri (2006) has shown that *empirical expectations* of the behavior of others are crucial to whether descriptive norms emerge and persist. If an agent expects a critical part of the population to follow a behavioral rule, then she will most likely follow the rule as well. Further, an agent may only become aware of the existence of a candidate alternative norm once it is sufficiently widespread in the population. If a large part of the group starts to comply with the rule, the agent reasonably expects that the behavior will spread to the entire group, and these expectations overrule an agent's independent preferences as a determinant of her individual behavior. Conversely, if the percentage of individuals abiding by the rule is lower than such a critical value, group behavior barely affects individual behavior. Social sensitivity should thus be treated as an *endogenous* variable crucially depending on the observed behavior in the group.
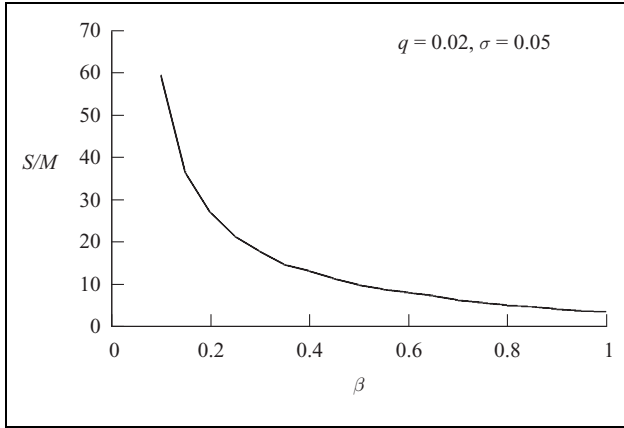
Figure 6. The effect of a contagion parameter on the number of people standing.

Both in the baseline case and the inertia extension, social sensitivity was considered exogenous. If it is low at the outset, then it will stay low, even if the norm spreads rapidly in the group. This delays the convergence process. It is therefore worthwhile to investigate whether our results remain robust under the feedback effects described above. To that end, we have to specify the dependency between the variables of the original model.

We keep the baseline equation (1) intact and only write social sensitivity σ as a function of the other variables.[6] As argued above, social sensitivity should be very low ($\approx \varepsilon$) if S/M is significantly below a critical value, and very high (≈ 1) if S/M significantly exceeds that value. It is natural to assume that the lower the intrinsic propensity q, the higher the threshold: if an agent strongly dislikes the behavioral rule, her empirical expectations of compliance with the rule will be higher, and the group will have to behave more homogenously in order to meet them. Thus, we might choose

$$\sigma\left(\frac{S}{M}\right) = \begin{cases} 1, & \text{if } S/M \geq 1 - q \\ \varepsilon, & \text{otherwise} \end{cases} \tag{4}$$

Figure 7 gives a graphic representation of the model. On the x-axis is the number of people standing up on the total audience, on the y-axis the values of sigma.

When the intrinsic propensity of an individual is high her decision to stand up is triggered by few individuals standing, whereas when it is low she needs to see more people to follow them as well.[7]

4.1 Numerical studies

In this set of simulations, we studied the effect of the intrinsic propensity on time of convergence. We examined values of q varying from 0.01 to 1 in steps of 0.01. For simplicity of our treatment, we do not introduce the inertia and contagion parameters, as we have previously examined them in isolation. Further, as we did not find crucial
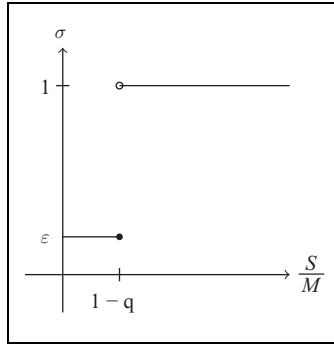
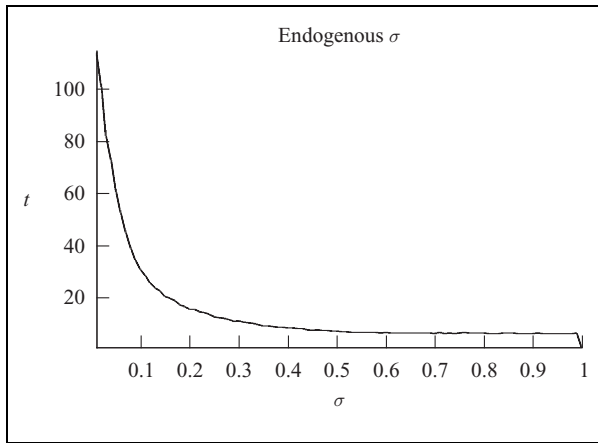**Figure 7.** Discontinuous model.

Figure 8. The effect of the endogenous sigma on the time to convergence.

differences between the incomplete and the complete information case, we only present the first case. As expected, and Figure 8 shows, we find that the time of convergence decreases for increasing values of intrinsic propensity.

5 The symmetric model

In our final model, we consider a generalization of our original model. As we have discussed previously, the models we have been considering thus far all have an implicit directionality: people start out sitting, and potentially stand. The only descriptive norm that can emerge is one of ovation. However, this assumption limits our ability to describe the emergence of descriptive norms more generally. As an example, consider the norm that governs how forks are used while eating. In Europe, a fork is used in the left hand, so as to enable the eater to use a knife in her right hand. In the United States, however, while forks are held in one's left hand while cutting food, they are then moved to the right hand for raising food to one's mouth. While either method of using forks is perfectly suitable

for consuming food in a polite and efficient manner, they are regionally segregated. In the United States, it is rare to see the European method, and the US method is rarely seen in Europe. What we can notice is that there is no particular reason to think that one method is prior to the other, so we cannot model this as a standing ovation. So how might we provide an explanatory framework for the emergence of norms when the potential behaviors are on equal footing?[8]

What we propose is a return to our baseline model (1), but with a few crucial changes. First, we introduce two agent types, each type having a preference for one of the two actions available to them. So type 1 agents prefer action 1, and type 2 agents prefer action 2. Second, we re-interpret the variable $0 < q < 1$, such that 0.5 represents the indifference point, rather than 0. On this new interpretation, 1 represents a strong preference for the action of one's type, and 0 represents a strong preference away from this action. Third, we allow agents to change their minds: whereas in previous models once an agent has chosen to stand, she must remain standing, in this model agents can reverse course and go back to their previous action. Finally, when we initialize the model, agents are randomly (and independently) assigned an initial starting action. So, unlike previous models, our starting state has half of the agents performing the first action (say, the European way of using forks), and half the agents performing the second action (like the US method of using forks). These changes allow us to investigate several things that could not be examined in previous models. In particular, since we are treating the two methods of using forks as symmetrical to each other, and we allow agents to change their minds, we should expect the models to exhibit more complex behavior. More importantly, however, we are now able to clearly separate cases of norm emergence from behavioral regularities, since we have differing preferences amongst agents.[9]

The model's equilibrium states can be broken down into three classes: descriptive norm emergence, large-scale behavioral regularities, and a mix of behaviors. Descriptive norm emergence is found when the entire population settles on a single action. In these cases, half of the population must be going against their intrinsic preferences, and instead their social sensitivity drives their decision-making. This can be contrasted against large-scale behavioral regularities, which are cases in which all agents of one type choose the same action, but choose a different action from agents of another type. So everyone who prefers European fork handling employs it, and likewise everyone who prefers US fork handling does so. In this case, we claim that individual preferences are the most powerful guide to decision-making, and so social sensitivity effectively drops out of consideration. Our final case is what is left over: a mix of influences, none of which is strong enough to completely guide behavior. In these cases, both preferences and social sensitivity are at work, neither of which is sufficiently strong to overpower the other. So we find an unsystematic mix of behaviors.

Let us now turn to a numerical analysis of this model.

5.1 Numerical studies

This model requires a different approach to our analysis. Rather than consider something like time to convergence, we must instead consider the probabilities of settling into the
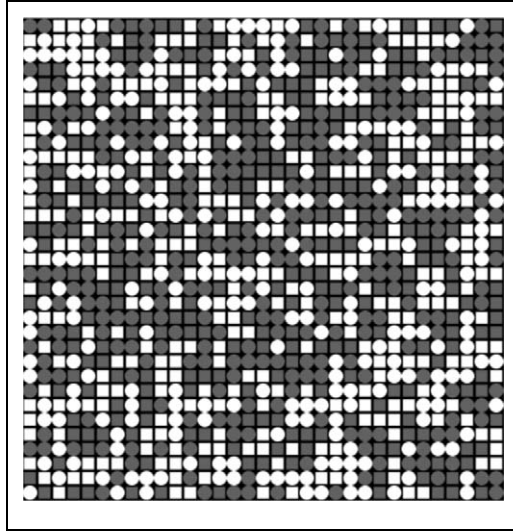
Figure 9. Symmetric model's initial state.

three different states for the different values of intrinsic preference and social sensitivity. The initial state is shown in Figure 9, as it appears in the Netlogo interface. The grid represents the theater audience, composed of two agent types: squares are agents who prefer standing and circles are agents who prefer sitting. White squares stay for those agents who perform their preferred action, in this case standing, otherwise they are black. Black circles stay for those agents who perform their preferred action, namely sitting, otherwise they are white.

In the set of simulations for the symmetric model we examined the probability for a norm to emerge according to the distributions of social sensitivity and intrinsic propensity between the audience. The norm emergence corresponds to a state in which all agents in the audience perform the same action, regardless of their intrinsic propensity. Graphically (Figure 10), this happens when all circles and squares are white (or when all circles and squares are black).

What we found was that the emergence of full descriptive norms is quite rare. Figure 11 represents the probability of descriptive norms emergence for increasing values of intrinsic propensity. Each curves corresponds to fixed values of social sensitivity. We can see that this probability increases as the agents become indifferent between the two actions (q ; 0.5) and in general for higher values of social sensitivity.

More common are large-scale behavioral regularities (Figure 12). These occur when agents perform their favorite action, e.g. when all squares are standing and all circles are sitting. These outcomes can happen for wider ranges of social sensitivity, so long as the intrinsic preference is more extreme in value. Figure 12 shows that the probability of behavioral regularities is lower for intermediate values of intrinsic propensity and for lower values of social sensitivity.
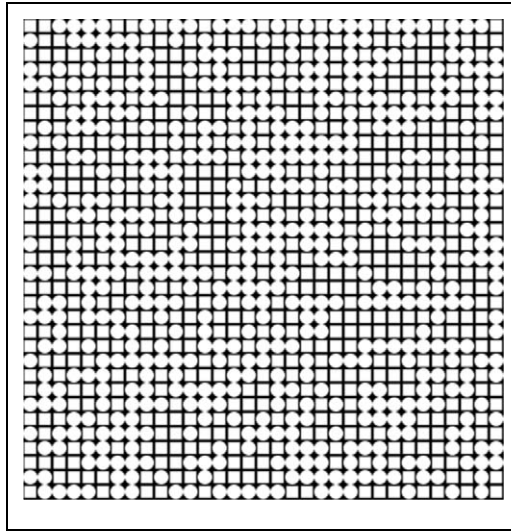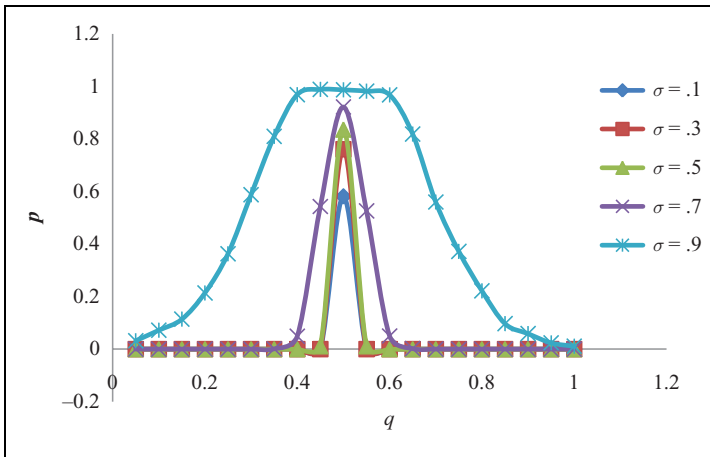
Figure 10. Full norm.

Figure 11. The probability of descriptive norms emergence.

Most common of all, however, are mixed outcomes, those in which some of the agents perform their preferred outcome and others do not (graphically, this corresponds to a situation similar to the initial state but with a different distribution of colors, according to those circles and square that modified their initial state).

This result seems to comport well with the real world: although descriptive norms are quite common and are found in a very wide variety of social situations, there are many more possible descriptive norms than there are actual descriptive norms. Most of our day-to-day behaviors are not norm-governed, even though many are.
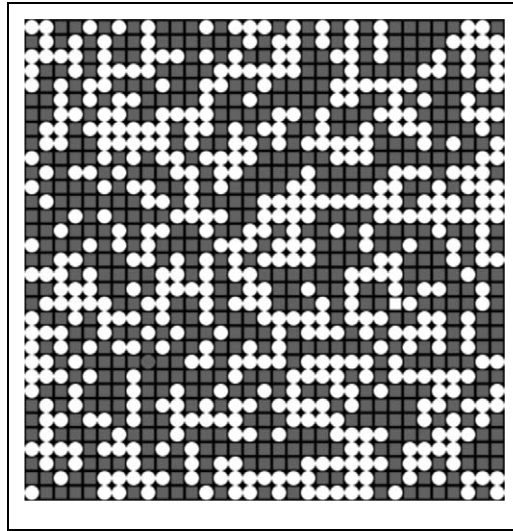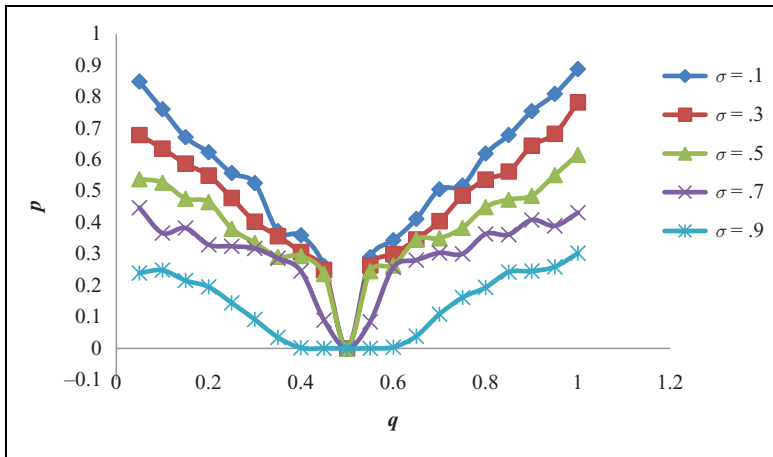
Figure 12. Behavioral regularities.

Figure 13. The probability of behavioral regularities emergence.

6 Conclusions

We have argued that a model of standing ovations can provide a useful framework for the investigation of the emergence of descriptive norms. While we do not claim that all descriptive norms have the character of standing ovations, we have tried to suggest that with a few modifications, we can transform a model of standing ovations into a general model of the emergence of descriptive norms. In particular, we wish to stress the qualitative match of results across the various models we present. The baseline, inertia, and

endogenous social sensitivity models all explore the convergence dynamics of a 'directed' transition from one behavior to another. Although they build in substantively different psychological assumptions about the agents involved, we find that these large perturbations do not shift us far away from our original baseline results.

While descriptive norms themselves most often are not fully captured by the baseline model, it can often be the case that these sorts of directed transitions *can* describe the propagation of information about the social context for behaviors. For example, Christians remove hats in Church to show proper deference, but not at sporting events. When they enter a novel environment, for which they may or may not have to signal deference, say going into a classroom or a museum for the first time, they may look to others for signals of what they should do. When we enter into a novel situation, we may not be sure which of our already-established norms ought to govern our behavior. A directed transition model, such as a standing ovation, might be a good representation of this sort of phenomenon.

The final model we consider, we contend, does capture the essential elements of the emergence of descriptive norms, given that it is possible for any behavioral rule, or none at all, to emerge as a norm. What is so striking about this last model is that it is only a minor modification of the original baseline model, but provides a dynamic that displays the relevant considerations for the potential emergence of a descriptive norm. We did not do anything to change the fundamental decision procedure: we simply allowed people to change their minds, and have preferences for more than one action. But with these small changes, we generalized the model, and enabled ourselves to discuss a much larger class of social phenomena.

This kind of modeling offers some insights into the nature of descriptive norms that might be difficult to arrive at otherwise. In particular, what we find, especially by studying our symmetric model, is that whether a norm emerges at all, let alone which norm it is, is remarkably contingent on factors that have nothing to do with the substance of the norm itself. Whether it is table manners or audience behaviors, or even how we dress, we do not follow them because they are somehow superior to their alternatives, but rather we follow them because a mix of social and personal factors happened to nudge us in one direction rather than the other. We often place value in these norms, but we should avoid making the mistake that this value comes from the action itself. Rather, we can see the value of an action coming from the fact that we have become accustomed to doing it.

Acknowledgements

Funding

Policy? Scope and Limits of Data-Driven Methods in Statistical Model Evaluation'
(grant number 016.104.079).

Notes

1. This is easily demonstrated for any single agent i as follows: In each round k, $P_i^{(k)} \geq (1 - \sigma_i) q_i$. Thus, the probability that the agent will remain seated after n rounds is less than or equal to $(1 - (1 - \sigma_i) q_i)^n \to 0$. Thus, the agent will eventually stand up with probability one. Since the group is finite, the group will almost surely stand up as well.
2. For ease of analysis, we report on those models in which all agents have the same parameter values. We examined mixed populations, but did not find differences that merited separate presentation.
3. For modeling simplicity, we assume that each agent counts herself as a spectator. In this way, we avoid the issue that the model would otherwise be undefined for agents in the front row.
4. The $e^{-\alpha n}$ term dominates in the long run and makes $P_i^{(n)}$ approach zero very quickly.
5. We have also explored a different extension of the model with a counter-force to the overall conformity. This second way assumes that some people increasingly resist standing up as more people stand. They act against the mainstream. We do not present this non-conformist model, as we did not find a significant deviation from the baseline model, even if this condition may be psychologically relevant.
6. This implies that σ is time- and agent-dependent, but for reasons of simplicity, we drop the subscripts in this exposition.
7. It is, however, not clear whether real social sensitivity is as discontinuous as this equation suggests. It seems more realistic that in many cases people have moderate individual preference coupled with moderate social sensitivity. So we "smooth" the function by introducing an additional parameter that governs the quickness of the transition. We did not find that this variation had a significant difference on the final result, so we only report on the discontinuous case, as the mathematics are more straightforward.
8. Young (2009) developed a model of innovation diffusion that assumes priority of one action over another, that shares some characteristics with Schelling (1971). While this model is rather elegant, it does not capture the possibility of equal footing for either norm, or the possibility of no norm emerging.
9. Recall that while descriptive norms rely on agents being motivated out of a desire to do what others do, behavioral regularities are simply cases in which individuals all perform the same action, but for independent reasons. They just all happen to prefer the same action.

References

Anderson L and Holt CA (1996) Information cascades experiments. In Plott C and Smith V (eds), *The Handbook of Results in Experimental Economics*. Amsterdam: North Holland, pp. 335–343.

Axelrod R (1984) *The Evolution of Cooperation*. New York, NY: Basic Books.

Axelrod R (1986) An evolutionary approach to norms. *The American Political Science Review* 80(4): 1095–1111.

Banerjee AV (1992) A simple model of herd behavior. *The Quarterly Journal of Economics* 107(3): 798–817.

Bernheim DB (1994) A theory of conformity. *The Journal of Political Economy* 102(5): 841–877.

Bicchieri C (2006) *The Grammar Of Society*. New York, NY: Cambridge University Press.

Bikhchandani S, Hirshleifer D and Welch I (1998) Learning from the behavior of others: conformity, fads and informational cascades. *The Journal of Economic Perspectives* 12(3): 151–170.

Binmore K (2004) Reciprocity and the social contract. *Politics, Philosophy and Economics* 3(1): 5–35.

Lewis D. (1969) *Convention: A Philosophical Study*. Cambridge, MA: Harvard University Press.

Miller J and Page SE (2004) The standing ovation problem. *Complexity* 9(5): 8–16.

Schelling TC (1971) Dynamic models of segregation. *Journal of Mathematical Sociology* 1: 143–186.

Schelling TC (1978) *Micromotives and Macrobehavior*. New York: W.W. Norton.

Skyrms B (1996) *The Evolution of Social Contract*. Cambridge, UK : Cambridge University Press.

Skyrms B (2010) *Signals: Evolution, Learning and Information*. New York, NY: Oxford University Press.

Sugden R (2000) Team preferences. *Economics and Philosophy* 16(2): 175–204.

Young HP (2009) Innovation diffusion in heterogeneous populations: contagion, social influence, and social learning. *American Economic Review* 99(5): 1899–1924.

Author biographies

Ryan Muldoon is a post-doctoral fellow in the Philosophy, Politics, and Economics Program at the University of Pennsylvania. Previously, he was a post-doctoral fellow at the Rotman Institute of Philosophy at the University of Western Ontario. His research focuses on diversity in the social contract, norm dynamics, and the division of cognitive labor.

Chiara Lisciandra is a Research Fellow at the Munich Center for Mathematical Philosophy and a Ph.D. Student in Philosophy at the Tilburg Center for Logic and Philosophy of Science. During her Ph.D. she spent a term as a visiting scholar at the Department of History and Philosophy of Science at the University of Pittsburgh. She holds a Master s degree from the University of Vita-Salute San Raffaele, Milan, which was preceded by a B.A. from the University of Milan. Her main research interests are the philosophy of science and social science and in particular, formal epistemology, the nature and dynamics of the emergence of norms, and cognitive models of norms compliance.

Cristina Bicchieri is the Sascha Jane Patterson Harvie Professor of Philosophy at the University of Pennsylvania, where she also directs the Philosophy, Politics and Economics Program. Her research focuses on behavioral research and the dynamics of social norms.

Stephan Hartmann is Chair and Head of the Munich Center for Mathematical Philosophy at Ludwig-Maximillians University. He was formerly Chair in Epistemology and Philosophy of Science in the Department of Philosophy at Tilburg University and Director of the Tilburg Center for Logic and Philosophy of Science. Previous to that he was Professor of Philosophy in the Department of Philosophy, Logic and Scientific Method at the London School of Economics and Director of LSE s Centre for Philosophy of Natural and Social Science. His current research interests include formal social epistemology, the philosophy and psychology of reasoning, no-alternatives arguments, intertheoretic relations, and probabilities in quantum mechanics.

Jan Sprenger is Resident Fellow at the Tilburg Center for Logic and Philosophy of Science and Assistant Professor at the Department of Philosophy, Tilburg University. His research focuses on philosophy of science, especially explanatory reasoning and statistical inference, formal epistemology, and group decision-making.