The final publication is available at Springer via
http://dx.doi.org/10.1007/s00191-016-0476-2

Evolutionary Learning and the Stability of Wage Posting Equilibria*

Robert Jump†

June 2015

Abstract

This paper presents results on the stability of the wage dispersion model presented in Mortensen (2003). Specifically, four learning processes are tested on a single parameterisation of the underlying model, and the most successful is submitted to a sensitivity analysis. The results illustrate an important problem in evolutionary dynamics first highlighted in Nelson and Winter (1982) - that to play a role in equilibrium, a strategy must be consistent with a previous disequilibrium. The results are ambiguous concerning the applicability of the equilibrium method of Mortensen (2003), as some of the learning processes are stable, whilst others are extremely unstable and exhibit complex dynamics.

Keywords: Price dispersion, Search market equilibrium, Reinforcement learning.

JEL Classification: B52, C63, D83, J31.

*Funding for this research was provided by the ESRC (grant number: ES/J500148/1). I would like to thank Mathan Satchi for all his help, and Jagjit Chadha and Paul Levine for useful comments on an earlier draft. I would also like to thank two anonymous referees and an editor for extremely useful suggestions. All remaining errors and omissions are the responsibility of the author.

†Department of Economics, Kingston University, Penrhyn Road, Kingston upon Thames, Surrey, KT1 2EE, UK. Email: r.jump@kingston.ac.uk.

1 Introduction

This paper examines the stability under learning of a simple wage posting model. Specifically, the static version of the Burdett and Mortensen (1998) wage posting model is considered, following Mortensen (2003: 16-20). Whilst a relatively small literature studies the stability under learning of dispersed price equilibria, such models are arguably not as important, or at least are not as widely used, as search models in the labour market. Given this, a natural question is whether or not the stability results of price setting models carry over to models of the labour market. If they do not, then the use of wage posting models to explain observed wage dispersion must be called into question.

The key paper examining the stability of dispersed price equilibria is Hopkins and Seymour (2002). They note a common characteristic of dispersed price equilibria, which they call the "rock-scissors-paper" property, that moves the behaviour of price setters away from the Diamond (1971) result of uniform monopoly pricing. If, for example, some consumers are relatively well informed, Bertrand-style competition allows a price setter to steal customers from a rival by charging a slightly lower price. However, eventually this strategy will lead to a price so low that it is more profitable to charge relatively badly informed consumers a high price; hence the law of one price breaks down. Hopkins and Seymour (2002) examine the price dispersion models of Varian (1980) and Burdett and Judd (1983), using a general learning process of the following form:

$$\Delta \mathbf{x}_t = Q(\mathbf{x}_t)\pi(\mathbf{x}_t). \tag{1}$$

Here, $\mathbf{x}$ is a vector whose components are the proportions of agents playing each strategy. Changes in the proportions of agents playing each strategy through time are then given by the positive-definite matrix $Q(\mathbf{x}_t)$ multiplied by the pay-off function $\pi(\mathbf{x}_t)$. In this way, relatively successful strategies expand in the population at the expense of relatively unsuccessful strategies. For a more recent paper that proceeds along these lines, see Lahkar (2011).

The process described by (1) subsumes a number of well known learning processes, including fictitious play, reinforcement learning, and imitation learning (*ibid.*: 1164). Given this, the authors are able to demonstrate the stability under learning of the dispersed price equilibria in question with a very general approach (with an important caveat regarding the proportion of informed consumers). However, the results rely on an assumption that all strategies present in the initial distribution are present in the limiting distribution - in other words, the support of the strategy set is the same as the support of the equilibrium distribution (*ibid.*: 1166). This is not the case for the model considered in this paper, as in general, the support of the strategy set will not be the same as the support of the equilibrium distribution. Whilst it is true that any distribution can be approximated by a distribution with full support (*ibid.*), there is a more general, and more interesting, problem at play. This problem is an important one in evolutionary dynamics, and appears to have first been observed by Richard Nelson and Sidney Winter:

> "There is . . . the problem that certain episodes of an industry's evolution may be characterized by negative profits for virtually all firms. For example, assume that there is a once-and-for-all drop in the demand for a product, or an increase in factor prices. Even if the profit ranking of routines were invariant

> with respect to prices, the firms that would have survived in equilibrium may drop out of business before equilibrium is achieved . . .
>
> . . . There are two analytically distinct problems here. The first is that the routines of extant firms determine, to some degree at least, the environment that selects on routines. The second is that in order to play a role in actual equilibrium, a routine must be consistent with survival in a previous disequilibrium." (Nelson and Winter 1982: 160)

In the context of price posting models, including the wage posting model examined below, all of the strategies that are optimal in equilibrium must be at least viable during the disequilibrium adjustment, and the disequilibrium adjustment itself must weed out those strategies that are not optimal in equilibrium. Given that the support of the equilibrium distribution will not be the same as the support of the strategy set, the best way to approach the learning problem would appear to be numerical simulation. This approach follows Cason et al (2005) and Waldeck and Darmon (2006). Cason et al (2005) study the stability under learning of the Burdett and Judd (1983) model, using replicator dynamics and gradient dynamics. The former is fairly standard in the learning literature, whilst the latter supposes that firms adjust their strategies according to profitability relative to a nearest neighbour. Interestingly, this dynamic adjusts to a low amplitude limit cycle around the equilibrium distribution. The replicator process, on the other hand, does not converge to the equilibrium distribution, but does converge to a stationary distribution qualitatively similar (i.e. key moments are the correct sign, in this case positive skew).

Cason et al (2005) conduct numerical simulations where the proportion of agents playing each strategy is updated recursively - that is, the model is specified in agent proportions, rather than individual agents. Waldeck and Darmon (2006), on the other hand, study the stability of the Varian (1980) price equilibrium using a fully specified microsimulation of reinforcement learning. They consider a model with 1000 buyers and 20 sellers, where each seller posts a price from a set of size 100 (buyers have a fixed sample search size, and a reservation price, and do not learn). The computational cost of such a model is relatively high, and for this reason, presumably, the authors only consider one learning algorithm. They conclude, similarly to Cason et al (2005), that the adjustment process does not converge on the equilibrium distribution, but that the limiting distribution to which the process does converge is qualitatively similar to the equilibrium (key moments are of the correct sign).

Cason et al (2005) and Waldeck and Darmon (2006) can be seen as early applications of techniques commonly used in agent based computational economics - the numerical exploration of complex systems of adaptive agents. This type of analysis is becoming increasingly popular in finance and economics, and has a particularly important link to evolutionary economics, with recent examples in Russo et al (2015) and Assenza et al (2015). As with Cason et al (2005) and Waldeck and Darmon (2006), the present paper is not an exercise in agent based modelling *per se*, but rather applies agent based modelling techniques to a model of evolutionary learning. Given this, the existing numerical studies of price dispersion models offer mixed results concerning the stability of dispersed price equilibria. Particularly, the stability of the respective equilibrium distributions appears to be sensitive to the exact learning algorithm chosen. Therefore, instead of considering a single learning process, the approach followed in this paper is to conduct numerical simulations of a number of processes. The models are specified in strategy proportions, as with Cason et al (2005), in order to reduce the computational cost of such a study. Section 2 describes the model, and four candidate learning processes. Section 3 presents the results, including an initial comparison of the learning processes, and a sensitivity analysis of the most successful. Finally, section 4 concludes, and suggests avenues for further research.

2 The Model

2.1 A Simple Wage Posting Model

The model described here is the "Pure Wage Dispersion" model of Mortensen (2003: 16 - 20), which is a simplification of the original Burdett and Mortensen (1998) model[1]. The model consists of a single period, in which m firms face n households. Each firm posts a single wage offer to a randomly chosen household, and each household accepts the highest offer received (if any). Denoting the marginal revenue product of labour as p, firms earn $p-w$ given recruitment. Given the above, the number of offers received by any given household will be binomially distributed, with "probability of success" equal to $1/n$, and "sample size" equal to m. In large markets, this can be approximated by the Poisson distribution, where $\lambda = m/n$ denotes the ratio of firms to workers, and x denotes the number of offers received:

$$\Pr(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}. \tag{2}$$

Given (2), workers are not able to choose from the whole distribution of wage offers, but equally, it is not certain that a worker will receive only one offer. This means that the law of one price will fail. To see this, consider a situation in which all firms offer $w = p$. No firm will offer a higher wage than this, as it will guarantee a loss. On the other hand, it is profitable to earn a positive expected profit by offering the reservation wage b, as the probability that the worker contacted will receive no other offer, $e^{-\lambda}$, is positive. Now consider the case where all firms offer b. In this case, it is profitable to deviate slightly by offering $b + \epsilon$, as this offer will be accepted with certainty. As the probability of the worker accepting b, given all other firms offering b, is strictly less than unity, this means that the deviating strategy is relatively profitable. This line of reasoning is just the Hopkins and Seymour argument described in the introduction, applied to wage offers rather than price offers, and relies on continuity in the strategy space.

The foregoing implies that the wage distribution will be non-degenerate, with lower support given by b. Clearly, this is a result of the inability of workers to view each firm's offer, or, equivalently, the inability of firms to contact every worker. Denoting the offer c.d.f as $F(w)$, the probability of acceptance for any given wage is equal to the probability that an offer w exceeds the x other offers received, $F(w)^x$, given the distribution of x:

$$P(F(w), \lambda) = \sum_{x=0}^{\infty} \frac{e^{-\lambda}\lambda^x}{x!} F(w)^x = e^{-\lambda[1-F(w)]}. \tag{3}$$

[1]The simplified model is similar to Butters (1977).

Given (3), the equilibrium offer distribution can be found by appealing to an equal profit condition[2]. As the reservation wage will only be accepted if it is the only offer received, the expected profit of offering the reservation wage is independent of the offer distribution, and determines the expected profit of all other offers:

$$\pi(p, w, 0) = (p - b)e^{-\lambda} = (p - w)e^{-\lambda[1-F(w)]} = \pi(p, w, F(w)). \tag{4}$$

Solving (4) for F then yields the unique offer c.d.f.,

$$F(w) = \frac{1}{\lambda} \ln \left(\frac{p - b}{p - w} \right), \tag{5}$$

which can be solved for $F = 1$ to yield the upper support of the distribution, $\bar{w}$:

$$\bar{w} = p - e^{-\lambda}(p - b). \tag{6}$$

(5) and (6) describe the wage offer distribution that results from the static Mortensen (2003) model. Note, as mentioned above, that the support of the equilibrium distribution differs from the strategy set when $p > b$ and $\lambda < \infty$ (assuming that the supremum of the strategy set is p; this seems intuitive). Although the model is, in a structural sense, extremely simple, the situation it describes is rather complex from the point of view of the individual firm. As a lower wage offer trades off a higher profit given acceptance against a lower probability of acceptance, an individual firm has to know the probability of acceptance, conditional on all wage offers, in order to make an informed decision in regards to its individual offer. Unless one assumes that firms are aware of the profit function described by (4), and thus the equilibrium offer distribution, *ex ante*, this problem does not have an obvious solution. Following the literature outlined in the introduction, the general approach is to utilise processes based on reinforcement learning and replicator dynamics, i.e. an evolutionary approach. Section 2.2 presents four different processes, which are subjected to an initial test in section 3.1 for a single fundamental parameterisation of the underlying model. The most successful process is then subjected to a sensitivity analysis in section 3.2.

2.2 Candidate Learning Processes

Consider a large number of firms, facing a large number of households, such that the ratio of firms to households is λ. The environment is analogous to that described above: each firm posts a wage offer to a randomly picked household, and each household subsequently accepts the highest offer received (if any). The marginal revenue product of a match is p, and the reservation wage b is public knowledge. The set of possible wage offers is an equally spaced grid: $S = \{w_1, ..., w_i, ..., w_l\}$, with $w_{i+1} > w_i$, $w_1 = b$, and $w_l = p$. Given the above, assume that the number of firms and households is much larger than l, such that there is,

[2] See the appendix for a detailed derivation of (3).

at any point in time, a large number of firms playing each strategy[3]. The time t average profit of each strategy is then given by the following:

$$\pi_{it}(p, w_i, \tilde{F}_t) = (p - w_i)e^{-\lambda[1-\tilde{F}_t(i)]}. \tag{7}$$

Here, $\tilde{F}_t$ denotes the cumulative distribution of firms across the strategy set at time t, which may or may not be different from the equilibrium distribution F. Likewise, $\tilde{F}_t(i)$ denotes the value of $\tilde{F}_t$ evaluated at strategy i, i.e. the proportion of firms playing strategies 1 to i inclusive at time t. The processes that are examined here borrow from the reinforcement learning literature in supposing that each strategy has a fitness measure associated with it, and that these fitness measures determine the proportion of firms playing each strategy. There are then two separate problems in constructing learning processes: first, how the proportions of firms playing each strategy are determined, given the fitness measures, and second, how those fitness measures are determined.

Learning Process:

Consider a function A that determines the proportions of firms playing each of the l strategies, and a function B that determines those strategies' fitness measures (these functions are defined in detail below). The general learning process can then be described by the following algorithm:

1. $\tilde{f}_{it} = A_a[\phi_{1t}, ..., \phi_{it}, ..., \phi_{lt}]$.
2. $\tilde{F}_t \leftarrow \tilde{f}_t$.
3. $\pi_{it} = (p - w_i)e^{-\lambda[1-\tilde{F}_t(i)]}$.
4. $\phi_{it+1} = B_b[\phi_{1t}, ..., \phi_{it}, ..., \phi_{lt}, \pi_{1t}, ..., \pi_{it}, ..., \pi_{lt}]$.

As above, $\tilde{F}_t$ denotes the cumulative distribution of firms over strategies, and $\tilde{f}_{it}$ denotes the proportion of firms playing the ith strategy, at time t. This is determined by A, which takes fitness measures as arguments, denoted ϕ_{it}. These fitness measures comprise the model's l state variables, and thus an initial distribution of fitness levels is required for the model to be fully specified. Given these values, which determine $\tilde{F}_1$, intra-period expected profits are updated for each strategy, which allows the fitness measures to be updated. This is determined by B, which allows $\tilde{f}_{t+1}$ to be calculated. Hence, the process can be iterated, given initial conditions, to examine its limiting distribution. Note that two functional forms for A and two functional forms for B are considered, hence the indices a and b above.

[3]Given this, we do not have to specify the exact number of firms and households, just the ratio λ. Note, as well, that l must be reasonably large, as the argument for the existence of the equilibrium distribution described by (5) and (6) relies on continuity. As the set of possible wage offers S in the computational model is not continuous, wage dispersion is not guaranteed a priori.

Strategy Selection Functions:

Two possible functional forms for A are considered:

$$A_1\text{: } \tilde{f}_{it} = \frac{\phi_{it}}{\sum_{i=1}^{l} \phi_{it}}.$$

$$A_2\text{: } \tilde{f}_{it} = \frac{e^{\phi_{it}/\tau}}{\sum_{i=1}^{l} e^{\phi_{it}/\tau}}.$$

Denoting by Φ the space of vectors $[\phi_{1t}, ..., \phi_{it}, ..., \phi_{lt}]$, and by Z the space of vectors $[\tilde{f}_{1t}, ..., \tilde{f}_{it}, ..., \tilde{f}_{lt}]$, A_1 and A_2 are mappings from $\Phi \rightarrow Z$. With A_1, in any given period, the proportion of agents playing strategy i is equal to the relative fitness of that strategy. This is straightforward, and follows the usual manner in which replicator dynamics and reinforcement learning determine the proportion of agents over strategies. With A_2, the proportion of agents playing strategy i is exponentially related to strategy i's relative fitness; that is, a strategy i that is twice as fit as a strategy j is played by more than twice the number of agents. The extent to which this is the case is determined by the intensity parameter τ, such that a lower τ increases the rate at which agents choose relatively profitable strategies. This is variously known as softmax selection (Sutton and Barto 1998: 30) or logit dynamics (Cason et al *op. cit.*).

Fitness Updating Functions:

Two possible functional forms for B are considered:

$$B_1\text{: } \phi_{it+1} = \begin{cases} \phi_{it} + \alpha(\pi_{it} - \phi_{it}) & \text{if } \phi_{it} + \alpha(\pi_{it} - \phi_{it}) > 0 \\ 0 & \text{if } \phi_{it} + \alpha(\pi_{it} - \phi_{it}) \leq 0 \end{cases}.$$

$$B_2\text{: } \phi_{it+1} = \begin{cases} \phi_{it} + \alpha(\pi_{it} - \bar{\pi}_t) & \text{if } \phi_{it} + \alpha(\pi_{it} - \bar{\pi}_t) > 0 \\ 0 & \text{if } \phi_{it} + \alpha(\pi_{it} - \bar{\pi}_t) \leq 0 \end{cases}.$$

Denoting by Π the space of vectors $[\pi_{1t}, ..., \pi_{it}, ..., \pi_{lt}]$, B_1 and B_2 are mappings from $\Phi \times \Pi \rightarrow \Phi$. With B_1, fitness levels are updated as per the standard reinforcement learning rule, where the fitness measure of strategy i in any period is an exponentially weighted moving average of past profitability. B_2 is known as a "reinforcement comparison" algorithm in the machine learning literature (Sutton and Barto 1998: 41), where the fitness measure of strategy i is updated by comparing that strategy's intra-period profit to the intra-period arithmetic average of all strategies' profits, $\bar{\pi}_t$. In the model considered here, this becomes a type of social learning, similar in spirit to replicator dynamics. The foregoing, by the different possible combinations of updating functions, gives four separate learning processes, which will be referred to as processes A_1B_1, A_1B_2, A_2B_1, and A_2B_2. Section 3.1 compares the convergence results of the four learning processes for a single parameterisation of the underlying model. In general, their performance is as expected, given the results of the existing literature. However, A_2B_1 is found to perform extremely well for certain parameterisations, and this is the process that is subjected to further analysis in section 3.2.

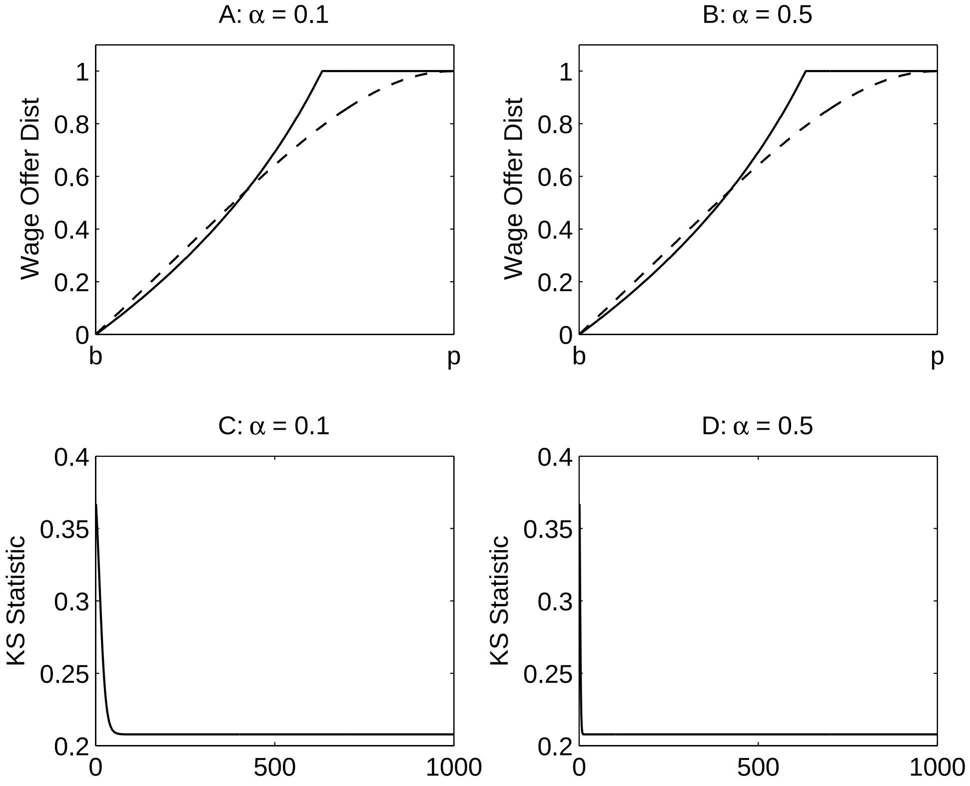

Figure 1: A_1B_1, Distributions and KS Statistics

3 Results

3.1 Comparative Analysis

The four learning processes defined above are iterated numerically to generate their limiting distributions[4]. In order to measure both the similarity between the limiting distributions and the equilibrium distribution, and the speed at which the processes converge on their respective limiting distributions, the Kolmogorov-Smirnov (KS) statistic is used (recall that F is the theoretical equilibrium distribution given by (5) and (6)):

$$KS_t = \sup_i |F(i) - \tilde{F}_t(i)|. \tag{8}$$

The KS statistic is the supremum of the set of absolute differences between F and $\tilde{F}_t$ in any given period, and the KS statistic at a process's limiting distribution is the greatest absolute difference between that distribution and the equilibrium distribution[5]. Instead of measuring "qualitative similarity" by recourse to an arbitrary set of moments, it will be defined only in a relative sense; that is, a process with a lower KS statistic at its limiting

[4]Pseudo-code for the simulations discussed in this section is presented in the appendix.

[5]Alternative distance metrics could have been used here; the KS statistic is chosen to allow comparability with the relevant literature, particularly Waldeck and Darmon (2006).

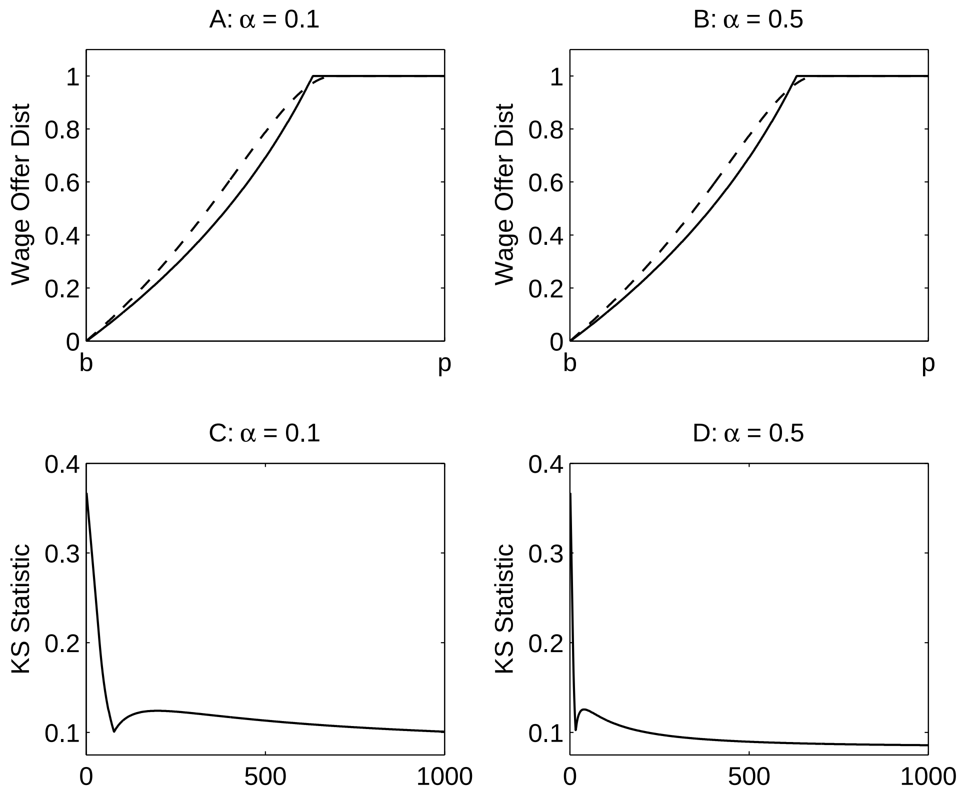

Figure 2: A_1B_2, Distributions and KS Statistics

distribution is more similar to the equilibrium distribution than a process with a higher KS statistic. In turn, as the KS statistic approaches zero, the limiting distribution will be said to "quantitatively match" the equilibrium distribution. The following four subsections describe the convergence results of the four learning processes, with the fundamental parameterisation as follows: $\lambda = 1$, $b = 1$, and $p = 2$. Although different learning parameters are considered, the size of the strategy set $l = 1000$ throughout, and initial fitness measures are equal to unity for each strategy. Note that the initial values for the fitness measures do not appear to have a significant effect on the results - this is discussed in section 3.2.

3.1.1 $\mathbf{A_1B_1}$

Panels A - D in figure 1 illustrate the convergence properties of A_1B_1. Panels A and B show the limiting distribution of the learning process (dashed line) against the equilibrium distribution (solid line) for $\alpha = 0.1$ and $\alpha = 0.5$ respectively. Panels C and D show the evolution of the KS statistic over iterations 1 - 1000 for the same parameter values. It is apparent, from visual inspection of panels A and B, that the limiting distribution of A_1B_1 is not particularly close to the equilibrium distribution, and does a particularly bad job of estimating $\bar{w}$, with no wage offers disappearing from the strategy set at all. That this is the case follows from a combination of factors. First, at the limiting distribution, each wage offer in the set S earns a non-negative expected profit. Given this, the fitness updating function B_1 means that the fitness of any strategy i cannot go to zero; combined with the

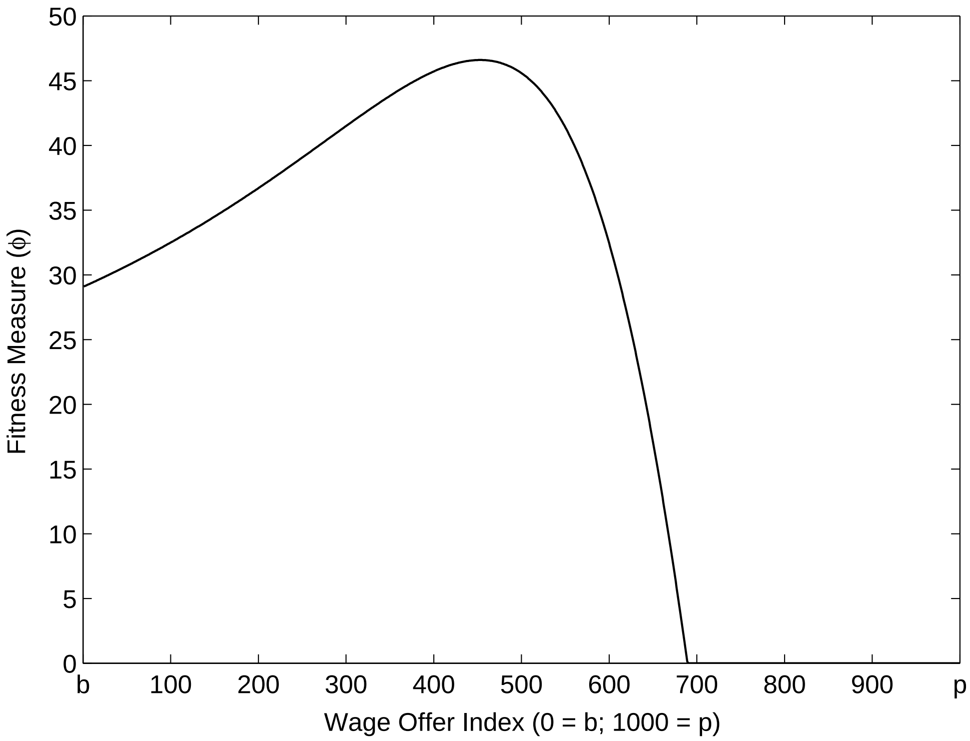

Figure 3: A_1B_2, Fitness Measures at Limiting Distribution

conservative strategy selection function A_1, this ensures that no strategy disappears from the offer distribution. Furthermore, the limiting distribution is invariant to α, the sole effect of which is to determine the adjustment speed, and the KS statistic at $t = 1000$ is approximately 0.21. Although A_1B_1 is very stable around its limiting distribution, therefore, it is not particularly successful at learning the equilibrium distribution.

3.1.2 $\mathbf{A_1B_2}$

Panels A - D in figure 2 illustrate the convergence properties of A_1B_2, for the same two parameterisations as above. As with A_1B_1, the limiting distribution is invariant to the choice of α, which governs the adjustment speed. With this learning process, however, the limiting distribution is much closer to the equilibrium distribution, with the KS statistic at $t = 1000$ approximately 0.085. Moreover, the process does a relatively good job of estimating $\bar{w}$, with the majority of wage offers above $\bar{w}$ disappearing from the limiting distribution. The increased success of this process compared to A_1B_1 is due to the potential for fitness measures to equal zero, given the fitness updating function B_2, and subsequently drop out of the offer distribution permanently, given strategy selection function A_1. Whereas, in A_1B_1, the fact that high wage offers earn a non-negative expected profit leads to those offers being present in the limiting distribution, with A_1B_2 the fact that high wage offers earn a non-negative expected profit does not lead to those offers being present in the limiting distribution, as those profit levels become considerably less than the average profit level at some point over the adjustment path, and thus their fitness falls to zero given B_2. This is further illustrated in figure 3, which plots the fitness measures over the strategy set at the limiting distribution

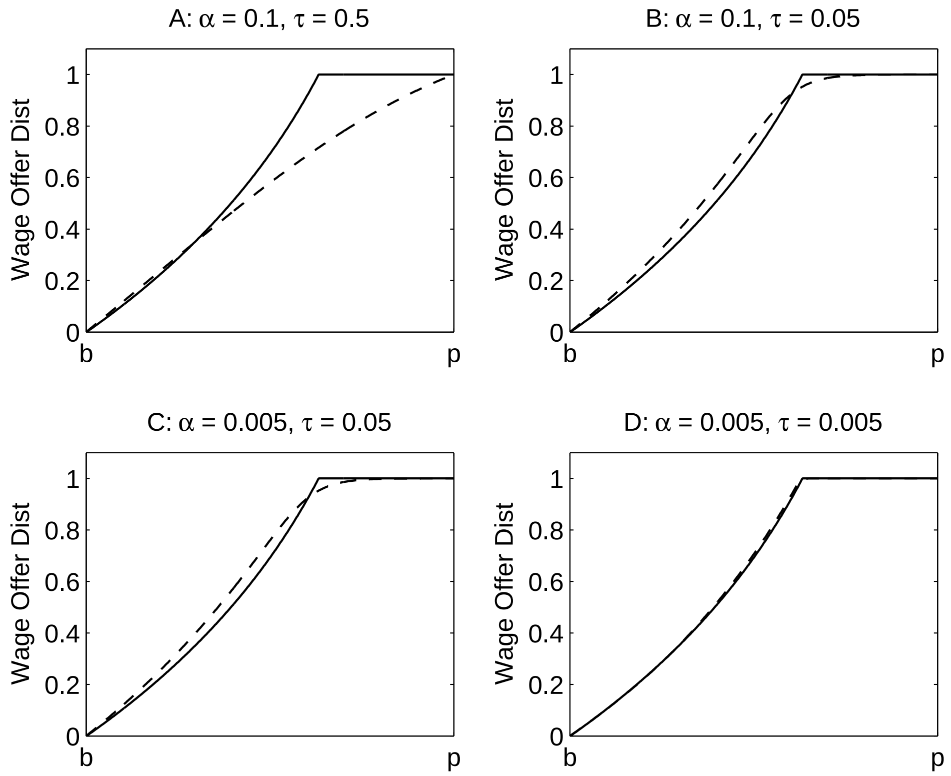

Figure 4: A_2B_1, Distributions

of A_1B_2 with $\alpha = 0.5$; each wage offer w_i in the strategy set S for which $i > 692$ has fitness measure $\phi_i = 0$, and is thus not played. Correspondingly, the wage offer distribution $\tilde{F}(i) = 1$ for all $i > 692$ in figure 2. However, where A_1B_1 smoothly approaches its limiting distribution, A_1B_2 exhibits a degree of cyclical adjustment in the KS statistic. This is indicated by the KS statistic rapidly falling to 0.1 in each parameterisation, after jumping up again and steadily approaching 0.085 over the remaining iterations. Correspondingly, A_1B_2 takes a significantly greater number of iterations to reach its minimum KS statistic than A_1B_1. Particularly, where A_1B_1 with $\alpha = 0.5$ exhibits extremely rapid convergence, A_1B_2 takes approximately 600 iterations. Despite this, A_1B_2 does a relatively good job of learning the equilibrium distribution, with particular success at estimating the upper support. This result is in line with the existing literature referred to above.

3.1.3 $\mathbf{A_2B_1}$

Unlike A_1B_1 and A_1B_2, A_2B_1 has two parameters governing the learning process: the adjustment parameter α as before, and the strategy choice parameter τ. Figure 4 shows the limiting distribution of the learning process for four parameter combinations: $\{\alpha = 0.1, \tau = 0.5\}$, $\{\alpha = 0.1, \tau = 0.05\}$, $\{\alpha = 0.005, \tau = 0.05\}$, $\{\alpha = 0.005, \tau = 0.005\}$. As before, α does not affect the limiting distribution (compare panels B and C). However, the strategy choice parameter does have an effect on the limiting distribution, and a judicious choice can greatly improve the ability of this learning process to match the equilibrium distribution. In general, it is the case that the KS statistic increases with τ; that is, the ability of A_2B_1 to learn the

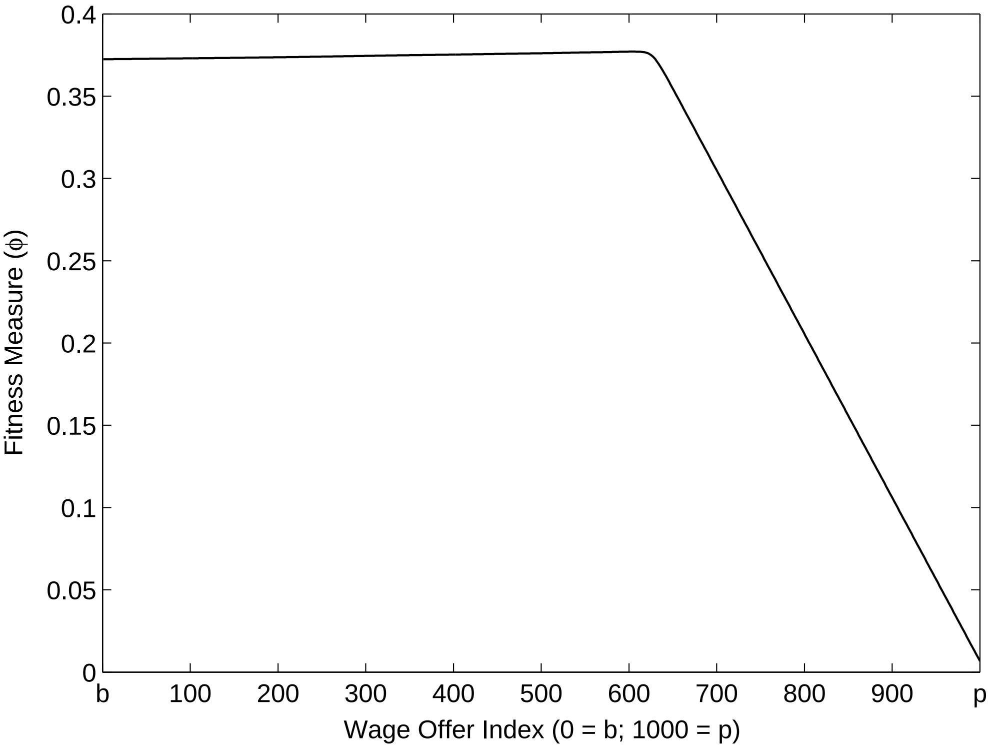

Figure 5: A_2B_1, Fitness Measures at Limiting Distribution

equilibrium distribution improves as τ decreases. This is an interesting result, as the rate of flow of agents from low to high fitness strategies increases as τ decreases. It also means that A_2B_1 works in a rather different manner to A_1B_2. Now, the combination of strategy selection function A_2 and fitness updating function B_1 means that the fitness measure of every wage offer in the strategy set S is positive in the limiting distribution, unlike in A_1B_2 where the fitness measures of high wage offers eventually fall to zero. The softmax selection of A_2, however, ensures that any wage offer with a fitness measure less than the average fitness measure is played very infrequently, and this effect increases as τ decreases. This is further illustrated in figure 5, which plots the fitness measures over the wage offer set at the limiting distribution of A_2B_1 with $\{\alpha = 0.005, \tau = 0.005\}$.

This learning process, in which fitness measures do not fall to zero, but those wage offers with relatively low fitness measures are played extremely infrequently, appears to perform better than A_1B_2, where fitness measures do fall to zero, but the remaining strategies are played in proportion to relative fitness. Unfortunately, as τ passes a certain threshold, A_2B_1 fails to converge at all, and can display extremely complex dynamics. An example is given in panel A of figure 6, which plots the KS statistic over iterations 1 - 150 for $\{\alpha = 0.5, \tau = 0.02\}$. Furthermore, the threshold value of τ at which the process becomes unstable is dependent on α, and panel B of figure 6 provides a stability plot for different combinations of α and τ, from which the relatively large region of unstable parameterisations is immediately apparent[6]. Despite this, the success of the stable parameterisations of A_2B_1 is unambiguously greater than the processes examined thus far. In fact, jointly decreasing α and τ results in an extremely low KS statistic - the parameterisation illustrated in panel D of

[6]In the stability plot, instability is defined as an absolute difference of 0.001 or more between the KS statistics at iterations 999 and 1000.

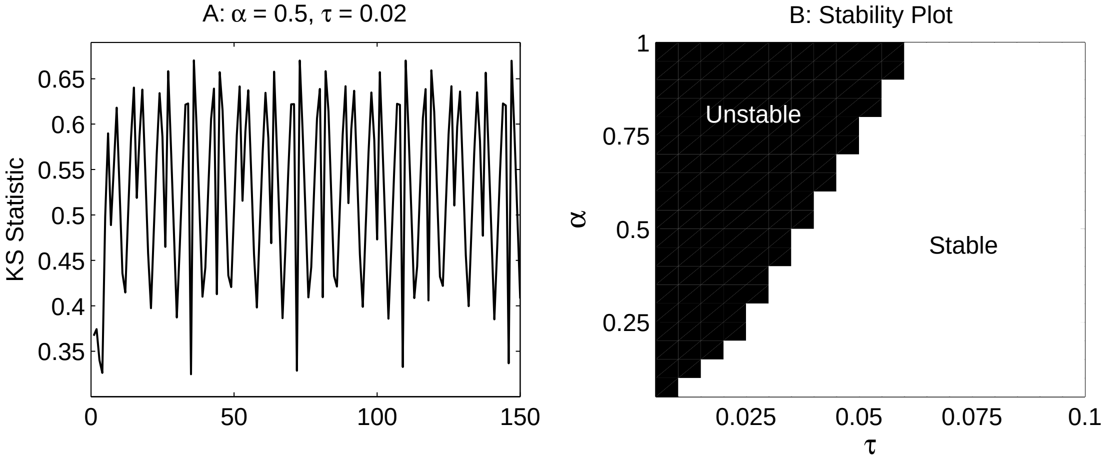

Figure 6: A_2B_1, KS Statistic and Stability Plot

figure 4, for example, has a KS statistic of 0.0115 at its limiting distribution. At this point, therefore, the tentative conclusion is that A_2B_1 has the ability to quantitatively match the equilibrium distribution described by (5) and (6).

3.1.4 $\mathbf{A_2B_2}$

The final learning process incorporates softmax selection into the reinforcement comparison fitness updating algorithm. As incorporating the former into the basic fitness updating algorithm significantly improved its performance, it might be imagined that A_2B_2 would emerge as the most successful process. Unfortunately, the tendency of A_2 towards instability appears to interact with the cyclical adjustment of B_2 in such a way that A_2B_2 does not converge to a limiting distribution for any combination of parameter values. Instead, the process produces explosive cyclical motion in the KS statistic. Panels A and B of figure 7 illustrate this for the parameterisations $\{\alpha = 0.1, \tau = 0.3\}$ and $\{\alpha = 0.5, \tau = 0.5\}$, respectively. A_2B_2, therefore, is not successful as a learning process in the context of the wage dispersion model considered here.

3.1.5 Discussion

To summarise, A_1B_1 is stable around its limiting distribution, but that distribution is relatively dissimilar to the equilibrium distribution, with a KS statistic of 0.21. At the other extreme, A_2B_2 produces explosive cyclical motion for all parameterisations, and thus fails completely as a learning process. The two relatively successful processes are A_1B_2 and A_2B_1. The former converges reliably, albeit slowly, with a KS statistic of 0.085 at its limiting distribution. In comparison, A_2B_1 can achieve an extremely low KS statistic by reducing α and τ jointly. Interestingly, this works by the strategy selection function A_2 taking considerable advantage of small differences in fitness measures, despite the fact that those fitness measures do not go to zero; this appears to work more successfully than playing strategies in proportion to fitness measures which may fall to zero. This is a delicate process, however, as after a certain threshold any reduction in τ with α fixed causes extreme instability.

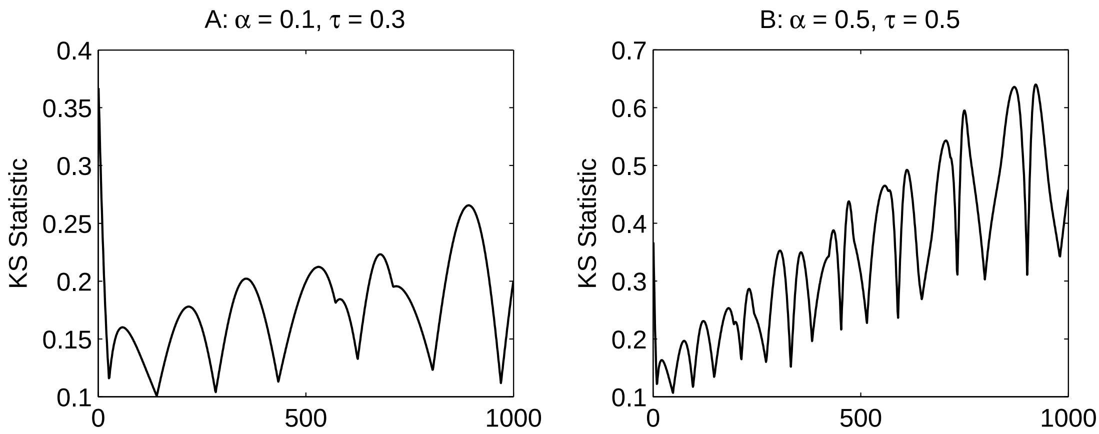

Figure 7: A_2B_2, KS Statistics

Given the above, it is not clear a priori which of the evolutionary learning processes A_1B_2 and A_2B_1 is the most successful. At this point, it is worth returning to the Nelson and Winter passage quoted in the introduction, above. In a more general process - with occasional structural change, for example - A_1B_2 may be expected to perform poorly compared with A_2B_1. This is because strategies in the former process permanently disappear from the set of possible strategies - that is, they are not necessarily "consistent with survival in a previous disequilibrium" (Nelson and Winter *op. cit.*). Whereas A_2B_1 - which works by retaining all strategies in the set of possible strategies, but playing the relatively profitably strategies extremely frequently - may be expected to perform better in a more general model. As such, and given its quantitative success over at least some of the parameter space, this is the process that is subjected to a sensitivity analysis in section 3.2. Particularly, its ability to the learn the equilibrium distribution for four different fundamental parameterisations is tested, and the region of unstable $\{\alpha, \tau\}$ combinations for those parameterisations is calculated.

3.2 $\mathbf{A_2B_1}$ Sensitivity Analysis

As demonstrated above, the KS statistic corresponding to the limiting distribution of A_2B_1 increases with τ. Given this, the threshold value of τ at which the process becomes unstable increases with α; hence for the fundamental parameterisation considered above, the KS statistic is minimised in the most south-westerly corner of the stable region of panel B in figure 6. When considering the sensitivity of the performance of this learning process, therefore, the primary question of interest is whether the region of unstable $\{\alpha, \tau\}$ combinations significantly differs for different fundamental parameterisations. The four parameterisations of the underlying model that A_2B_1 is tested against are as follows:

P1: $\lambda = 0.5$, $p = 2$, $b = 1$.

P2: $\lambda = 1.5$, $p = 2$, $b = 1$.

P3: $\lambda = 1$, $p = 3$, $b = 1$.

P4: $\lambda = 1$, $p = 2$, $b = 1.5$.

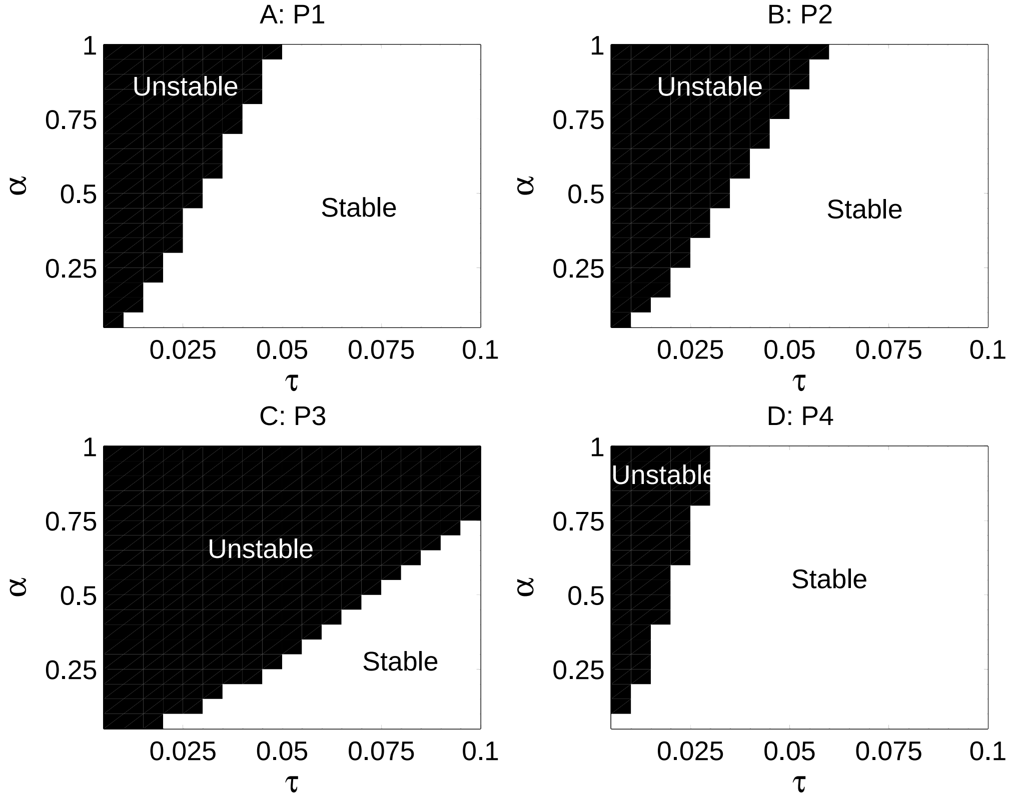

Figure 8: A_2B_1, Stability Plots

That is, in comparison to the parameterisation considered in section 3.1, P1 decreases λ, P2 increases λ, P3 increases $(p - b)$, and P4 decreases $(p - b)$. Figure 8 presents stability plots for each of these parameterisations, calculated in the same way as panel B of figure 6, above. As can be seen in panels A and B of figure 8, varying λ with p and b fixed does have an effect on the region of unstable $\{\alpha, \tau\}$ combinations, but the effect is not particularly pronounced. In contrast, varying $(p - b)$ with λ fixed has a significant effect; reducing the difference between the marginal revenue product and the reservation wage decreases the unstable region substantially, whilst increasing this difference enlarges the region. It is not clear what the economic intuition for this result is, although it is worth noting that $(p - b)$ determines the size of the strategy set, whilst λ does not, and it is possible that increasing the size of the strategy set increases the tendency towards instability. In addition, it is worth noting that the instability generated still corresponds to cyclical motion in the KS statistic, as in panel A of figure 6.

The foregoing indicates that the success of the learning process is relatively unaffected by the choice of λ, and rather more sensitive to the choice of p and b. However, for an appropriate choice of α and τ, the limiting distribution to which A_2B_1 converges can still achieve an extremely low KS statistic for both P3 and P4. This is illustrated in figure 9, which compares the limiting distribution and equilibrium distribution for P3 and P4, with $\{\alpha = 0.005, \tau = 0.01\}$ and $\{\alpha = 0.003, \tau = 0.003\}$, respectively. The KS statistic for the former is 0.0154, and the KS statistic for the latter is 0.0146. Finally, although the results are not presented here, the process is largely unaffected by the fineness of the strategy set. Reducing l substantially (e.g. < 100) does affect the stability properties, but mainly in the time taken to convergence rather than the fact of convergence itself. Increasing l past 1000,

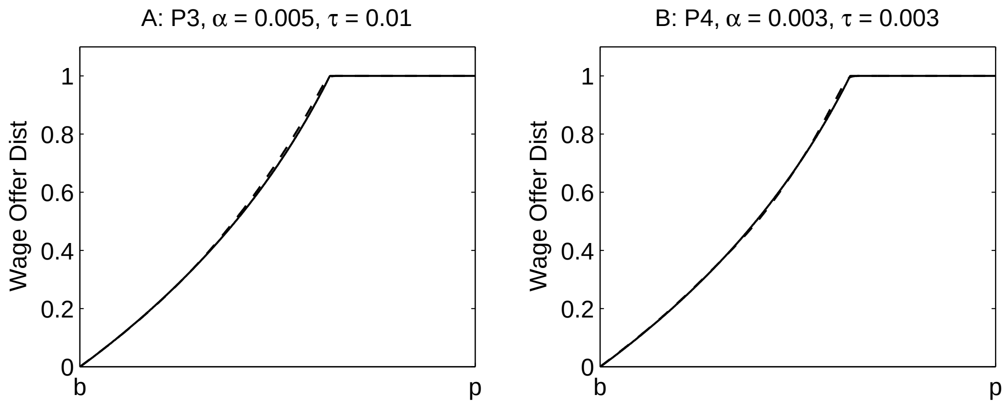

Figure 9: A_2B_1, Limiting and Equilibrium Distributions

on the other hand, has no material effect on the results. Similarly, randomising the initial distribution of fitness measures, rather than specifying an equal initial fitness measure for each strategy, appears to have no effect on the limiting distribution.

The conclusions of this section are, it is fair to say, mixed. Not only is the most successful process considered in section 3.1 unstable over a relatively large parameter region, but this region is itself affected by the fundamental parameterisation of the underlying model. An important consequence of this is that, for a given parameterisation of A_2B_1, this learning process will not dominate the alternative processes examined in section 3.1 for any random fundamental parameterisation. Thus there is an issue, so far unexamined here, concerning learning under parameter uncertainty. On the other hand, as mentioned above, this learning process may be able to deal better with occasional structural breaks, as no strategies permanently leave the set of possible strategies. Given this, it remains the case that for a judicious $\{\alpha, \tau\}$ combination, A_2B_1 can achieve an extremely low KS statistic at its limiting distribution for a variety of fundamental parameterisations. As such, the justification remains for concluding that the most successful process considered here can converge on a limiting distribution that quantitatively matches the equilibrium distribution of Mortensen's simple version of the Burdett and Mortensen (1998) wage dispersion model. Moreover, the second most successful process, A_1B_2, still does a reasonable job of learning the equilibrium distribution, and does so in a completely different manner.

4 Concluding Remarks

The four candidate learning processes considered in this chapter vary widely in their ability to learn the equilibrium distribution of Mortensen's simple wage posting model. The two basic processes, A_1B_1 and A_1B_2, are in line with results reached by Cason et al (2005) and Waldeck and Darmon (2006), that is, convergence to a limiting distribution qualitatively similar to the equilibrium distribution. The least successful process, A_2B_2, generates explosive cyclical motion for every parameterisation, which is a failure that has not been reported in the existing literature. The most successful process, on the other hand, can converge on a limiting distribution that can quantitatively match the equilibrium distribution. Although this learning process suffers from instability over a rather wide range of the parameter space, the instability appears to be bounded, and for a judicious choice of parameters the process is extremely successful.

This study highlights an important problem in evolutionary dynamics first highlighted in Nelson and Winter (1982) - that to play a role in equilibrium, a strategy must be consistent with a previous disequilibrium. To recap, the support of the equilibrium distribution of the Mortensen (2003) model is not the same as the strategy set. Given this point, two evolutionary learning processes considered in the present study have the ability to learn the equilibrium distribution of the Mortensen (2003) model relatively successfully. The first does so by eliminating a large number of strategies from the strategy set, such that the support of the process's limiting distribution is similar to the support of the equilibrium distribution, whilst the second does so by playing relatively successful strategies extremely frequently, with no strategies disappearing from the strategy set. Whilst the latter would seem to address the Nelson and Winter problematic, it does so at the expense of potential instability and complex dynamics.

The implications of these results for the Mortensen (2003) equilibrium model are ambiguous. On the one hand, learning processes have been found that approximate the equilibrium distribution relatively well, lending support to the equilibrium model as an approximation to a more general underlying model. On the other hand, the achievements of the learning processes are extremely varied, and it is not obvious which is the most empirically relevant. At the same time, some learning processes exhibit complex dynamics, and do not converge to stable limiting distributions at all. Given this, a possible line of research is to find a learning process that achieves the success of A_2B_1 without the associated instability. However, it is not a priori clear whether or not such a process exists; it may be the case that one faces a trade-off between accuracy and reliability in this type of model.

Finally, there are two interesting possibilities that would increase the empirical grounding of the model. First, the stability of more complicated, empirically relevant versions of the Mortensen (2003) model could be examined. In particular, the stability under learning of dynamic versions of the model could be studied. Second, the simple wage posting model could be conducted in a lab setting, and learning processes could be inferred from observed behaviour. As well as increasing the empirical content of the study, the observed learning behaviour might also provide evidence for processes that achieve convergence without instability. The explorations of these possibilities is left to future research.

Appendix

Deriving the Probability of Acceptance

Given the probability of offers received in (2), the probability of acceptance is equal to the probability that an offer w exceeds other offers received, $F(w)^x$, given the distribution of x. To derive (3), the following steps are required:

$$P(F(w), \lambda) = \sum_{x=0}^{\infty} \frac{e^{-\lambda}\lambda^x}{x!} F(w)^x$$

$$\Rightarrow P(F(w), \lambda) = \sum_{x=0}^{\infty} \frac{e^{-\lambda}[\lambda F(w)]^x}{x!}$$

$$\Rightarrow P(F(w),\lambda) = \sum_{x=0}^{\infty} \frac{e^{-\lambda[1-F(w)]}e^{-\lambda F(w)}[\lambda F(w)]^x}{x!}$$

$$\Rightarrow P(F(w),\lambda) = e^{-\lambda[1-F(w)]}e^{-\lambda F(w)} \sum_{x=0}^{\infty} \frac{[\lambda F(w)]^x}{x!}. \tag{9}$$

Finally, note that the infinite sum in (9) is equal to $e^{\lambda F(w)}$. So the probability of acceptance simplifies to $P(F(w),\lambda) = e^{-\lambda[1-F(w)]}$, as in (3).

Pseudo-Code for Simulations

The following pseudo-code describes the method by which the models in section 3 are simulated. Matlab code is available from the author on request.

1. Initialise parameters: T, p, b, l, α, λ.
2. Define $\bar{w}$ according to (6).
3. Define the grid of possible wage offers: $S = \{w_1, ..., w_i, ..., w_l\}$, with $w_{i+1} > w_i$, $w_1 = b$, and $w_l = p$.
4. Compute the equilibrium distribution F:
   (a) If $w_i < \bar{w}$, then $F(w_i) = \frac{1}{\lambda} \ln\left(\frac{p-b}{p-w_i}\right)$.
   (b) If $w_i \geq \bar{w}$, then $F(w_i) = 1$.
5. Simulate the model. For each time period 1 to T:
   (a) Given the fitness distribution ϕ inherited from the previous period, compute the proportions of firms over offers $\tilde{f}$ following either A_1 or A_2.
   (b) Cumulate $\tilde{f}$ to yield the offer distribution $\tilde{F}$.
   (c) Given $\tilde{F}$, compute the profit distribution $\pi(w_i) = (p - w_i)e^{-\lambda[1-\tilde{F}(i)]}$.
   (d) Given the fitness distribution ϕ inherited from the previous period, and the profit distribution π computed at step (c), compute next period's fitness distribution ϕ' following either B_1 or B_2.
   (e) Calculate this period's KS statistic: $KS = \sup_i |F(i) - \tilde{F}(i)|$.
6. Step 5 results in an l by T matrix of the offer distribution $\tilde{F}$ at each time period and a vector of KS statistics of length T. These can be used to plot the figures in the main body of the paper.

Note that the equilibrium distribution and offer distribution are defined on proportions of firms playing each strategy, rather than number of firms, hence the actual number of firms does not have to be specified (but the number of strategies, l, does have to be specified). Also note that an initial fitness distribution (fitness level per strategy) has to be specified to simulate the model (these initial fitness levels comprise the model's l state variables). Finally, note that the simulation is deterministic.

Compliance with Ethical Standards

This study was funded by the UK Economic and Social Research Council (grant number: ES/J500148/1). The author declares that no conflict of interest exists.

References

[1] Assenza, T., Delli Gatti, D., and Grazzini, J. 2015. Emergent Dynamics of a Macroeconomic Agent Based Model with Capital and Credit. *Journal of Economic Dynamics and Control*, **50**, 5-28.

[2] Burdett, K., and Judd, K. 1983. Equilibrium Price Dispersion. *Econometrica.* **51**, 955-969.

[3] Burdett, K., and Mortensen, D. 1998. Wage Differentials, Employer Size, And Unemployment. *International Economic Review.* **39**, 257-273.

[4] Butters, G. 1977. Equilibrium Distributions Of Sales And Advertising Prices. *The Review of Economic Studies*, **44**, 465-491.

[5] Cason, T., Friedman, D., and Wagener, F. 2005. The Dynamics Of Price Dispersion, Or Edgeworth Variations. *Journal Of Economic Dynamics And Control.* **29**, 801-822.

[6] Diamond, P. 1971. A Model of Price Adjustment. *Journal of Economic Theory.* **3**, 156-168.

[7] Hopkins, E., and Seymour, R. 2002. The Stability Of Price Dispersion Under Seller And Consumer Learning. *International Economic Review.* **43**, 1157-1190.

[8] Lahkar, R. 2011. The Dynamic Instability of Dispersed Price Equilibria. *Journal of Economic Theory.* **146**, 1796-1827.

[9] Mortensen, D. 2003. *Wage Dispersion: Why Are Similar Workers Paid Differently?* Cambridge MA: MIT Press.

[10] Nelson, R., and Winter, S. 1982. *An Evolutionary Theory Of Economic Change.* Cambridge, MA: Harvard University Press.

[11] Russo, A., Riccetti, L., and Gallegati, M. 2015. Increasing Inequality, Consumer Credit and Financial Fragility in an Agent Based Macroeconomic Model. *Journal of Evolutionary Economics.* DOI: 10.1007/s00191-015-0410-z.

[12] Sutton, R., and Barto, A. 1998. *Reinforcement Learning: An Introduction.* Cambridge MA: MIT Press

[13] Varian, H. 1980. A Model Of Sales. *American Economic Review.* **70**, 651-659.

[14] Waldeck, R., and Darmon, E. 2006. Can Boundedly Rational Sellers Learn To Play Nash? *Journal Of Economic Interaction And Coordination.* **1**, 147-169.