

Chapter 8: Regulation

At any time during the working day, I can get up from my desk, walk down to the cafeteria and find a container full of hot coffee from which I can pour myself a cup. The fact that the coffee is there waiting for me is not a consequence of careful preparation for my arrival by the cafeteria staff. I could go across to the next building, where I have never been before, walk into the basement café and sitting there waiting for me would be a similar container also filled with coffee. Not only coffee, but food, clothes, houses and everything else I need for modern living appears perfectly regulated for my needs. When I want something it is there waiting for me. When things aren't readily available—for example, nice houses being difficult to find; stocks of the latest games machine running out just before Christmas; or there being no coffee in the cafeteria at 11am—they become the subject of intense discussion about how the suppliers should act to rectify the situation. These situations are then often quickly rectified, as demand increases or decrease supply follows suit, as business try to fill gaps in the market. Consumers expect and receive supply that is regulated to suit their needs.

Regulation of supply and demand does not require central planning by me or anyone else. I don't have to call down to the café in advance ask them to switch on the percolator; the cafeteria owner doesn't have to know when the next boat of coffee beans is coming from South America; and the shipping agent doesn't need to check that new plants are already in the ground for the next year's crop. Through a series of local economic interactions I am provided with a regular supply of coffee. It was an amazement and an understanding of how unmanaged and unguided activities of humans lead to equilibriums in supply and demand that led Adam Smith to describe the economy as being guided as if by an invisible hand. For me, this amazement and understanding is best expressed in the opening sections of Thomas Schelling 's book 'Micromotives and Macrobehaviour' (1978). Schelling identifies the importance of the fact that "the market works", not just in economics, but in all forms of collective behaviour.

This chapter is about when markets work and when they don't work. More broadly it is about systems which involve regulatory feedback. Regulatory feedback, which is also referred to as negative feedback, is where a system responds in an opposite direction to a perturbation. For example, if a café has many visitors, the café employees make more coffee. If queues become too long customers stay away. The individual agents, the employees and the

customers, behave in a way that eventually leads to an equilibrium queuing time. Regulatory feedback usually, though not always, performs a balancing act that stabilises systems, bringing them to equilibrium. When it works effectively, regulatory feedback balances supply and demand, not only in our own society but also in the workings of other animal societies.

8.1 Co-operative regulation

Workers in insect societies often share a common goal of achieving a stable response to their environment, usually in the form of some optimal balance in their intake of water, sugar and other resources. For example, honey bees need to construct comb in which to store incoming nectar. Deciding when to build this comb poses a challenge for the bees, because construction is energetically costly and nectar intake is highly variable. To address this challenge, the bees only begin construction of new comb when there is both a high rate of nectar flow into the colony and the available comb drops below a threshold level (Kelley 1991; Pratt 1999). This strategy ensures that comb is available over a wide range of foraging conditions, even when sudden food 'bonanzas' become available (Pratt 1999; Pratt 2004).

Simple rules, similar to those of building when nectar flow is high and comb availability is low, are used to regulate a whole range of tasks within honey bee colonies (Seeley 1995). A major challenge in understanding this regulation is pinpointing how individual bees gather information about which tasks need to be performed. How do bees, with only a limited experience of the overall status of the colony, know when to begin a particular task? In comb building these rules remain unknown, although experiments on the construction of drone cells (comb for rearing male bees) suggest that individual builder bees may monitor the proportion of empty cells (Pratt 1998a; Pratt 1998b). In general, a number of theoretical models have emphasised how efficient division of labour, i.e. allocation of worker bees among tasks, can emerge from individuals using local information about their environment (Anderson & Ratnieks 1999; Anderson & Ratnieks 2000; Bonabeau et al. 1998a; Franks & Sendova-Franks 1997; Gordon 1996; Gordon et al. 1992). Individuals change their propensity to perform a particular task in response to the information they collect about the state of the environment and as a result the colony as a whole regulates its allocation of workers to tasks.

The entrance to a honey bee colony, often referred to as the dancefloor, is a market place for information about the state of the colony and the environment

outside the hive. Studying interactions on the dancefloor provides us with a number of illustrative examples of how individuals changing their own behaviour in response to local information allows the colony to regulate its workforce (Seeley 1995; Seeley 1997). For example, upon returning to their hive honey bees which have collected water search out a receiver bee to unload their water to within the hive. If this search time is short then the returning bee is more likely to perform a waggle dance to recruit others to the water source (Lindauer 1954). Conversely, if this search time is long then the bee is more likely to give up collecting water (Kuhnholz & Seeley 1997). Since receiver bees will only accept water if they are thirsty or require water to pass on to the bees which nurse the brood, this unloading time is correlated with the colony's overall need of water. Thus the individual's water forager's response to unloading time up- or down-regulates water collection in response to the colony's need.

Similar regulatory interactions also determine how honey bees increase the overall level of activity within the colony (Seeley et al. 1998); decide when to build drone comb (Pratt 1998b); decide whether to scout for new or exploit known food sources (Beekman et al. 2007); decide when to collect pollen (Fewell & Winston 1992); and decide whether to nurse younger or older developing larvae (Schmickl & Crailsheim 2002). Seeley (1995) gives an authoritative review of these and many other regulatory feedback mechanisms within the honey bee colony. One of the most striking examples of regulatory feedback is seen in nectar processing. When returning from a successful foraging trip, a forager bee performs either a waggle dance or a tremble dance. The choice of dance depends on the time it has searched for a bee to which to unload the nectar it has found. Waggle dances result in the recruitment of more foragers, reflecting the colony's need for more nectar, while tremble dances result in the recruitment of more nectar receivers reflecting an increased influx of nectar. Combined, these two regulatory feedbacks ensure that nectar flow is not delayed by a shortage of either foragers or receivers.

In many ant species, contacts between workers are used to regulate the division of labour. Gordon et al. (1993) showed that ants regulate their degree of antennal contacts, aggregating more when density was low and less when density was high. While these experiments were carried out in rather artificial conditions, Gordon hypothesised that contact rates could provide a general explanation of how workers regulate their division of labour (Gordon et al. 1992). Experimental evidence now supports this hypothesis. For example, the safe return of morning patrollers to a red harvester ant nest results in an increase in the

number of foragers leaving the nest to collect food (Gordon 2002; Greene & Gordon 2007a; Greene & Gordon 2007b). Similarly, ants which encounter nestmates engaged in refuse pile maintenance are more likely to engage themselves in refuse work (Gordon & Mehdiabadi 1999). In leaf cutter ants there is an additional regulation of refuse work whereby non-refuse workers aggressively force contaminated workers to remain and work in the garbage area (Hart & Ratnieks 2001; Hart & Ratnieks 2002).

Regulatory feedback is not necessarily a result of direct contacts between individuals, but can also be mediated through interactions with resources and/or pheromones. Maileux et al. (2000b) found that if *Lasius niger* ants find food that allows them to ingest a desired volume they leave pheromone trail and recruit nest-mates, but if they cannot obtain this volume they return to the nest without recruiting. This simple rule of thumb prevents recruitment of an excess of foragers to a site with only small amounts of food or where food has been depleted. This rule is extended such that when food is made up of small sub-units the ants scout locally around their first local discovery and then recruit to this site if there exist other nearby sub-units (Maileux et al. 2003b). Other species of ants may lay pheromones to prevent other ants going to areas they have already explored and failed to find food. For example, Robinson et al. (2005) found evidence that Pharaoh's ants mark unrewarding foraging paths.

In all the above examples of regulatory feedback, individuals do not have global knowledge of the state of the colony or the distribution of food. Rather they have made their own samples of the available food or the time they had to wait to unload and regulated their own behaviour in an appropriate way. For example, it may well be that a particular forager has by chance a short unloading time for a commodity that is not needed within the colony. However, if a large number of bees are simultaneously attempting to unload this same commodity, then on average their unloading time will be long and the number of individuals collecting this commodity will be down-regulated. Local sampling, performed in parallel by large numbers of individuals, allows the colony to accurately tune its average response to changes in the environment.

8.2 Over-compensation and chaos

In the above examples we can distinguish two types of regulatory feedback: passive and active. Passive regulatory feedback involves individuals adjusting

their internal probability of performing some action based on their own success. For example, the decision to abandon the utilisation of a food source is a function of an individual's own success in obtaining food there. Active regulatory feedback involves individuals producing a cue or a signal that in turn changes the probability of other individuals performing some action. For example, negative and positive pheromones are signals that respectively down- or up-regulate the number of ants taking a particular path to food. Active regulatory feedback may be either cue- or signal- based. For example, the decision by a bird to search for food in a particular place may be copied from that of other individuals (i.e. cue-based), even though the individual making the decision was completely unaware of being copied. I classify this as active feedback, even though there was no deliberate communicative action on the part of the copied individual.

Passive regulatory feedback usually results in a stable equilibrium, while the active regulatory feedback can result in over-compensation. Box 8.A describes a model which illustrates this point, showing how active regulatory feedback produces oscillations whereby the population never reaches equilibrium and can even become chaotic. In the model in Box 8.A there are three factors that are required to produce over-compensation and chaos in regulatory feedback. The first is that feedback is active, if individuals simply retire from a food source when it becomes overcrowded and return independently from each other to assess whether it is exploitable then collectively they will not over- or under-shoot the equilibrium. The second factor is that information is local. Individuals that have sampled a single food unit cannot determine whether their experience reflects the overall state of the resource. If individuals were able to assess the entire resource then they would have a fuller picture of the effects of recruitment. The last factor is that there is a time delay between the observation and the regulatory response. The generation of instability and chaos depends on discrete time steps. If these are taken away, the oscillations are dampened out.

{ [Box 8.A about here](#) }

{ [Figure 8.1 about here](#) }

{ [Figure 8.2 about here](#) }

Insect societies do sometimes over-compensate for changes in the environment. For example, when starved ants are offered food they typically recruit strongly to

it at first, leading to over-crowding at the food source (Mailleux et al. 2003b; Pasteels et al. 1987). Once the food source is overcrowded recruitment is reduced, but often not until after some individuals have arrived to find the food overcrowded (Detrain & Deneubourg 2006). Given that such over-compensation can lead to chaotic oscillations, why is active regulatory feedback so common in the interactions of social insects? A first answer to this question lies in the advantages of active feedback when a system is a long way from equilibrium. In chapter 3 we saw that positive feedback is highly effective in transmitting new information. In dynamically changing environments, positive feedback can communicate changes in the environment without requiring every individual to experience the change itself.

Box 8.A Over-compensation and chaos

Consider individual agents each of which aims to exploit a resource, such as a flower patch or a Thursday night music bar. Assume that the resource is composed of n smaller units, e.g. individual flowers within a flower patch for foraging honey bees or chairs in a bar, and further assume that each of these units can only be exploited by one individual at a time. The division of the resource in to smaller units means the information individuals obtain about the resource is local. Individuals sample one unit and have to make a decision of what to do next on the basis of their experience at this unit. We assume that within the resource, individuals choose units entirely at random. This assumption implies that, provided n is reasonably large, the number of workers choosing a particular food unit is Poisson distributed, i.e. the probability that k individuals choose resource unit is

$$p_k = \frac{(x_t / n)^k}{k!} e^{-x_t / n}$$

See, for example, Brännström & Sumpter (2005) for a derivation of the Poisson distribution in this way.

I now discuss separately two ways in which individuals might respond to their experience in deciding whether or not to go to visit the resource. Under passive regulatory feedback, we assume a constant flow α of individuals who spontaneously decide to test the resource. We assume that if two or more individuals choose the same unit then they both conclude the resource is overcrowded and decide not to try to exploit it on the next time step. We can then write the equations for the number of individuals visiting the resource x_t through time t as

$$x_{t+1} = \alpha + p_1 n = \alpha + x_t e^{-x_t / n} \equiv f(x_t) \quad (8.A.1)$$

Figure 8.1 plots a time series of iterations of equation (8.A.1) along with a cobweb diagram showing how these iterations converge on a stable equilibrium, x_* , which satisfies $x_* = f(x_*)$.

While we cannot write down a simple closed form expression for x_* , it is relatively straightforward (although I don't do it here) to show it exists and lies between 0 and n . Differentiating f with respect to x we get

$$f'(x) = (1 - x/n)e^{-x/n}$$

Thus, $-1 \leq f'(x) \leq 1$ for all $0 \leq x \leq n$ and thus the slope of $f(x)$ near to the equilibrium is less than one. This observation implies that the equilibrium is stable (for more details of the methodology underlying these conclusions see, for example, Strogatz 1994). In words, passive regulatory feedback results in the number of individuals at the resource stabilising at the unique value of x_* , independent of the parameter values α and n .

Active regulatory feedback includes some form of recruitment to the resource, e.g. pheromone trails left by ants, dances by bees, or spreading the word to friends by humans. Assume that individuals use the following two simple rules: (1) if an individual chooses a resource unit that no other individual chooses they conclude that the resource has excess capacity and recruit to the resource an average of $b-1$ other individuals who weren't previously at it; alternatively, (2) individuals choosing a unit chosen by another individual conclude that the resource is over exploited and decide not to come back on the next time step. For large n these rules give the following equation for the number of individuals visiting the resource x_t through time t as

$$x_{t+1} = bp_1 n = bx_t e^{-x_t/n} \equiv g(x_t) \quad (8.A.2)$$

Figure 8.2 plots a time series of iterations of equation (8.A.2) along with a cobweb diagram of how consecutive populations change, for different values of b .

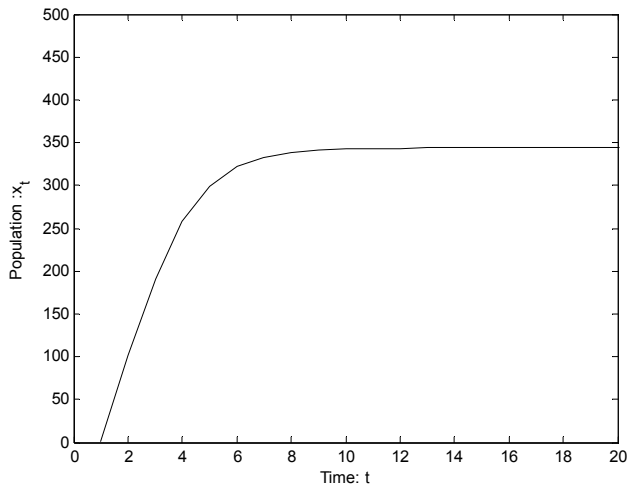
For small b , iterations of equation (8.A.2) converge on a stable equilibrium, $x_* = n \ln(b)$, which satisfies $x_* = g(x_*)$. However, differentiating g and evaluating at x_* gives

$$g'(x_*) = 1 - \ln(b)$$

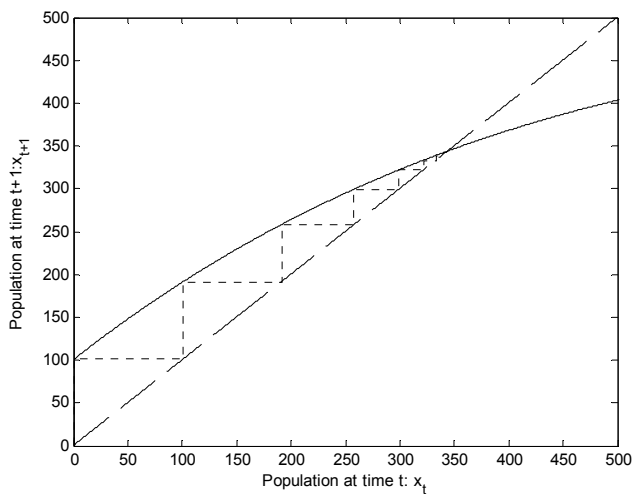
If $\ln(b) > 2$ then the slope at the equilibrium is less than -1 , implying that the equilibrium is not stable when $b > e^2$. Figures 8.2(a) and (b) give an example $b = 6$. For b slightly larger than e^2 the population will oscillate around the equilibrium $x_* = n \ln(b)$ but never come to rest there. An example of this is seen in figure 8.2(c) and (d) for $b = 8$. When the population at the resource is slightly below the equilibrium, recruitment over compensates and the population becomes larger than the equilibrium. When the population is larger than the equilibrium abandonment again overcompensates and moves to below the equilibrium. The cycle of over-compensation continues indefinitely.

Things become even less stable as b increases further. Figures 8.2(e) and (f) give an example $b = 20$. Here the population oscillates wildly, with populations sometimes large and at other times small. Far from being regulated, the population never settles to anything near to an equilibrium. In fact, equation (8.A.2), more widely known as the Ricker map, is an example of a chaotic

dynamical system (May 1976; Strogatz 1994; Sumpter & Broomhead 2001). If we were to start with two similar but slightly different population sizes, then within a few generations these differences would become amplified in a way that would make it impossible to reliably predict future population sizes. Active regulation with strong feedback leads to chaotic population dynamics. See Sumpter & Broomhead (2001) for further investigation of this model.



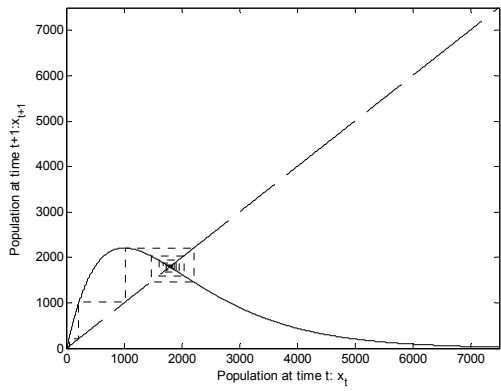
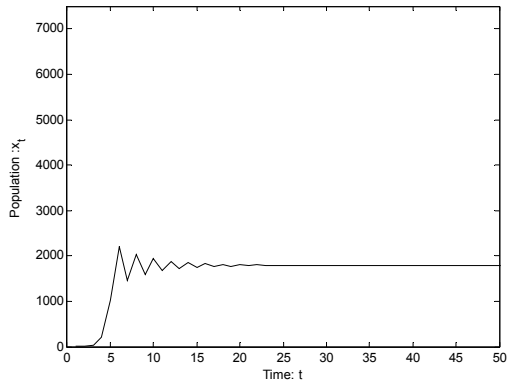
(a)



(b)

Figure 8.1: Illustration of how a population modelled by equation 8.A.1 changes through time. The parameters are $a=100$ and $n=1000$. The time series (a) shows that the population of individuals exploiting the resource equilibrates. The cobweb diagram (b) provides a plot of equation 8.A.1 (solid line) as well as the line $x_t = x_{t+1}$ (dashed line). The point x_* which satisfies $x_* = f(x_*)$ is the equilibrium population of individuals exploiting the resource. The dotted line shows how consecutive iterations of f move towards this equilibrium.

Even when positive feedback is used to actively up-regulate the number of individuals engaging in a particular task, under most natural conditions equilibrium is usually reached. Indeed, the three factors that are needed for over-compensation to occur are unlikely to be present simultaneously. In particular, positive feedback is not usually



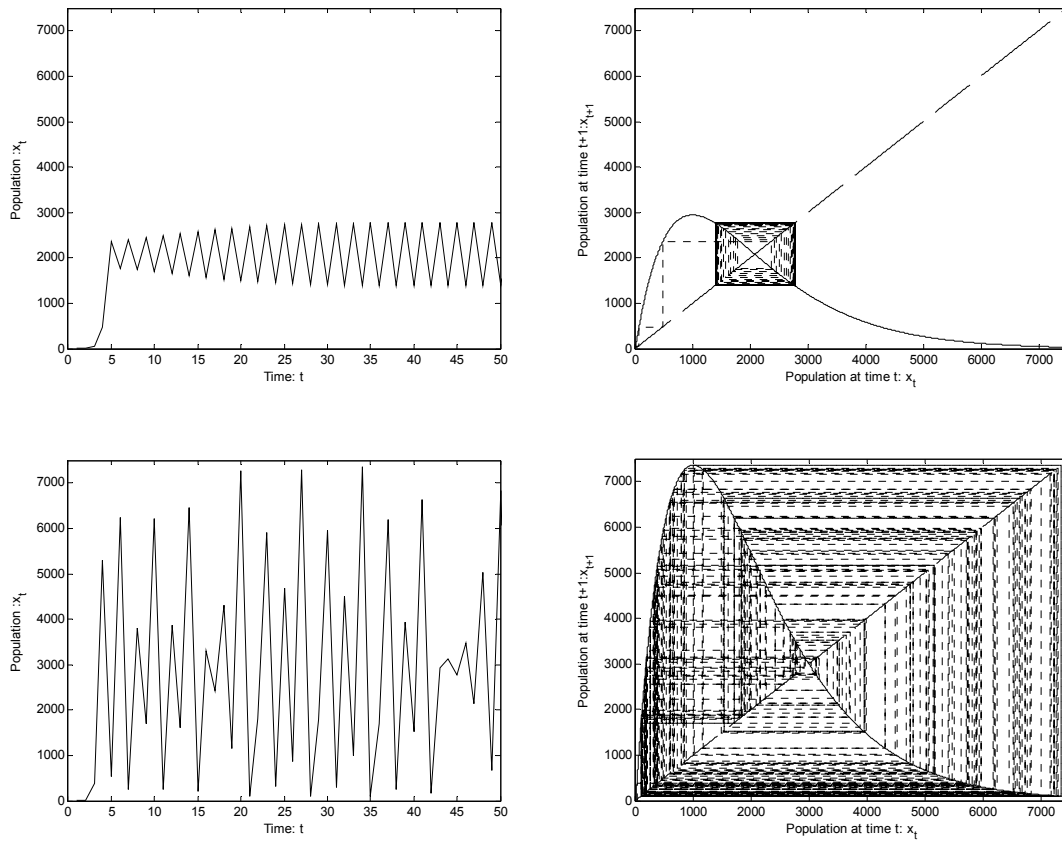


Figure 8.2: Numerical solution of how a population modelled by equation 8.A.2 changes through time. The parameters are $n=1000$ in all plots with (a,b) $b=6$, (c,d) $b=8$ and (e,f) $b=20$. The time series (a,c,e) shows how the population of individuals exploiting the resource changes through time. The cobweb diagram (b,d,f) provide a plot of equation 8.A.1 (solid line) as well as the line $x_t = x_{t+1}$ (dashed line). The point x_* which satisfies $x_* = f(x_*)$ is the equilibrium population of individuals exploiting the resource. The dotted line shows consecutive iterations of f . When $b=6$ the population equilibrates, when $b=8$ the population cycles periodically and when $b=20$ it is chaotic.

particularly strong within many insect societies. Seeley (1995) emphasises the economy of communication within honey bee colonies. His picture of a honey bee colony is one of an ensemble of individuals that rather infrequently exchange information with each other. Instead, each adjusts its behaviour in response to changes in their shared environment. Thus although active regulatory feedback is observed in honey bee colonies, it is not usually sufficiently strong to produce over compensation. Where positive feedback is strong, for example in ant foraging, other passive regulatory mechanisms, such

as retirement in response to overcrowding, operate to damp down initial over-compensation.

8.3 Selfish regulation

Honey bees and other social insects regulate variables in which they share a common interest, usually those which aid the successful reproduction of the queen. However, regulation is not limited to situations where individuals have a common interest in a particular variable reaching equilibrium. In a queue for coffee a particular equilibrium queue length is not necessarily regulated by the café's management or the customers, it emerges from their respective and different aims.

Microeconomics is primarily concerned with finding the equilibrium price of commodities, where supply equals demand. Given supply and demand curves as a function of price, the equilibrium market price is the point at which these two curves cross. Schelling (1978) emphasises that the fact that an equilibrium is reached should not be confused with the idea that the equilibrium is 'good' or in some way optimal for those involved. At a café, it may be that regulatory feedback has led to people staying away from the café because the queues were too long. These people staying away make queues shorter for everyone else, but it is perhaps not the optimal solution for the average customer or the cafe management. Much of the theory of economics is about designing markets that reach an equilibrium that optimises some criteria, be it maximising economic growth or minimising carbon emissions. Microeconomics is thus a powerful tool for solving these problems and forms the backbone of economic theory and practice (Fama 1970; Fama 1991) and is the subject of a large number of textbooks (Krugman & Wells 2004; Perloff 2007)

The theory of microeconomics is grounded in the idea of the rational agent, who attempts to maximise its own utility, which usually but not always means financial benefit. The interactions between rational agents provide a regulatory feedback that brings the market to equilibrium. If an item is over-priced, demand falls and as a result so does price. If the item is under-priced then demand exceeds supply and the price increases. Regulatory feedback on the price of a commodity brings supply and demand to equilibrium. This mechanism is not only theoretically grounded, but is consistent with our everyday experience. That prices in supermarkets are stable from week to week is testimony to market equilibrium. A store raising or lowering its prices by large amounts relative to those of its competitors would not remain in business for very long.

While prices of many everyday commodities may be stable, many economic systems are a long way from equilibrium. The growth of western economies over the last 100 hundred years has been characterised by large fluctuations over the time scale of years, where periods of accelerated growth are preceded by recession and then a return to growth (Ball 2004; Ormerod 1998). Furthermore, variations in the price of financial markets on shorter time scales of one minute to one week are power law distributed (Mantegna & Stanley 1995; Potters et al. 1998; Stanley et al. 2001). Rather than prices being normally distributed noise around an equilibrium, as would be expected in a stable market, large fluctuations in prices occur on a daily basis. Financial markets can only be considered in equilibrium if viewed on the time scale of months. Both longer and shorter time scales show non-normally distributed fluctuations.

The long time scale fluctuations, often referred to as the business cycle, pose a challenge to the rational agent hypothesis for regulation. If there exist predictable cycles in economic growth then rational agents would be able to exploit these cycles for their own financial gain, thus reducing and eventually removing the cycles. There are two main schools of thought aimed at resolving this paradox (Vercelli 1991). The first of these, following the microeconomic theory of rational agents, sees the business cycle as something which is externally generated by, for example, lags in the time between clusters of technological innovations and the capital generated by these innovations (Kydland & Prescott 1982; Lucas 1975). The alternative school of thought sees the business cycle as an intrinsic property of economic activity, explaining it either in terms of macroeconomic variables (Krugman 2005; Maynard Keynes 1936) or in terms of agents with limited information or, what is known by economists as bounded rationality (Arthur 1994; Conlisk 1996; Ormerod 1998).

A full discussion of how and why the business cycle arises is well beyond the scope of this book. For my purposes, the important observation here is the link between models of cycles in economics and mechanistic explanations of regulatory feedback. Arthur (1994; 1999) proposed a toy economic model, which he called the El-Farol bar problem, for investigating how different mechanisms might or might not lead to equilibrium. Consider a bar which has a music night every Thursday. We can define a payoff function, $f(x)$, which measures the 'satisfaction' of individuals at the bar attended by a total of x patrons. One example of such a payoff function is $f(x) = k - x$. So that individuals

going to the bar positively benefit if attendance is less than or equal to k , but would have done better to stay at home (which we assume has payoff 0) if attendance is greater than k . The El-Farol bar problem is an example of an n -player social parasitism game (section 10.2) and as such we can find the evolutionary stable strategy for bar goers. In this case the evolutionarily stable strategy is to attend the bar with probability k/n . Individuals who go to the bar any less than this risk missing a good night out, while those going more often will find it too crowded.

While a probability of visiting the bar of k/n may be evolutionarily stable, whether this equilibrium is reached depends on the mechanisms by which individuals learn whether or not they should attend the bar. We can draw a parallel between the El-Farol bar problem and the model in Box 8.A. In particular, we can see the bar as a resource and tables at the bar as resource units and think of the active feedback as individuals in the non-attending population as going to the bar if they hear that a friend went there and found a vacant table. With this interpretation, if bar goers make their decisions independently of each other then bar populations will stabilise, while if there is copying and active regulatory feedback then bar populations may over-compensate for previous observations. In particular, if the positive experiences of those attending the bar are communicated to large numbers of non-bar goers (i.e. b is large) then the population of bar goers can oscillate or become chaotic. Individuals who copy others can produce attendance levels that fluctuate wildly around the equilibrium (figure 8.2e,f).

The study of the El-Farol bar problem, and that of the related minority game (Challet & Zhang 1997), has focussed on passive regulation. The key question is how individuals that are boundedly rational, being equipped with a small number of strategies and limited memory of previous bar visits, can choose the best bar attendance strategy. On each round of bar attendance each agent adopts the strategy which would have maximised its payoff on previous rounds. When individuals have a small number of strategies to choose between and limited memory of past interactions they do not always converge to the evolutionarily stable strategy of attending the bar k/n of the time. As the memory of agents increases, the agents become more efficient and average attendance becomes close to k/n , with only small fluctuations away from this equilibrium (Challet & Zhang 1998; Savit et al. 1999). Interestingly, agents with intermediate memory produce smaller fluctuations in attendance than those with very long memory. This phenomenon occurs because individuals with a long

memory can effectively adopt the random, evolutionarily stable attendance strategy, while those with an intermediate memory cannot. The conclusion is that agents with limited memory in some cases reduce and in other cases increase a market's volatility, relative to that produced by all-knowing rational agents.

In the El-Farol bar problem there is an advantage to not following the herd, i.e. going to a bar on occasions when others are likely to attend. In this case, active regulation or copying others is unlikely to be a desirable strategy. In general however, and as we saw in chapters 3 through 5, copying can be a good strategy for making decisions in environments where information is limited. In financial markets, there is empirical evidence that financial analysts follow the buy/sell recommendations of their peers (Walter & Weber 2006; Welch 2000), although it is not clear whether this is naive herding or due to correlations in the information used by the analysts (Bernhardt et al. 2006). Copying or herding is the basis of a large number of models of financial markets (Avery & Zemsky 1998; Devenow & Welch 1996; Kirman 1993; LeBaron 2006; Lux 1995; Sornette 2003a). As such, these models are similar to those of preferential attachment (Box 2.C) and positive feedback (Box 3.A) discussed earlier in this book. The properties which emerge from these models, such as power law distributions and sudden changes in group dynamics in response to small changes in model parameters, also correspond to those observed in real markets.

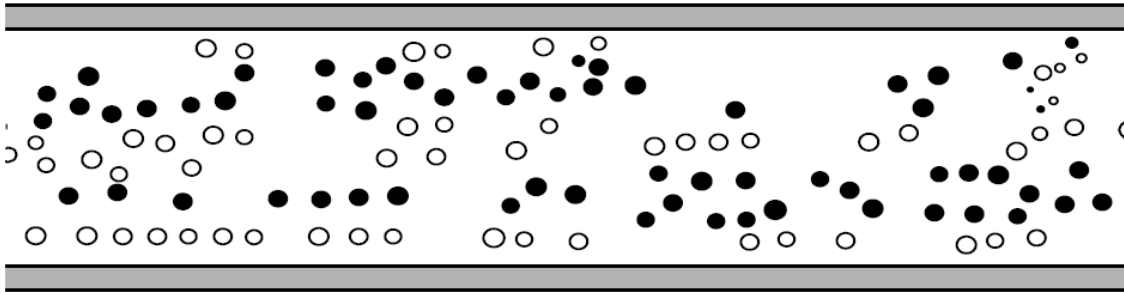
There is still a great deal of debate about the key mechanisms underlying market stability and fluctuations. However, the 'herding' models and other models arising from theoretical physics take us away from the view of the self-regulating noisy market, towards a view of a market that is sometimes stable, sometimes unstable, but always complex (LeBaron 2006; Shiller 2000; Shiller 2003; Sornette 2003b). The wealth of data available on financial markets makes them an ideal system for studying human collective behaviour, and they will continue to fascinate scientists for many years to come.

8.4 Congestion

Many species of ants form well-defined trails between food and the nest (see chapters 3 and 7). These trails are used simultaneously both by outbound and inbound ants, potentially leading to congestion on the trail and reduced traffic flow. In the leaf cutting ant, *Atta cephalotes*, encounters between ants moving in opposite directions slow their average walking speed by 16% for inbound and 21% for outbound ants (Burd & Aranwela 2003). Given the importance of rapid

delivery of food to the colony, we might expect these ant species to evolve mechanisms for efficient flow on trails.

(a)



(b)

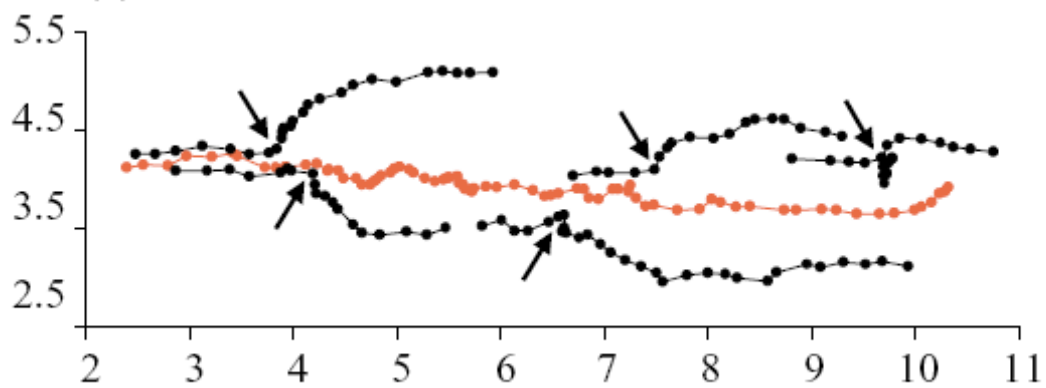


Figure 8.3. (a) Model predictions of pedestrians forming unidirectional lanes of traffic. The black and white circles represent pedestrians moving in opposite directions (reproduced from Helbing and Molnár, 1995). (b) Traffic flow on army ant trails. The figure shows the path taken by an nest-bound ant interacting with

five outbound ants. Arrows indicate the points at which the ants interacted (reproduced from Couzin & Franks, 2003).

Minimisation of collisions is also important in human pedestrian traffic. The design of walkways and safety exits involve optimally regulating the flow of humans through their environment (Batty et al. 2003b; Helbing 2001; Helbing et al. 2000; Helbing et al. 2007).

Theoretical models predict that individual ants or humans do not have to adopt particularly sophisticated movement rules in order for an efficient trail organisation to emerge. Helbing & Molnar (1995) modelled lane formation in pedestrians, using a self-propelled particle approach (see chapter 5). Each pedestrian particle was equipped with only two rules: walk in a particular direction (either left or right) and avoid collisions with others. At high densities of pedestrians moving down a corridor in opposite directions, collisions are frequent at first. Collisions are reduced for an individual pedestrian if it, through a random sequence of collision avoidances, finds itself behind another individual moving in the same direction. These 'traffic lanes' are stable, since leaving the lane will result in increased collisions. Eventually the traffic segregates into unidirectional lanes that provide highly efficient traffic flow (Helbing 2001; Helbing & Molnar 1995). An example of the resulting lane structure is shown in figure 8.3a.

Lane formation occurs on the fast moving and densely populated trails of the army ant, *Eciton burchelli*. The traffic separates into three lanes, of which the inner lane consists primarily of nestbound ants and the outer lanes of outbound ants (Couzin & Franks 2003). This three-lane formation, as opposed to the multi-lane formation predicted by Helbing & Molnar's model, results from a larger turning angle when avoiding collisions by outbound ants. These ants then move to the edge of the trail while inbound ants continue down the centre (figure 8.3b).

A related finding of Helbing and co-workers' model is that, when constrained to move through a narrow door the flow of traffic will oscillate between the two directions (figure 8.4). This temporal organisation again arises because those entering the door behind another individual are less constrained than those

attempting to enter a door from which individuals are leaving. Dussutour et al. (2005b) tested this prediction by restricting foraging *Lasius niger* ants to a bridge with a width sufficient to allow only two ants to stand side by side. Alternating clusters of inbound and outbound traffic used the bridge. The ants exhibited an additional element of co-operation, over and above that defined in the Helbing model. Ants at the bottleneck give way to ants already travelling towards them on the bridge, waiting until a gap arises in the flow and the direction switches (Dussutour et al. 2005). This temporal organisation meant that the flow rate on the narrow bridge was equal to that on a bridge more than 3 times as wide.

In ant foraging, the regulatory feedback of avoiding collisions is complemented by positive feedback provided by pheromone recruitment. Positive feedback is a form of active regulation, which increases the number of ants taking a particular route. Interactions with others act as passive regulation which prevents ants taking

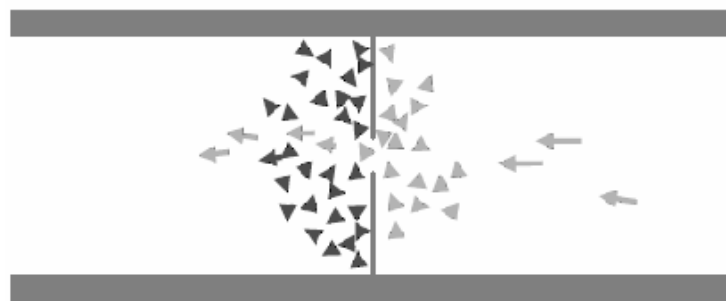
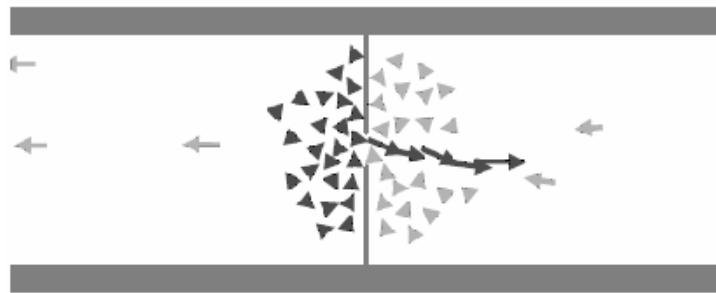


Figure 8.4: Two snapshots of a simulation of passage through a narrow door show oscillations in the build up and flow in opposite directions (reproduced from Helbing et al. 2001).

overcrowded routes. Dussutour et al (2004) took the pheromone recruitment model presented in Box 3.A and extended it by adding a term for overcrowding. In the standard model symmetry breaking bifurcation occurs, where for sufficiently high flows of ants one of the bridges was used more often than the other. With the inclusion of a term for overcrowding, however, as the width of the bridges decreased a bifurcation occurred whereby at high flow rates the bridges were exploited equally again. Dussutour et al. (2004) tested these predictions with *Lasius niger* ants and confirmed that as bridge width was narrowed the ants switched from using one to two bridges.

Ant trails are good examples of how simple strategies of avoiding collisions can lead to efficient traffic flow at the group level. However, Burd (2006) points out we should not overestimate the importance of flow of ants as a measure of efficiency of these trails. Encounters between in- and outbound ants are important sources for transfer of materials (Anderson & Jadin 2001) and, potentially, information (Burd & Aranwela 2003).

Preliminary experiments on human pedestrian traffic and 'field' observations appear to confirm many of the predictions of Helbing's models (Helbing et al. 2005; Kretz et al. 2006a; Kretz et al. 2006b; Kretz et al. 2006c). One of the most interesting outcomes of these experimental studies is that the sum of the flows in both directions of bidirectional traffic is higher than the flow of similar densities of pedestrians moving in a single direction (Helbing et al. 2005; Kretz et al. 2006a). The challenge now is to disentangle sociological explanations for phenomena,

e.g. people walk more rapidly when approaching individuals moving in the opposite direction, from explanations based on collision avoidance and interaction radii. An area where this work has had high impact and useful consequences is control at events such as football matches or religious gatherings which attract large crowds (Batty et al. 2003a; Batty et al. 2003b; Helbing et al. 2007).

8.5 Segregation and self-sorting

Humans and other animals often regulate who they interact with. Schelling proposed a series of models aimed at disentangling the mechanisms by which individuals become segregated (Schelling 1969; Schelling 1971; Schelling 1978). He uses the rather provocative example of racial segregation to illustrate his model, but his approach illuminates how individuals become sorted in everything from age and income to their hobbies and consumer preferences. One version of this model is described in Box 8.B. The model assumes a city neighbourhood made up of populations of both blacks and whites. Each individual has its own tolerance level for the ratio of whites to blacks. The tolerance level across individuals can be thought of as a distribution, with some individuals more tolerant than others. Schelling assumed that individuals will leave the neighbourhood if the race ratio is out of the limits of their tolerance and move in if the ratio is within their tolerance limit. Through this process a pattern of segregation emerges for the racial make-up of the neighbourhood.

The striking aspect of Schelling's model is that even a tolerance distribution which might initially appear relatively tolerant (in the wider meaning of the word) can lead to segregation. In the example in figure 8.5a, 50% of white individuals would accept living in a neighbourhood consisting of half whites and half blacks. The model predicts, however, that the only stable ratios of blacks to whites are single race neighbourhoods (figure 8.5c). Figure 8.5d shows that when 75% of individuals accept a 1:1 ratio then this ratio is stable, but the all white and all black neighbourhoods still remain stable if the initial populations lie nearby. The message of this model is clear: the fact that people tolerate a degree of racial mixing does not imply that segregation is avoided. Likewise, the existence of complete segregation does not imply complete intolerance on the part of the people living in the segregated neighbourhoods. Although individuals may think they are regulating their relocation decisions in a way that will generate an integrated society, the outcome can be highly segregated.

Real tolerance distributions are not as simple as those assumed in Schelling's model and they are known to differ between whites and blacks. Numerous studies based on both questionnaires and measurements of peoples' behaviour show that in the USA during the second half of the 20th century very few whites would tolerate living in an all black neighbourhood and a minority would tolerate living with more than 50% blacks. A sizeable minority of blacks would tolerate living in an all white neighbourhood and nearly all would tolerate a 50% white neighbourhood (Bruch & Mare 2006). Furthermore, in a study by Farley (1978), 79% of whites in Detroit said they would be very comfortable in neighbourhoods that were not more than 20% black.

Clark (1991) used the then available survey data to make empirical preference distributions to test Schelling's model. Analysis of four American cities showed results qualitatively similar to that of Schelling's original model: mixed race equilibriums are unstable, while all black and (nearly) all white equilibriums are stable. One difference between the empirical data and the original model is that a small number of blacks can stably inhabit a predominately white neighbourhood. Another difference is that the region where both races move into an area is smaller than that predicted in figure 8.5c. The conditions under which both whites and blacks will move into an area are very limited, and ultimately Clark (1991) predicts racial segregation.

Box 8.B Segregation model

Schelling (1969) proposed a model of racial dynamics within a city neighbourhood inhabited by people who are either black or white. He assumed that white people are willing to live with a particular ratio of blacks to whites, and this ratio is different amongst different individuals. An example of a cumulative distribution for the white's 'tolerable' race ratio is given in figure 8.5a. Under this distribution all whites tolerate being in a neighbourhood that contains only whites, half of them would happily be in a neighbourhood with a 1:1 equal ratio, and a small number of them would tolerate an almost 1:2 whites to blacks, leaving them in a minority. This tolerance distribution thus reflects differences within the population of whites for tolerating blacks in their neighbourhood.

We write W , respectively B , as the number of whites, resp. blacks, living in the neighbourhood out of a potential population of N_W , resp. N_B , whites, resp. blacks, who could choose to live in the neighbourhood. Thus W/N_W is the proportion of whites and B/N_B is the corresponding proportions of blacks living in the neighbourhood (note that the proportions are calculated relative to the potential inhabitants of each race rather than the proportions of blacks vs. whites living in the neighbourhood). We can now write an expression, that which is shown in figure 8.5a, for the proportion of whites tolerating a neighbourhood with a particular neighbourhood: $P_W = 1 - B/rW$, where $1:r$ is the ratio above which no white will move in to a neighbourhood.

There are various ways in which the tolerance distribution can influence the behaviour of individuals and thus the dynamics of the racial mix of a neighbourhood. Schelling (1969) assumed that if the actual proportion of whites living in the neighbourhood is greater than the proportion that would tolerate the current ratio, i.e. $W/N_W > P_W$, then whites would start to leave the neighbourhood, starting with the least tolerant. Similarly, if the actual proportion of whites living in the neighbourhood is less than the proportion that would tolerate the current ratio, i.e. $W/N_W < P_W$ then whites would move into the neighbourhood. Rearranging terms in these equations we see that whites will move out of a neighbourhood if $B > rW(1 - W/N_W)$ and into a neighbourhood if $B < rW(1 - W/N_W)$. The curve $B = rW(1 - W/N_W)$, plotted in figure 8.5b, gives the equilibrium at which whites will no longer move in or out of a neighbourhood.

If only whites make movement decisions on the basis of the racial makeup of a neighbourhood then the neighbourhood's composition will eventually reflect the tolerance of the white population for mixed race neighbourhoods. However, if blacks simultaneously make movement decisions based on race ratio then this outcome changes. Figure 8.5c shows the effect of whites and blacks simultaneously making movement decisions when both races have identical tolerance distributions for each other and $r=2$. When the neighbourhood at first contains a small number of whites and blacks with a near 1:1 ratio then both blacks and whites will move in to the neighbourhood. However, if the ratio is biased slightly in one direction, say with a small majority of blacks, then as the population of the neighbourhood increases whites will start to move out and more blacks will move in. As this process continues, the ratio will change to a larger majority of blacks and even more whites will move out and blacks move in, distorting the ratio still further. The stable ratio of blacks to whites is then 1:0. A similar argument applies if the initial majority are whites, with a stable ratio of 0:1. Neighbourhoods with a small black majority become all black, and those with a small white majority become all white. This situation changes when both races show a higher tolerance. Figure 8.5d shows that when $r=4$ the 1:1 ratio is locally stable, although the 1:0 and 0:1 ratios also remain locally stable. Which of these equilibrium ratios occurs depends upon the initial ratio of blacks to whites.

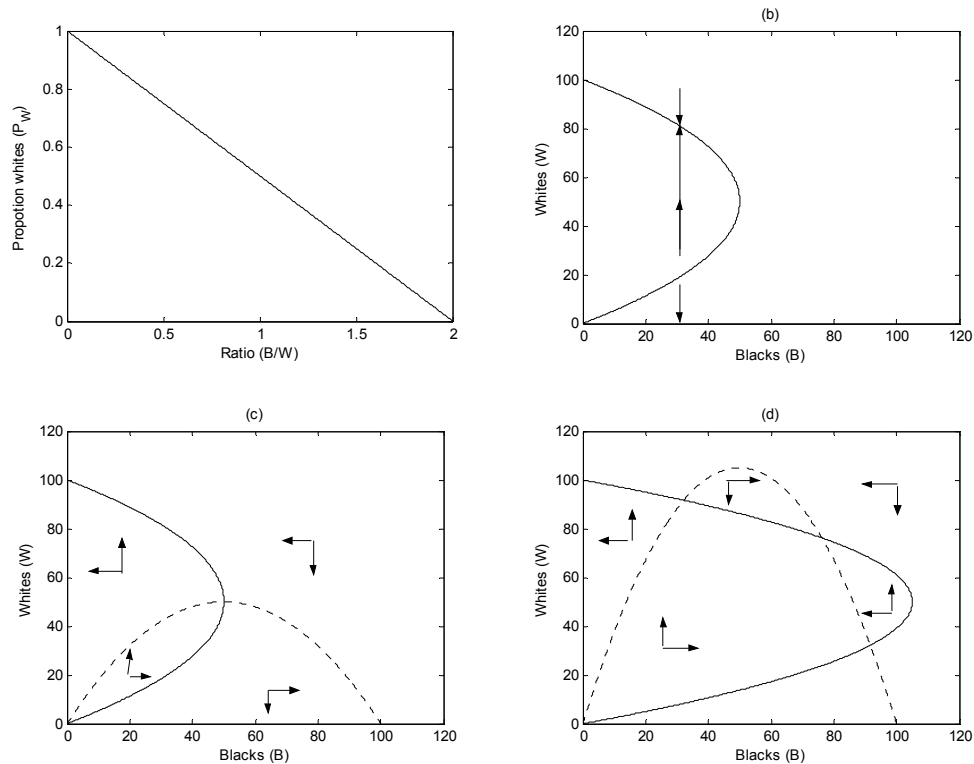


Figure 8.5: Schelling's neighbourhood segregation model. (a) Cumulative distribution for white's 'tolerable' race ratio. In this case $P_W=1-B/rW$, with $r=2$. (b) Equilibrium level at which whites neither move in or out of a neighbourhood. The arrows indicate whether for a particular population size of whites, the white population will increase (i.e. individuals move in) or decrease (i.e. individuals move out). (c) Equilibrium level at which whites neither move in or out of a neighbourhood, plotted together with the equilibrium level at which blacks neither move in or out of a neighbourhood. The arrows show the direction in which the population changes. The population stabilises at either a 1:0 or 0:1 race ratio. $r=2$ for both whites and blacks. (d) The same as c, but for $r=4$. In this case, the ratio 1:1 becomes stable.

The question that individuals were asked in Clark's study was which ratio of blacks to whites they *preferred*. In his original formulation, however, Schelling uses the term *tolerance* to denote a cut off point between wanting to leave and wanting to move into a particular neighbourhood. For example, a white individual who would ideally live in and actively seek out a neighbourhood with 50% blacks, and is neutral to living with up to 75% blacks but prefers not to live

with a higher ratio than this is said to have a tolerance of 75%. Clark's notion of preference does not account for such neutral tolerance. Studies by Farley (1978) and Farley et al. (1994) presented interview subjects with a sequence of cards on each of which was drawn a simple representation of neighbourhoods consisting of 15 houses, a proportion of which were coloured black, to represent black occupants, while the others were white. They asked whether the subjects would move into the area represented by each card, increasing the number of non-like neighbourhoods on each of the cards shown. Figure 8.6 shows the outcome of Schelling's model given the 'moving in' distributions established in these studies. In this case, the mixed equilibrium would be stable, although only marginally so, under the results of the questionnaire in 1992 (figure 8.6a) but not in 1978 (figure 8.6b). In reality, the current racial makeup of the city of Detroit, whose residents were the subject of Farley's study, is not consistent with a mixed equilibrium. The 2001 census showed that Detroit has over 80% black and less than 13% white inhabitants, with many whites having moved to the suburban outskirts.

When moving away from illustrative models of segregation and towards applications to real world data, it is important to remember the difference between correlation and causation. There is little doubt that when asked about a neighbourhood with a particular racial makeup, interview subjects base their opinion not simply on their like or dislike for a particular race but on their idea of: the quality and prices of housing and schools; crime levels; and social problems in areas that have that racial structure at the time of the interview. Disentangling whether it is race per se, or correlated variables, or even perception of correlated variables that determines people's relocation decisions is a challenging and important problem, and one about which Schelling's model says very little (Charles 2003). It is also important to bear in mind the mathematical limitations of Schelling's model. Most importantly, the model assumes a fixed population of blacks and whites which choose whether or not to live in a particular area. It doesn't say what happens to individuals who choose to live elsewhere or, due to high house prices in certain neighbourhoods, have no choice about where to live. These limitations aside, Schelling's model powerfully illustrates that weak preferences can generate strong segregation, making it all the more difficult to counteract it.

The relationship between preference at the level of the individual and aggregation or segregation at the level of the group is a complex one. For example, Rivault et al. (1998) found that different strains of cockroaches preferred the odour of their own strain. However, in experiments in which

cockroaches of the two strains were put together in an arena with two shelters, they aggregated all under the same shelter (Ame et al. 2004). In this case aggregation, instead of segregation, occurred because although each strain preferred its own odour it was weakly attracted to the odour of the other strain. This weak attraction was amplified when the strains were put together. Millor et al. (2006) showed using a model (which is a two strain extension of that in Box

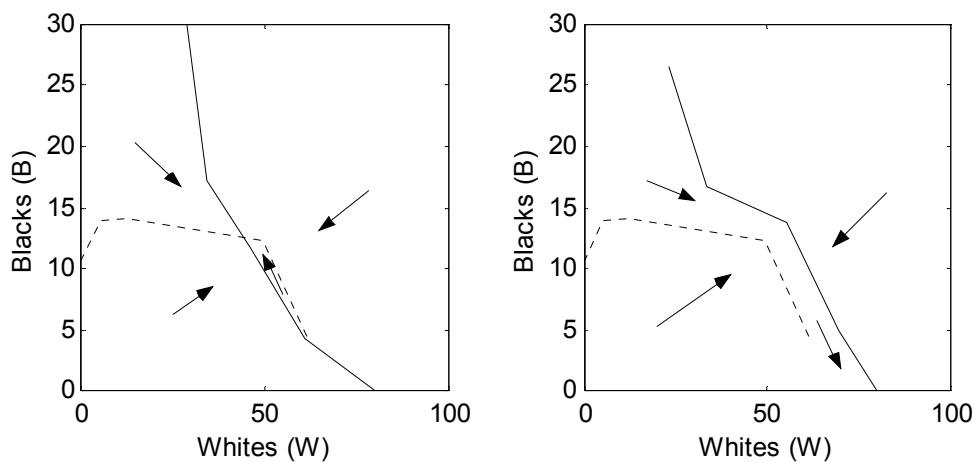


Figure 8.6: Tolerance curves from questionnaires collected from Detroit residents from (a) 1978 and (b) 1992 (Farley 1978; Farley & Frey 1994). Equilibrium level at which whites neither move in or out of a neighbourhood, plotted together with the equilibrium level at which blacks neither move in or out of a neighbourhood. The arrows show the direction in which the population changes.

3.A) that whether segregation or aggregation occurs between two strains that are weakly attracted to each other depends upon the size of the groups and the relative between- and within-strain attraction. Large groups of individuals with weak between-strain attraction will segregate, while smaller groups of individuals with strains that are more strongly attracted to each other will aggregate.

The theoretical results are partially supported by experiments which placed in the same arena equal numbers of two different species of cockroach, *Periplaneta fuliginosa* which is strongly attracted only to the odour of conspecifics and *Periplaneta Americana* which is weakly attracted to the odour of conspecifics as well as to that of *P. fuliginosa* (Leoncini & Rivault 2005). In these experiments segregation was a more common outcome than aggregation. However, smaller groups of ten cockroaches aggregated more often (in 38% of trials) than larger groups of twenty cockroaches (which aggregated in 19% of trials). The most interesting conclusion of these theoretical and empirical studies is that aggregation and segregation both arise from the same set of individual rules, with initial distribution, group size and between-group preference playing an important role in outcome. As with segregation in human societies, we should not conclude that just because two groups segregate that they are necessarily intolerant of one another.

Active regulation on the basis of particular characteristics is seen within fish shoals. Croft (2003) showed that when shoals of guppies split, they actively segregated in terms of body length. They were, however, also sorted on the basis of their response to their environment. Larger fish were more often found further from the surface of the water. These observations have interesting consequences for the social structures of shoals (Croft et al. 2006; Croft et al. 2005). Segregation leads to particular forms of social networks, which in turn determines how information flows through a population (Newman 2003; Watts & Strogatz 1998).

8.6 An Invisible hand?

Adam Smith's invisible hand which guides the economy to equilibrium remains a powerful metaphor for thinking about regulation. Many economical systems are close to equilibrium and function very efficiently without centralised control. One of the most remarkable examples of an invisible hand we have seen in this chapter is the lane formation in ants and humans. Simply by avoiding collisions, individuals self-organise in lanes which allow for efficient flow. Unfortunately, the invisible hand is not always as steady as we might hope. We have seen how active regulation can lead to over-compensation and chaos and how even a small tendency to prefer associations with like individuals can lead to strong forms of segregation. It is amazing how stable the social world is to perturbations, but it is also worth remembering how easily it can spiral out of control.