An agent-based model that relates investment in education to economic prosperity

Yuqing Tang[1], Simon Parsons[1,2], and Elizabeth Sklar[1,2]

[1] Department of Computer Science
Graduate Center, City University of New York
365, 5th Avenue, New York, NY 10016, USA
ytang@gc.cuny.edu
[2] Department of Computer & Information Science
Brooklyn College, City University of New York
2900 Bedford Avenue, Brooklyn, NY 11210, USA
{parsons,sklar}@sci.brooklyn.cuny.edu

Abstract. This paper describes some experiments with an agent-based model designed to capture the relationship between the investment that a society makes in education in one generation, and the outcome in terms of the health of the society's economy in ensuing generations. The model we use is a multiagent simulation derived from an equation-based model in the economics literature. The equation-based model is used to establish parameterized sets of agent behaviors and environmental characteristics. Agents are divided into three chronological categories: students, adults and pensioners; and each responds to and affects the environment in different ways, in terms of both human and physical capital. We explore the effects of different parameter settings on the education investment of a society and the resulting economic growth.

1 Introduction

We are working towards creating tools that can be used in determining the effects of particular choices in education policy. Our aim is to be able to use such tools to inform the debate about initiatives like the US "No Child Left Behind" Act [10], and illuminate the controversies that such initiatives have created. To this end we have been extracting predictive models from sets of data related to human education, and implementing predictive models [12, 14].

Typically, data on education is collected in one of two ways. It is either very large, aggregrate data sets over entire populations (like whole cities, school districts, states or provinces) or it is very small, localized experimental samples. In both cases, the data is usually analyzed using standard statistical methods. Often, the most highly publicized statistics are the simplest, for example the mean and standard deviation of standardized test scores in mathematics and language arts. These values are frequently the ones used to make policy decisions. More occasionally, the data is analyzed in such as way as to examine how multiple factors influence each other, such as the relationship between student-teacher ratios and test scores, dollars per student and test scores, or class size and test scores.

Where this data is extracted into models, it is formulated in traditonal terms, as sets of interelated differential equations. In contrast to such models, commonly called *equation-based models* (EBMs), we are building agent-based models (ABMs) which are constructed in terms of a set of autonomous interacting entities. Such models have been successfully used to generate useful predictions about the behavior of populations made up of individuals [11], especially where such individuals make their own decisions about how to act [15].

A particular strength of agent-based models [3] is that they allow one to identify *emergent phenomena*. Emergent phenomena result from the actions and interactions of individual agents, but are not directly controlled by the individuals. Indeed, they have an existence that is partly independent of those individuals—the classic example of an emergent phenomenon is a traffic jam, which, while caused by the actions of drivers moving in one direction, may even travel in the opposite direction.

Emergent phenomena simply do not show up in EBMs, but knowing about them can be crucial. Bonabeau [3] gives a nice example of emergent behavior with the agent-based model used by NASDAQ to identify the effects of changing some of the market rules. This model showed that reducing "tick" size (the minimum possible change in the price of an offer to trade) would lead to a larger bid-ask spread (the difference between offers to buy and offers to sell), a result that was completely counter-intuitive. Cases of emergent behavior also appear in [2, 5, 7], and in our prior work [14]. Such findings are also echoed in ecology, as in [6, 13] for example, and agent-based models have been used quite extensively in ecology where they go by the name "individual-based models". Since one can not only examine the behavior of individuals in an agent-based model, but can also look at the statistics across a population, agent-based modeling can help bridge the gap between macro and micro data sets, and thus provides the perfect tool for our work.

In this paper, we describe results of our work on one specific agent-based model, a model developed from an equation-based model published in the economics literature [9]. The model relates the effectiveness of education to economic productivity, and money spent on education to the effectiveness of education. It therefore provides a means to tie models like those we have developed in our previous work [12, 14] — models which concentrate on the education obtained by individuals — into the wider economic picture.

The remainder of this paper is organized as follows. Section 2 describes the model that we have implemented, both the equation-based original and the agent-based model we derived from it. Section 3 then describes a set of experiments that we performed using the model, Section 4 gives the results, and Section 5 analyzes them. Section 6 then concludes.

2 The model

The model that we have implemented is taken from the work of Laitner [9]. This section describes the main features of this model, and the main aspects that needed to be adapted to create an agent-based model.

2.1 Laitner's equation-based model

The setting for the model is a simple economy that has two sectors. Each of these sectors produces one good. The goods are:

- units of education that are used to train individuals in the population; and
- units of a numeraire good.

"Numeraire" is defined as "a basic standard by which values are measured, as gold in the monetary system" [4]. In [9] it is a good that is produced (see below) and then traded for things that individuals consume. Presumably these things are produced by a different economy that trades with the one we are studying.

The individuals that inhabit this economy live for three time periods, periods in which they are students, adults and pensioners. Consider an individual who is a student during period $t-1$. She spends this period living with her parent[3] and studying. Parents provide the numeraire good that supports the child during this period, but the child selects her own units of schooling, borrowing the money to finance this. In the period t, the now adult individual forms her own household, rears a child (paying for the child's consumption but not the child's schooling), and chooses how much of the numeraire good, c_t^l, that she earns during this period will be consumed by the household during the same period. In the period $t+1$, the individual is a pensioner, and chooses her consumption for that last period, c_{t+1}^2, from the numeraire good than she has saved. An individual's utility, u, is:

$$u = (1-\alpha)\ln(c_t^l) + \alpha \ln(c_{t+1}^2) \tag{1}$$

where $\alpha \in (0,1)$, and all individuals have the same α.

While working, period t in our example, our individual earns w_t per unit of human capital she possesses. Her human capital depends on her innate ability and the amount of schooling she chose as a child. An individual with ability a who purchased e_{t-1} units of education will have human capital:

$$h = a\left(\frac{(e_{t-1})^\gamma}{\gamma}\right) \tag{2}$$

where $\gamma \in (0,1)$ and all individuals have the same γ. The relationship between e and a allows education to raise human capital, but in a way that is subject to the law of diminishing returns. Innate ability is randomly assigned at the birth of individuals, with values being taken from a stationary distribution given in [9].

The model does not include inheritance and bequests, so every individual has to pay for her consumption and education out of what she earns during the period t during which she works. If r_t is the interest rate on savings made during period $t-1$ and held until period t, every individual is constrained by

$$c_t^1 + \left(\frac{c_{t+1}^2}{r_{t+1}}\right) + p_{t-1}r_t e_{t-1} \leq w_t a\left(\frac{(e_{t-1})^\gamma}{\gamma}\right) \tag{3}$$

[3] In this model, every individual has one child and raises that child alone — in terms of real world accuracy, this is equivalent to the model in [8], implemented as an agent-based model in [14], which assumes every family is a perfect nuclear family of a mother and a father and a son and a daughter.

where p_{t-1} is the cost of a unit of education in period $t-1$. In other words, the total amount that an individual consumes, including their education, suitably discounted over time, must be less that their earnings. Any earnings that are not consumed in an individual's lifetime are lost.

In every period, m individuals are born, and so there are $3m$ individuals in total in every period in time — m of these are being educated, m are working, and m are retired.

Considering the sector of the economy that produces the numeraire good, the model assumes constant returns to scale, so that the output per individual in a given generation is:

$$n = (K_t^n)^{\beta^n} (\lambda_t^n H_t^n)^{1-\beta^n} \tag{4}$$

where $\beta^n \in (0,1)$, K_t^n is the physical capital the sector has per working individual at time t, and H_t^n is the average human capital per individual in the generation that is currently working. $\lambda^n > 1$ models the tendency of technological change to increase the effect of human capital in the sector of the economy that generates the numeraire good. The other sector of the economy produces education. Here we have:

$$e = (K_t^e)^{\beta^e} (\lambda_t^e H_t^e)^{1-\beta^e} \tag{5}$$

and the model allows for λ^e and β^e to be different from, or the same as, λ^n and β^n, their counterparts in the numeraire sector of the economy.

For both the numeraire and education sectors, the assumption is that all physical capital is consumed in a single period, so the numeraire good produced in period t has to equal all consumption plus the physical capital used at time $t+1$.

Taken together, these equations and the values of the constants provided in [9] provide a rather standard economic model.

2.2 The agent-based model

As described in [3, 11, 14], it is possible to generate agent-based models from equation-based models, by equipping individuals with decision processes that make decisions in line with the equations.

For this model, the decisions faced by an individual are:

1. How much education to purchase[4].
2. What proportion of wages to save.

The first of these is, in essence, an investment decision. Given (2), for a given level of ability, the more education that an individual purchases, the greater their productivity. All other things being equal — and in particular, the performance of the other individuals in the economy — this greater productivity will turn into greater production of numeraire goods, and, once the cost of education is paid off, greater utility for the individual. Because (2) captures diminishing returns, an individual who spends too much

[4] Since every unit of education that an individual undergoes must be paid for from its later wages, it seems appropriate to think of choices about education as a purchase.

on education will not recoup their investment. The second decision is the same decision faced by anyone who has considered their own retirement — how much of one's lifetime earnings, minus cost of living and any debts accumulated, should be saved for retirement rather than spent while one is working.

In addition to these decisions, there are decisions faced by the economy as a whole. In the current version of the model, these decisions are taken by a single agent, representing the government. These are:

1. What proportion of numeraire production should be turned into physical investment.
2. What proportion of physical investment should be put into the numeraire sector rather than the education sector.
3. How to allocate workforce between the education and numeraire sectors.

The first of these decisions can be considered as the effect of taxation — some money is taken out of the income pool and is distributed by the government.

The second decision determines how much of this taxation is invested in education rather than into the production of numeraire goods — this provides the K_t^n and K_t^e in (4) and (5). Since this investment amortises over a single time-step — which is reasonable given that each time step represents a third of a lifetime, or approximately 25 years given average life expectancy in the United States — it needs to be renewed at every timestep. With this second decision under the control of some central authority, the model looks like a command economy. A more capitalist model, in which firms compete for investment from individuals and use that to provide physical investment for the numeraire sector, while leaving the government to deal with education investment, is a topic for future work.

Allowing the government to directly determine what proportion of workers to place in each sector also looks like something one would expect to find in a command economy. However, all governments exercise some control over aspects like this through their actions — in many economies the government has a large say in the organization of the education sector and can encourage people to work in the education sector by, for example, spending money raised through taxation to increase the wages of education workers.

We have implemented a number of ways that each of these decisions can be taken, and these are explored in the next section, which also gives a description of the experiments we have run.

3 Experiments

The experiments that we will describe here were intended to explore the properties of the model described in the previous section, examining whether our agent-based version could run successfully. That is, whether the decision-making functions with which we had equipped the model were sufficient to create a healthy economy and to approximate the behaviors of real economies. Before we explain how the experiments were run, we need to say what the decision functions are.

In our current implementation, individuals only have one decision to make because the proportion of wages that are saved is kept fixed. The decision they have to take, then, is how much education e_d to purchase, and the implementation provides two ways for individuals to do this:

1. Randomly: e_d is drawn from a normal distribution with mean 13 and standard deviation 2.1. This is the distribution used in [8, 14], and was originally taken from recent US census data.
2. Maximum utility: e_d is chosen by:

$$e_d = argmax_e \left(w_t \cdot a \cdot \frac{e^\gamma}{\gamma} - p_{t-1} \cdot r_t \cdot e \right) \tag{6}$$

As described above, the government has to decide:

1. $pr_{n,k}$: the proportion of numeraire production to be used as physical investment in the next period;
2. $pr_{k,e}$: the proportion of physical investment to be put into the education sector; and
3. $pr_{h,e}$: the proportion of the working population to move into the education sector.

The implementation provides several ways that each of these decisions can be made. There are three ways to decide on $pr_{n,k}$:

1. Constant proportion: $pr_{n,k}$ is set to 0.4.
2. Self adjustment: if numeraire production exceeds demand then $pr_{n,k}$ is decreased by 5%, otherwise $pr_{n,k}$ is increased by 1%.
3. Z policy: The policy that [9] uses for this decision.

Laitner's Z policy first computes an intermediate variable Z_t which describes a relationship between physical investment K_t and capital value H_t, then the policy computes an estimate of H_{t+1} from the education students have received at time t and their ability, and then computes the physical investment K_{t+1} from the estimate of Z'_{t+1}. These computations are the following:

$$Z_t = \frac{K_t}{\lambda^t \cdot H_t} \tag{7}$$

$$Z'_{t+1} = \theta \cdot (Z_t)^\xi \tag{8}$$

$$\theta = \left(\frac{\alpha \cdot (1-\gamma)}{\lambda^{1/(1-\gamma)}} \cdot \frac{1-\beta}{1+\gamma \cdot [(1-\beta)/\beta]} \right)^{1-\gamma} \tag{9}$$

$$\xi = (\beta - 1) \cdot (1-\gamma) + 1 \tag{10}$$

$$K'_{t+1} = Z'_{t+1} \cdot \lambda^{t+1} \cdot H_{t+1} \tag{11}$$

$$pr_{n,k} = \frac{K_{t+1}}{n} \tag{12}$$

The implementation includes two methods for choosing $pr_{k,e}$:

1. Constant proportion: $pr_{k,e}$ is set to 0.1.

2. Self adjustment: if education production exceeds demand, $pr_{k,e}$ is decreased by 5%, otherwise $pr_{k,e}$ is increased by 1%.

and there are two methods implemented for choosing $pr_{h,e}$:

1. Constant proportion: $pr_{h,e}$ is set to 0.1.
2. Self adjustment: if education production exceeds demand, then $pr_{h,e}$ is decreased by 5%, otherwise $pr_{h,e}$ is increased by 1%.

We ran experiments for each combination of these decision mechanisms.

4 Results

The results of these experiments, which were run over 100 timesteps, or just over 30 generations, are given in Figures 1 and 2, which show, for each economy:

- The average utility of individuals.
- The total earnings of all individuals in the economy, along with their savings for retirement, and the unpaid debt for their education.
- The education that is produced, per individual in the economy, along with the average demand for education.
- The number of numeraire goods that are produced, per individual in the economy, along with the average demand. Demand is measured by the amount of goods and individual chooses to consume, an amount that may not be satsified if the economy does not produce enough.
- The wage rates, broken down across the numeraire and education sectors.
- The number of individuals who cannot generate enough wages during their lifetime to pay for their education and their consumption as a worker or as a retiree, broken down across the numeraire and education sectors.

By all these measures, the economy in Figure 1 (Experiment 10) is healthy. The overall utility of individuals grows over time, as do wages (which reflect production). Education production flucuates over time, but fits well with demand — note that when demand exceeds supply, then individuals only receive a proportion of the education they want, and the surplus demand is spread across the population. Numeraire production grows over time. Wages in the numeraire sector grow steadily over time, as do those in the education sector, but these latter are also affected by spikes in demand. Finally, no individuals go bankrupt.

In contrast, the economy in Figure 2 (Experiment 18) is dramatically unhealthy. Once we get past the start-up effects, which are responsible, for example, for the same modest jump in average utility in both Figure 1 and 2 (note that Figure 1 (a) and Figure 2 (a) are on rather different scales), utility enters a long slump, total earnings are static while debt mounts, demand for education consistently outstrips supply by a factor of around 3, average wages have a downward trend, and after about six generations (20 timesteps) become insufficient to support the whole population — indeed after around 15 generations (40–50 timesteps) the entire population cannot meet its needs. The only apparent bright spot is that numeraire production exceeds demand, but this is because individuals do not have enough money to consume any of the goods — at the end, production is 40 times less than that in the healthy economy.

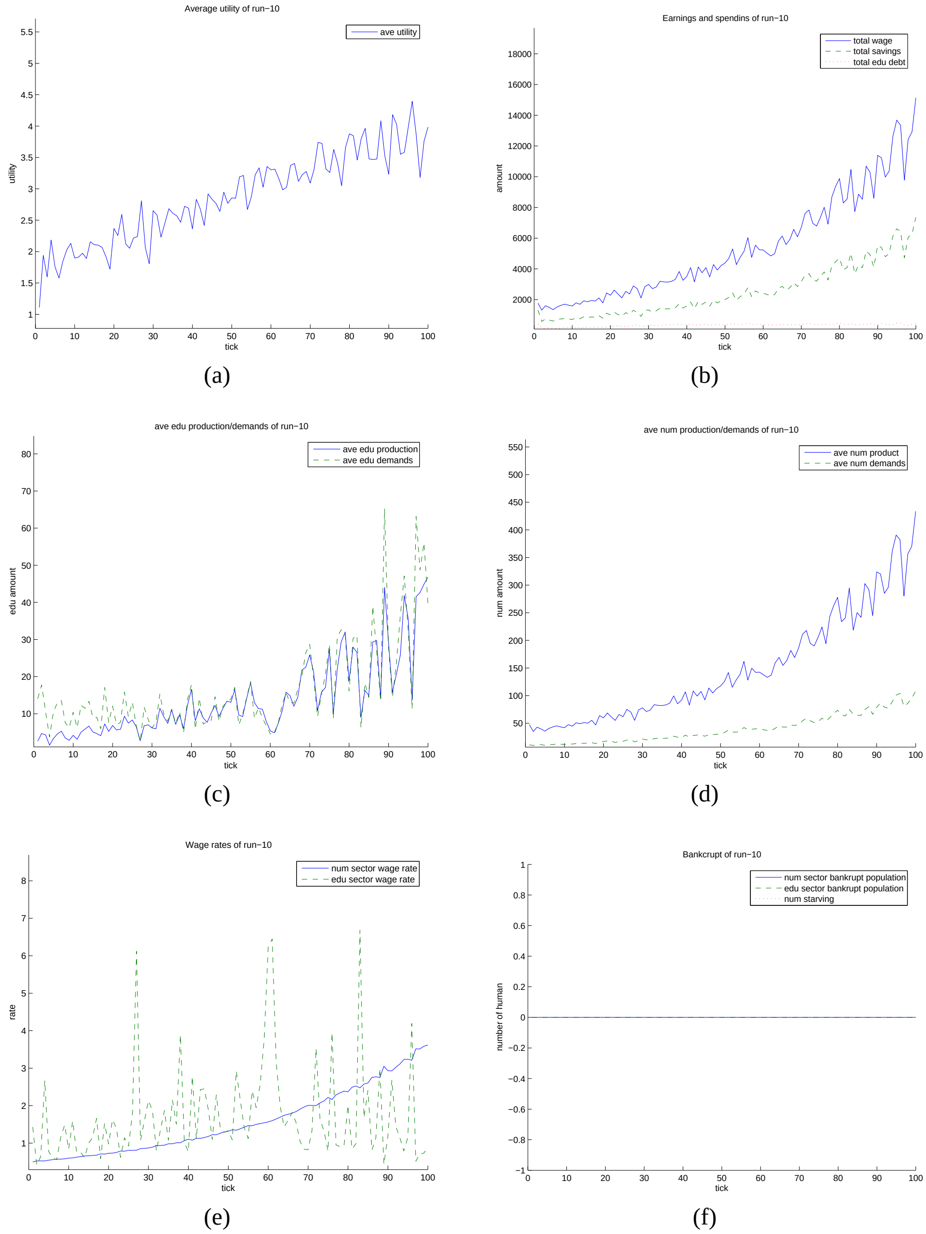

Fig. 1. Experiment 10, an example of a healthy economy under the model. (a) Average utility of individuals. (b) Earnings and savings. The solid line shows total earnings. The dashed line shows total savings. The dotted line shows debt due to education. (c) Education production per individual. The solid line shows actual production. The dashed line shows demand. (d) Numeraire production per individual. The solid line shows actual production. The dashed line shows demand. (e) Wage rates. The dashed line shows wages in the education sector. The solid line shows wages in the numeraire sector. (f) Bankruptcy. The solid line shows the number of workers in the numeraire sector who are bankrupt. The dashed line shows the corresponding number for the education sector. The dotted line shows the number of individuals who cannot afford to consume.

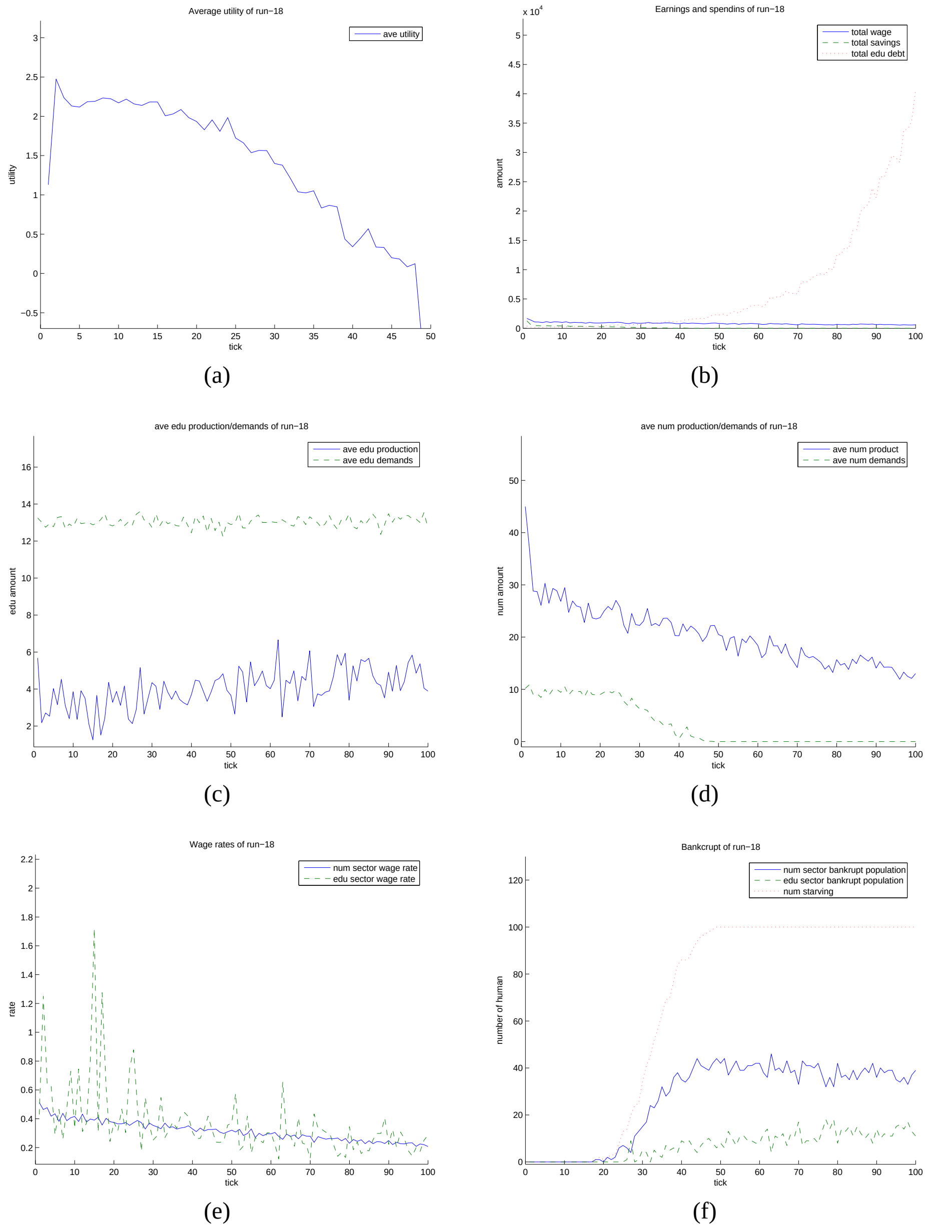

Fig. 2. Experiment 18, an example of an unhealthy economy. (a) Average utility of individuals. (b) Earnings and savings. The solid line shows total earnings. The dashed line shows total savings. The dotted line shows debt due to education. (c) Education production per individual. The solid line shows actual production. The dashed line shows demand. (d) Numeraire production per individual. The solid line shows actual production. The dashed line shows demand. (e) Wage rates. The dashed line shows wages in the education sector. The solid line shows wages in the numeraire sector. (f) Bankruptcy. The solid line shows the number of workers in the numeraire sector who are bankrupt. The dashed line shows the corresponding number for the education sector. The dotted line shows the number of individuals who cannot afford to consume.

5 Discussion

The results in the previous section are taken from only two examples of the 24 outlined in Section 3, but they are typical. To back up this claim, Figure 3 gives the average production of numeraire good and the demand for that good (which is a useful measure of economic health) against the demand and supply of education for 10 of the models. The results are presented in pairs, so Figure 3 (a) and (c) give numeraire production and education production for one model, Figure 3 (b) and (d) for the next model, and so on.

The broad trends shown in Figures 1 and 2 are repeated in these other models — the results in Figures 1 and 2 are those in Figure 3 (f) and (h) and Figure 3 (n) and (p) respectively. All of the other runs have results that fall into the same two broad classes — not only are all healthy economies healthy in exactly the same way, but all unhealthy economies are unhealthy in the same way.

The question, of course, is "why do the failing economies fail?", and it seems to us that the reason for the failure is clear from Figure 3. All the economies that fail have a consistently unmet demand for education. Over time, if economies lack the ability to educate the workforce, productivity falls, there is no basis for capital investment, and so demand for education remains unmet.

Of course, this feedback effect is written into the equation-based model, so it is no great surprise that it surfaces in the agent-based model. Indeed, we would be worried if it did not. However, note that in all the economies, even the successful ones, the demand for education initially outstrips supply. It is those economies responding to this mismatch by pumping resources into education and thus growing education production, that manage to bootstrap themselves out of the initial surplus demand for education (which will, of course, limit the productivity of the education sector since future educators themselves will be less productive if their education demands are not met). Interestingly, all the economies in Figure 3 that fail are economies that use the self-adjustment mechanism to set investment. This mechanism is much more short-term than the others, cutting investment at the first suggestion that production exceeds demand. It is tempting to interpret the failure of this approach in the models depicted as a failure for short-termism in economic policy, but we need to run more experiments before we can be confident in making any judgement on this.

6 Summary

This paper has described the creation of an agent-based model of an economy from an equation-based model, and the results of some experiments intended to establish the behavior of the model under a range of conditions. These experiments have shown that the model tightly couples investment in education to production, and, through production, to the overall health of the economy.

Our next step with this model is to extend it towards the policy evaluation tool that we described in the introduction. To do this, we first envisage combining it with the model we described in [14] — an agent-based model that was developed from the equation-based model in [8]. The model in [14] will give us a mechanism that individuals use to determine the level of education that they desire (a level that is based on that

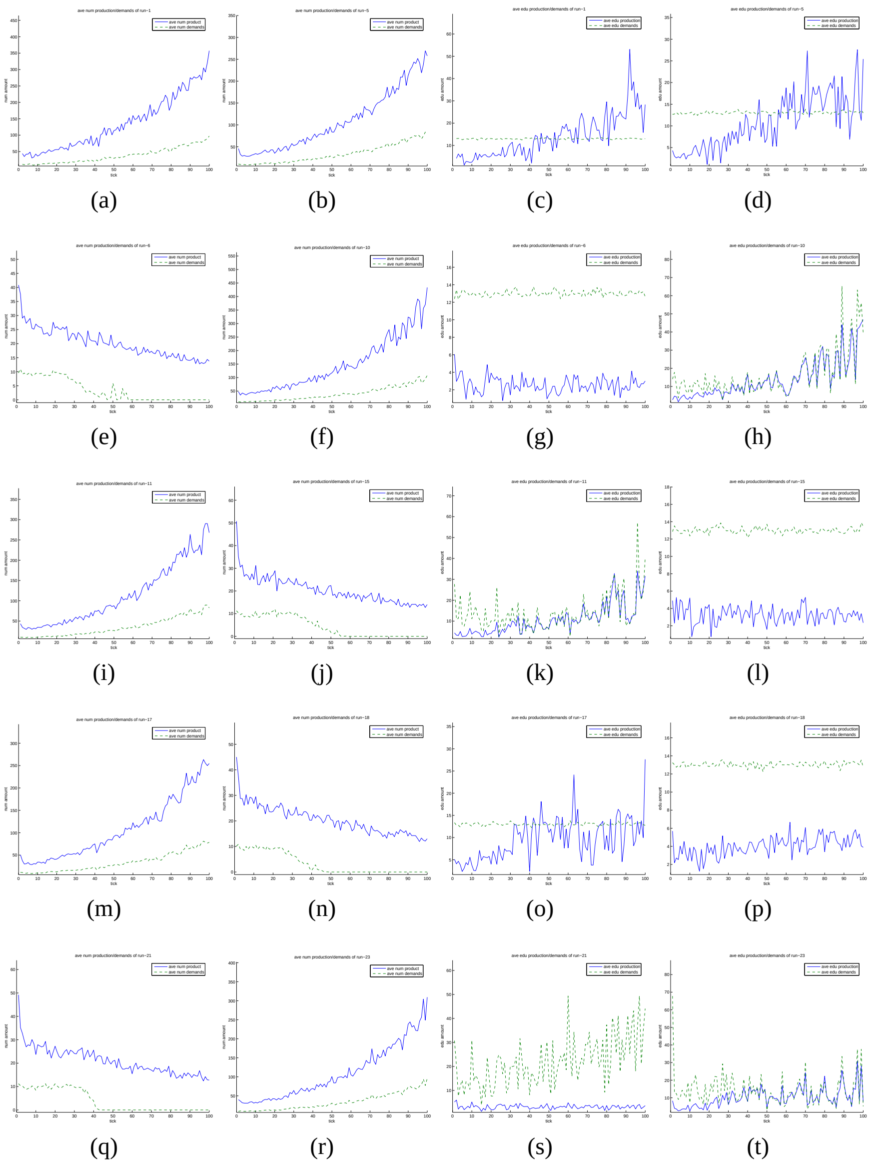

Fig. 3. The relationship between the education produced by the economy and the production of the numeraire good in selected experiments. In all graphs, demand is given by a dotted line, supply by a solid line. The left two columns give the numeraire production, the right two columns the education production. (a) and (c) are taken from the same economy, as are (b) and (d), and so on.

of their parents), a model that, as [8] describes, is a good fit for real data. With that done, we want to couple in models like that in [12] which relate policy changes in education, like class size, to the quality of education that is provided.

Acknowledgments

This work was partially supported by NSF #ITR-02-19347 and by PSC-CUNY #68525-00-37.

References

1. L. Antunes and K. Takadama, editors. *Multi-Agent-Based Simulation VII, International Workshop, MABS 2006, Hakodate, Japan, May 8, 2006, Revised and Invited Papers*, volume 4442 of *Lecture Notes in Computer Science*. Springer Verlag, 2007.
2. R. N. Bernard. Using adaptive agent-based simulation models to assist planners in policy development: The case of rent control. Working Paper 99-07-052, Sante Fe Institute, 1999.
3. E. Bonabeau. Agent-based modelling: Methods and techniques for simulating human systems. *Proceedings of the National Academy of Science*, 99(3):7280–7287, May 2002.
4. S. B. Flexner, editor. *The Random House Dictionary of the English Language*. Random House, New York, NY, 1987.
5. A. Greenwald and J. Kephart. Shopbots and pricebots. In *Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence*, pages 506–511, Stockholm, Sweden, August 1999.
6. V. Grimm. Ten years of individual-based modeling in ecology: what have we learned and what could we learn in the future? *Ecological Modeling*, pages 129–148, 1999.
7. J. Kephart and A. Greenwald. Shopbot economics. *Autonomous Agents and Multi-Agent Systems*, 5(3):255–287, 2002.
8. M. Kremer. How much does sorting increase inequality? *The Quarterly Journal of Economics*, 112(1):115–139, Feb. 1997.
9. J. Laitner. Earnings within education groups and overall productivity growth. *The Journal of Political Economy*, 108(4):807–832, Aug. 2000.
10. No Child Left Behind. `http://www.whitehouse.gov/news/reports/no-child-left-behind.html`, July 2001.
11. H. Van Dyke Parunak, Robert Savit, and Rick L. Riolo. Agent-based modeling vs. equation-based modeling: A case study and users' guide. In *Proceedings of the Workshop on Multiagent-based Simulation*, pages 10–25, 1998.
12. Elizabeth Sklar, Mathew Davies, and Min San Tan Co. SimEd: Simulating Education as a MultiAgent System. In *Proceedings of the Third International Conference on Autonomous Agents and MultiAgent Systems (AAMAS-2004)*, pages 998–1005, 2004.
13. R. V. Sole, J. G. P. Gamarra, M. Ginovart, and D. Lopez. Controlling chaos in ecology: From deterministic to individual-based models. *Bulletin of Mathematical Biology*, 61:1187–1207, 1999.
14. Y. Tang, S. Parsons, and E. Sklar. Agent-based modeling of human education data. In *Proceedings of the Fifth International Conference on Autonomous Agents and MultiAgent Systems (AAMAS-2006)*, 2006. An extended version of this paper appeared in [1].
15. M. Wooldridge, N. R. Jennings, and D. Kinny. The Gaia methodology for agent-oriented analysis and design. *Autonomous Agents and Multi-Agent Systems*, 3(3):285–312, 2000.