

(THE EVOLUTION OF)
POST-SECONDARY EDUCATION:
A COMPUTATIONAL MODEL
AND EXPERIMENTS

Andreas Ortmann
Sergey Slobodyan

CERGE-EI

Charles University
Center for Economic Research and Graduate Education
Academy of Sciences of the Czech Republic
Economics Institute

Working Paper Series
(ISSN 1211-3298)

355

**(The Evolution of)
Post-Secondary Education:
A Computational Model
and Experiments**

Andreas Ortmann
Sergey Slobodyan

CERGE-EI
Prague, June 2008

**ISBN 978-80-7343-155-6 (Univerzita Karlova. Centrum pro ekonomický výzkum
a doktorské studium)**

ISBN 978-80-7344-144-9 (Národohospodářský ústav AV ČR, v.v.i.)

(The Evolution of) Post-Secondary Education: A Computational Model and Experiments

Andreas Ortmann and Sergey Slobodyan*

CERGE-EI[†]

Abstract

We propose a computational model to study (the evolution of) post-secondary education. “Consumers” who differ in quality shop around for desirable colleges or universities. “Firms” that differ in quality signal the availability of their services to desirable students. As long as they have capacity, colleges and universities make offers to students, who apply and qualify.

Our model generalizes an earlier literature (namely, Vriend 1995) in an important dimension: quality, the model confirms key predictions of an analytical model that we also supply, and the model allows us to systematically explore the emergence of macro regularities and the consequences of various strategies that sellers might try.

We supply three such exercises. In our baseline treatment we establish the dynamics and asymptotics of our generalized matching model. In the second treatment we study the consequences of opportunistic behavior of firms and thus demonstrate the usefulness of our computational laboratory for the analysis of this or similar questions (e.g., the problem of early admission).

In the third treatment we equip some firms with economies of scale. This variant of our matching model is motivated by the entry of for-profit providers into low-quality segments of post-secondary education in the USA and by empirical evidence that, while traditional nonprofit or state-supported providers of higher education do not have significant economies of scale, the new breed of for-profit providers seems to capture economies in core functions such as curricular design, advertising, informational infrastructure, and regulatory compliance. Our computational results suggest that this new breed of providers is likely to continue to move up the quality ladder, albeit not necessarily all the way up to the top.

JEL Classification: C63, D21, D83, I21, L15

Keywords: post-secondary education, for-profit higher education providers, computational simulations

*{andreas.ortmann, sergey.slobodyan}@cerge-ei.cz. We thank participants of the WHEIA workshop in Trieste, Italy, the SCE international meeting in Amsterdam, Netherlands, and seminar participants at CERGE-EI for valuable comments. Special thanks to Dmitry Kolyuzhnov and two referees for an earlier version of this paper. Special thanks also to Martin Dieringer who helped

Abstrakt

Navrhujeme výpočetní model pro studium (vývoje) vzdělávacího systému po středním stupni. “Spotřebitelé”, kteří se liší v kvalitě, nakupují od vysokých škol a univerzit. “Firmy”, které se také liší v kvalitě, nabízejí studentům své služby. Pokud mají volné kapacity, vysoké školy a univerzity dávají nabídky studentům, kteří se přihlásí a projdou přijímacím řízením.

Náš model zobecňuje dřívější literaturu (jmenovitě Vriend 1995) z důležitého hlediska: kvalita, potvrzující důležité předpovědi analytického modelu, který rovněž nabízíme, což nám umožňuje systematicky studovat výskyt makro regularit a důsledky různých strategií, které mohou jednotlivý prodejci zkoušet.

Nabízíme tři takovéto případy. V našem základním přístupu popisujeme dynamiku a asymptotiku našeho základního srovnávacího modelu. V druhém přístupu studujeme důsledky oportunistického chování firem a tudíž ukazujeme užitečnost naší výpočetní laboratoře pro analýzu, popřípadě podobných problémů (např. problém předčasného přijetí).

V třetím přístupu přidáme některým firmám ekonomiku škál. Tato varianta našeho srovnávacího modelu je motivována vstupem proziskových poskytovatelů do části vzdělávacího odvětví s nízkou kvalitou vzdělávání po středním stupni ve Spojených státech a je také motivována empirickými pozorováními, že zatímco tradiční neziskový popř. státní provozovatelé vzdělání nedisponují významnou ekonomikou škál, nově přicházející proziskový poskytovatelé jí zřejmě disponují v základních funkcích jako je obsazení pracovních pozic, reklama, mezinárodní infrastruktura a plnění regulací. Naše výpočetní výsledky naznačují, že tento nový výskyt poskytovatelů bude pravděpodobně pokračovat ve zvyšování úrovně kvality, ačkoli ne nezbytně celou cestu až na vrchol.

with early versions of the computer code. Earlier versions of this article listed as co-author Samuel S. Nordberg, however, place and time constraints made continued collaboration impossible; we thank him for significant input in the early stages of this project. The first author acknowledges financial support by the Alfred P. Sloan Foundation and the Center for Adaptive Behavior and Cognition of the Max-Planck-Institut fuer Bildungsforschung in Berlin, Germany, which he was visiting while working on early versions of the computer code.

†CERGE-EI is a joint workplace of the Center for Economic Research and Graduate Education, Charles University, and the Economics Institute of the Academy of Sciences of the Czech Republic.

Address: Politických vězňů 7, P.O.Box 882, 111 21, Praha 1, Czech Republic

1 Introduction

Post-secondary education in the USA, formerly known there as higher education, has undergone dramatic changes over the past decade. The new label reflects the increasing orientation of traditional higher education providers toward vocationalism (see e.g., Breneman, 1994; Sperling and Tucker, 2006; Washburn, 2005), and the emergence of a new breed of higher education providers: publicly traded, degree-granting providers of post-secondary education (e.g., Ortmann, 1997; Ortmann, 2001; Sperling, 2000; Ruch, 2001; Kirp, 2003; Newman, Couturier, and Scurry, 2004; Pusser, 2005; Breneman, Pusser, and Turner, 2006) that we shall call for-profits from here on. These for-profit “mutants” now represent about 10 percent of the post-secondary education institutions in the USA.¹

That for-profits have managed to invade the higher education sector is little short of sensational. Higher education in the USA was, and for the most part still is, an industry whose private not-for-profit and public segments were, and still are, subsidized through significant tax and regulatory breaks, see Facchina, Showell, and Stone (1993), as well as – especially at the higher end – significant donations and the endowments they generate. In addition, not-for-profit and public institutions of higher education in the USA do not have to pay investors a reasonable return. Thus for-profits were, and are, clearly handicapped. How then could they succeed? This is the key question we address below.

For-profits invaded higher education initially by providing services to market niches such as information technology training and continuing education/workplace training for adults (e.g., Sperling, 2000; Ruch, 2001; Sperling and Tucker, 2006; Ortmann, 2006). In terms of the classification proposed by Zemsky, Shaman, and

¹The major publicly traded, degree-granting providers of post-secondary education in the USA (by way of their stock market symbols, APOL, CECO, COCO, DV, EDMC, ESI, STRA, and WPO) will generate about \$10 billion in revenue in 2006 which represents about 5% of the higher education market as traditionally understood. The divergence between market share in terms of number of institutions and in terms of revenue reflects the particularities of the ways for-profits operate: Typically they own centralized administrative and curricular development facilities and no-frills “campuses”/learning centers that are often located in malls for easy access. For more details, see Ortmann (2001, 2006), Ruch (2001), and Sperling and Tucker (2006).

Iannozzi (1997), for-profits initially entered post-secondary education through segments in which one typically also finds community colleges. Since then for-profits have successfully moved up to segments in which one typically also finds state universities. It is thus an interesting question whether this invasion of ever higher segments of post-secondary education by for-profit “mutants” will come to a halt, or whether for-profits will ultimately invade the top echelon of higher education as we knew it. To put it starkly, could a liberal arts college — arguably the paragon of what higher education once stood for — be organized as a for-profit institution?²

Towards a better understanding of these issues, we propose a computational model, or laboratory, that, in principle, could be calibrated with data from post-secondary education in the USA (e.g., the data on which the VIRTUAL U simulation is based).³ However, while we believe that our model captures the key aspects of post-secondary education, we prefer to think about it as a culture-dish that allows us to explore how macro regularities might emerge through the repeated local interactions of boundedly rational, heterogeneous agents. We will demonstrate the usefulness of such a laboratory through three computational exercises. We also relate our model to the simple theoretical model shown in the appendix that gives us confidence that our baseline results are sound.

Following exhortations in the literature to concatenate new computational models with predecessors (e.g., Axelrod, 1997), we first “reverse-engineered” and then generalized (especially as regards the classifier system) Vriend’s (1995) influential model of decentralized markets consisting of locally interacting boundedly rational and heterogeneous agents. Indeed, we have been able to replicate reasonably well Vriend’s results (e.g., the service ratio approaching to 1, approximately one third of consumers patronizing previously attended firms, etc.)

Since his model presented a decentralized market, with buyers and sellers not

²We are agnostic on the issue of whether a liberal arts college *should* be organized as a for-profit institution. Interestingly, a for-profit college with classic curriculum is scheduled to open in 2007 (see Anonymous, 2006).

³VIRTUAL U is an ambitious attempt to build a Sim City-like simulation of higher education in the US. It draws on real-world data in parameterizing the underlying simulation machines. See <http://www.virtual-u.org/> for more details.

stratified by quality (contrary to buyers and sellers of post-secondary education in the USA), we generalized Vriend’s model in that key dimension. In our baseline treatment (“Q-model” or T1), we explored the dynamics and asymptotics of this generalized matching model. The results of that baseline treatment confirm the results of an analytic model that we supply in the appendix.

To further demonstrate the usefulness of our computational laboratory, in a second treatment (T2) we study the consequences of opportunistic behavior of colleges and universities (e.g., admittance of unqualified students for fiscal reasons). The results demonstrate that our computational model lends itself to the study of various other related issues, such as viable quality improvement strategies for colleges and universities or the impact of various forms of early admissions (see Avery, Fairbanks, and Zeckhauser, 2003). We could also study, and in fact have done so in a previous version (see Ortmann, Slobodyan, and Nordberg, 2003), the emergence of behaviorally different consumers (“hoppers” rather than traditional “patronizers”).

Finally, in a third treatment, we study the key question that motivated our study originally: How could severely handicapped for-profit firms succeed in an environment that, at first glance, seemed utterly hostile? To this purpose we equip one firm with a cost structure that features initially higher cost but also economies of scale once a certain number of customers has been attracted by the firm. This variant of our matching model (“QES-model” or T3) is motivated by the entry of for-profit providers into low-quality segments of post-secondary education in the USA and empirical evidence (detailed in section 2.3) that, while traditional not-for-profit or state-supported providers of higher education do not have significant economies of scale, the new breed of for-profit providers seems to capture economies in core functions such as curricular design, advertising, informational infrastructure, rent-seeking, and regulatory compliance (e.g., Ortmann, 2001; Ruch, 2001).⁴ Our computational results suggest that this new breed of providers is likely to continue

⁴While our study is motivated by recent developments in post-secondary education in the USA, similar developments today can even be observed in transition economies that historically were much less open to curricular and other educational innovations, see Kraft and Vodopivec (2003).

to move up the quality ladder. There is, however, no guarantee that they will march all the way to the top.

The article is structured as follows: Section 2 presents the matching model, a discussion of our computational agents, and the details of experimental design and implementation. Section 3 presents, among other things, findings on the equilibrium distributions of firms across the quality spectrum under various treatments. Section 4 provides a brief discussion of related literature and a rationalization of our approach. In Section 5 we proffer some concluding remarks. The appendix contains an analytically tractable simplified version of our baseline model which suggests our computational results are robust to details of computational implementation.

2 Structure of the matching model

2.1 Summary of the matching model

Buyers (prospective students and/or their parents) and sellers (colleges and universities) of post-secondary education try to match optimally in a decentralized market for a number of periods.⁵ In the first period, buyers are randomly and uniformly distributed along a quality spectrum that is normalized to the interval $[0, 100]$.⁶ Likewise, in the first period, sellers are randomly and uniformly distributed along a quality spectrum that is normalized to the interval $[0, 100]$.

Buyers and sellers are modeled as boundedly rational decision makers that select actions probabilistically; we will rationalize this assumption in Section 4 below. We explain how exactly buyers and sellers make decisions in the following section.

Table 1:

⁵For the remainder of the text we use as synonyms the words buyers, consumers, and students, on the one hand, and sellers, firms, and colleges and universities, on the other hand.

⁶We realize that it would be desirable to allow the distribution to change over time, reflecting for example the influx into post-secondary education, and especially into for-profits, of students who are the first in their families to go to college. However, given our current focus, no additional insight would be gained from such a modification.

FIRMS	CONSUMERS
– make production and signaling decisions	
– signal	
	– choose firms, apply to one
– accept or reject consumers	
	– if rejected, choose another firm
...	...
– calculate profits, adjust quality	

Table 1 details, top-down, the timeline of interactions of sellers and buyers, i.e., the matching protocol, in each period. In every period, firms first make production and signalling decisions. Firms then signal potential buyers by picking a random buyer (with replacement) and checking her or his quality. Only those buyers that are within a pre-specified range of quality ($[Q - \Delta^C, Q + \Delta^F]$) get signalled, up to the pre-determined number of signals that the firm has chosen to send in that period. This reflects the practice of colleges and universities to admit only those students that fulfill certain minimum quality standards and to diligently track the yield of various advertising and recruiting channels (*i.e.*, not to waste recruiting efforts on candidates that can be expected to be out of reach or undesirable). Implicitly, we assume that firms know buyers' preferences regarding the quality of school they are willing to accept and thus avoid sending signals to consumers that are out of reach. This explains the upper limit $Q + \Delta^F$ of the quality range to which firms send signals.

Consumers then choose their firm from the offers. Only those firms become candidates that are above a pre-specified quality that equals buyers' own quality minus Δ^F . This reflects the practice of the overwhelming number of students not to go to colleges and universities with quality significantly less than the desired quality level. Symmetrically endowing consumers with knowledge of firms' behavior rules (making only firms with $Q \in [Q - \Delta^F, Q + \Delta^C]$ candidates) makes no difference, as no consumer of quality Q receives a signal from a firm with quality greater than

$Q + \Delta^c$.

Since, typically, a student will be signalled by several colleges or universities, the question arises how he or she prioritizes among multiple offers. We assume that consumers collect all offers and put those firms that satisfy a minimum quality on a list of desired firms. For consumers who PATRonize, this list consists of only one firm whose quality they do not check because they must have done so at some point in the past⁷ and because quality changes typically do not happen suddenly. All consumers (including those who PATRonize) then “apply” to their desired firm(s). This matching process involves two random processes as follows: First, a consumer is randomly drawn; second, that consumer randomly draws a firm from her list of desired firms. As soon as such a firm can and wants to provide, a match is accomplished and it is another randomly drawn consumer’s turn.⁸ Firms do not discriminate between consumers who patronize and those responding to offers.⁹

Buyers and sellers are characterized by preferences and internal states, behavioral rules (= rules for selecting actions out of the current choice set), the number of behavioral rules, internal rules (= rules for selecting and modifying rules), and specifications of the decision makers’ interactions with the world. Table 1 summarizes these characteristics which are discussed in more detail below.

Table 2

⁷Consumers are forced to take a KNOWN action in the first period; hence they are able to PATRonize from the second period on.

⁸Think of a student who collects all the information she gets in a large folder and, when the times come to apply, takes (randomly) the first one that fulfills her aspiration level. If this attempt fails, the student randomly selects another firm of sufficiently high quality out of the folder. We’ll rationalize that procedure in section 4. Other procedures are, of course, possible. For example, rather than selecting firms randomly, consumers might call on schools according to their quality.

⁹This is, in a sense, in contrast to Kirman and Vriend (2001) where loyal customers could receive more or less preferential treatment. The implications of loyalty on the part of sellers remains an issue for future research. Again, we believe that this issue is not of material relevance for the issues we are interested in here.

CONSUMERS	FIRMS
PREFERENCES	
Δ^F	Δ^C
INTERNAL STATES	
Weights; own Q, firm attended last period, list of schools that are desirable	Weights; own Q, demand, avg. Q of consumers, profit, average profit
BEHAVIORAL RULES	
IF (SAT, no SAT, indifferent) AND (INFO, no INFO, indifferent) THEN (PATR, KNOWN)	(production, signal)
NUMBER OF RULES	
18	20
INTERNAL RULES	
<i>Rules for Selecting Rules</i> stoch. auction; reinforcement	<i>Rules for Selecting Rules</i> stoch. auction; reinforcement <i>Rules for Changing Rules</i> production and signaling adjustment, GA
MATCHING PROTOCOL	
Specification of firm selection	Specification of consumer selection

The preferences of buyers and sellers are defined by the minimal quality of a counterpart they are willing to consider: Buyers will go only to firms that meet a given quality threshold (defined as own quality Q minus Δ^F); sellers are interested only in those consumers who meet a given quality threshold (defined as own quality Q minus Δ^C).

The internal states of buyers and sellers are defined as follows: Buyers keep track of the strengths (“weights”) of their behavioral rules, their own quality Q ,¹⁰ the index of the firm which they attended last period, and a list of schools that

¹⁰Currently, our consumers do not change their quality, i.e., exactly what school they attend has no consequence for their educational outcomes. Firms thus face a quasi-fixed distribution of consumers in quality space. We call this distribution quasi-fixed because the distribution of patronizers across firms changes over time. We note that there is no consensus about the value that colleges and universities add to human capital formation, see Altonji and Dunn (1996), Behrman, Rosenzweig, and Taubman (1996), and Tamura (2001).

are desirable (i.e., have a minimum quality $Q - \Delta^F$); sellers analogously keep track of the strengths (“weights”) of their behavioral rules, their own quality Q ,¹¹ the realized demand for their services, and current as well as average past profit.

Buyers’ behavioral rules each have a conditional and an action part (a Classifier System). The conditional part determines if a rule will be activated (to be explained below) given the current state of the world while the action part encodes possible actions. Specifically, rules have the following form:

IF (SAT, no SAT, Indifferent) **AND** (INFO, no INFO, Indifferent) **THEN** (PATR, KNOWN)

Here SAT denotes a buyer’s satisfaction (being served last period), and INFO records whether a buyer has received signals from firms in the current period. Think of this signal as an invitation to apply. Some rules use a coarser representation of the state of the world and are indifferent to SAT, INFO, or both. Buyers have two actions available to them: they can try to patronize the firm they attended (PATR) or try to go to a firm that signalled them (KNOWN).¹² Buyers who are not able to take the PATR or KNOWN action (because firms do not accept them) do not get matched. Buyers start with a complete set of $3 \cdot 3 \cdot 2 = 18$ rules which remain unchanged over the course of our computational experiments.

Unlike the classifier system representing the behavior of buyers, the behavioral rules for sellers encode pairs of integer numbers, one representing the number of units

¹¹A firm’s quality is updated according to the following rule, $Q = w_1 \cdot Q_{avg} + w_2 \cdot \pi$, where Q is the firm’s quality, Q_{avg} the average quality of its consumers, π the firm’s profits, and w_1 and w_2 are weights. Since the weight on profits is rather low, we essentially model the quality of a college or university as the average of the quality of its students. This follows well-established precedent in the literature, e.g., Lazear (2001), Rothschild and White (1995). We conjecture that factoring in the quality of faculty would not affect our qualitative results for all reasonable parameterizations.

¹²Adelman (2000) is an eminently readable sketch of the emerging “parallel universe of postsecondary credentials ... an education and training enterprise that is transnational and competency-based, confers certifications not degrees, and exists beyond governments’ notice or control.” (p. 20) We note that competency-based certification is also propagated by institutions such as Western Governors University (<http://www.wgu.edu/>) which has made considerable headlines by offering its prospective students the advantage that skills and knowledge acquired at other universities, on the job, or just through life may be counted toward one’s WGU degree. What all these developments point to is a new kind of student — “hoppers” we call them — who takes classes here and there and then consolidates her or his portfolio at a school of her or his choice. In our model, hoppers are modeled as consumers who never use the PATR action. Their behavior is analyzed in Ortmann, Slobodyan, and Nordberg (2003).

produced and the other the number of signals to be sent. Every integer is coded by a bitstring of length 10; therefore, numbers from 0 to $2^{10} - 1 = 1023$ can be coded. We restrict production to $[0,255]$ because the average number of consumers per firm is 100 and we have not observed large deviations from that number except in the initial adjustment process.

In each period, sellers produce slots whose availability is then signalled to prospective (and desirable) buyers. Note that rules in this sense translate directly into actions. There are twenty such rules that are initialized randomly (for every bit in a string, a fair coin is tossed to determine whether it is 0 or 1) so as to represent various production–signaling combinations.¹³

2.2 Experimental implementation: Stochastic auction, reinforcement, and evolution of rules

Buyers and sellers select their actions probabilistically in a stochastic auction into which all rules (for buyers, only the rules matching their current state) are entered. However, these rules are not entered equally weighted. Rather, their weights reflect their past “success”: the better they performed in the past, the more weight they get.¹⁴

With a small probability that we call “discard probability”, every rule’s total bid can be discarded. This procedure makes sure that the “best” rule typically wins the auction but that inferior rules have a small chance of winning as well. (In essence this is the idea of a trembling hand well established in the literature.) In the following the winning rule is called the active rule. The rule that was active in

¹³The two left-most bits of the production part of every rule are set to zero, so that a firm’s production is indeed between 0 and $2^8 - 1$.

¹⁴In such an auction, every rule submits a “bid” proportional to its weight or “strength” w plus *i.i.d.* error, $b_1w + \varepsilon$. The basic bid b_1w represents a stake that a rule is willing to pay for the right to win in the auction. This stake is higher the higher a rule’s weight is. Following Holland (1992), the winning rule pays its basic bid, of which b_2b_1w is transferred to the previous winning rule. Lettau and Uhlig (1999) show that there is an important connection between this so-called bucket-brigading and dynamic programming: In particular, in the steady state the current rule’s weight equals current payoff plus a share of the next period rule’s weight. Thinking of weight as value, and of rule as state, this is an analogue of a convergent Bellman equation.

the previous period obtains a share of the active rule’s basic bid.

The selection of the initial strength of a rule, its possible range (from zero to one in our case); the standard deviation of the auction’s error term, ε ; the speed with which that standard deviation decreases over time;¹⁵ and the discard probability all influence two characteristics of the stochastic process generated by the stochastic auction: the expected number of active rules (not more than three or four in our case) and the variance of the number of rules that will be called to duty on a regular basis.

Strengths of rules are restricted to $[0,1]$. This, together with the discard probability and the decreasing standard deviation of the auction’s error term, ε , is done to prevent, early in the simulation, the emergence of “runaway” rules that might lead to premature convergence.

After the stochastic auctions have determined the buyers’ and sellers’ rules for the current period, matching is implemented as described in the previous subsection and payoffs to buyers and sellers are realized. For a buyer, the payoff equals one if she is served this period and zero otherwise. For a seller, the payoff equals the ratio of the current profits to average profits over the last 200 periods,¹⁶ times $\delta \in [0, 1]$.¹⁷ Next, each buyer’s and seller’s payoff is multiplied by $b_1(1 - b_2)$ and this product is added to the active rule’s weight.¹⁸

The stochastic auction and reinforcement mechanism described above closely resembles various forms of probabilistic enforcement learning proposed in the literature. However, our evolutionary programming technique is more than simple individ-

¹⁵This is done by Vriend (1995) to reduce stochastic disturbances of the system as time progresses.

¹⁶This is motivated, first, by the parameterization in Vriend (1995) and, second, by our desire to stabilize our computational model within a reasonable run length.

¹⁷Multiplication by δ means that rules which produce average profit every period cannot achieve the maximum strength; instead, their weight converges to $100\delta\%$ of it. This construction facilitates the never ending emergence of strategies that aim to beat the average performance: no seller rule will be used forever and eventually new combinations of (production, signalling) pairs will be experimented with. This “new broom effect” speeds up adjustment to a rapidly changing environment.

¹⁸Together with bucket-brigading, this ensures that weight of the rule used every period converges to one for buyers and δ for sellers.

ual reinforcement learning. To model the behavior of firms, we used a combination of the Steepest–Ascent Hill–Climbing algorithm and the GENITOR algorithm.¹⁹ This evolutionary technique is arguably the simplest programming technique and as such is a desirable baseline, e.g., see Chen, Duffy, and Yen (2002) and the critique of Valente (2002).

The Steepest–Ascent Hill–Climbing part of our algorithm is implemented as follows: firms update the production part of their rules in every period as in Vriend (1995). If a firm’s demand (number of consumers that applied to a firm in the current period) differs from its production, production is adjusted by 10% of the difference or 1 unit if 10% of the gap is less than 1.

The GENITOR part of the algorithm is implemented by generating one completely new rule every 50 periods. Every firm’s rules are ordered by their weight and two “parent” rules are selected from the top quarter (top five rules). A standard uniform crossover operator is applied to the binary strings—parent rules—and one of the two “children”, randomly selected, is retained. Then, we mutate every bit of the child string and replace a randomly selected rule from the bottom half (bottom ten rules) with the child, which is assigned a weight equal to the average of its parents’ weights. We could use even simpler procedures, like the child being a linear combination of the two parents’ rules, but it would be essentially inconsequential.

Since the consumers’ classifier system is complete, there is no need to evolve it further. Unlike firms, consumers have a set of rules that does not change over time. (Of course, the strengths of the rules might change.)

After buyers and sellers have been matched, firms compute their revenues, costs, and profits. They also update their quality as the weighted average of the quality of students who have chosen to enroll and current profits, with weight on profits

¹⁹In the GENITOR algorithm, rules are ranked according to their fitness, and the probability of selecting a particular rule is proportional to its rank. Every n^{th} period, two evolutionary operators (crossover *and/or* mutation) are applied to produce a new rule, which is inserted into the existing ranking and replaces an old rule. One of the advantages of the GENITOR algorithm, according to Chattoe (1998) and Whitley (1989), is the relative stability of the ranking, which results in stable actions. Chattoe (1998) argues furthermore that the GENITOR algorithm closely resembles the real–world decision–making process in companies and humans.

being relatively small.²⁰ While this approach to determining the quality of colleges and universities — essentially defining the quality of a school as the average of the quality of the students that it attracts — is admittedly simplistic, it captures, in our view, the most important aspect of what determines the quality of an institution. Specifically, it allows us to study the trade-off any typical college faces on the margin of admitting a rich but not-too-smart, instead of a poor, but brilliant, student. Below, we call such admittance of unqualified students for fiscal reasons opportunistic behavior.

This process repeats round after round. The matching process, in other words, is a dynamic process that evolves over a number of periods. The dynamic process is defined algorithmically in terms of the behavioral rules of our agents, their internal states and preferences, their repeated interactions, and — through internal behavioral rules that govern how rules are selected and changed — the evolution of rules toward some stable outcome.

The program code is described in the working paper version of this manuscript, Ortmann, Slobodyan, and Nordberg (2003).

2.3 Experimental design: Parameters and treatments

Following exhortations in the literature to concatenate new computational models with predecessors, e.g., Axelrod (1997), we parameterize our baseline Q-model almost completely with the parameters Vriend (1995) chose, with two notable modifications (signaling costs are much lower in our model, and arguably were unrealistically high in Vriend’s model; the variance of the error term in the consumer stochastic auction is higher, allowing our consumers more experimentation which arguably reflects search behavior better in the market that we study here; see Boylan 1998). Table 3 below details all relevant parameters common to our three treatments (the Q-model, the Q-model with moral hazard, and the QES-model) and also relates these parameters

²⁰The formula for updating the quality is $Q = w_1 \cdot Q_{avg} + w_2 \cdot \pi$. $w_1 = 0.95$ in all simulations, and w_2 is calibrated by requiring the average firm quality to be equal to 50 which produces w_2 about 0.1. Other ways of calibration are, of course, possible but seem less natural.

to those employed by Vriend (1995).

Table 3

Run-length	3000	Vriend (1995)
Production cost C_Y (maximal quality)	.25	— . — . —
Signal cost C_S (maximal quality)	.025	Vriend = .08
Price P (maximal quality)	1	Vriend (1995)
Average number of consumers per firm	100	— . — . —
Maximum acceptable quality gap, consumers	10	NA
Initial rule weight, firms (consumers)	0.3 (0.5)	Vriend (1995)
Steady state weight δ of an average rule, firms	0.65	NA
Stdev, auction error term, firms, $N(0, R)$	0.075↓0.03	Vriend (1995)
Stdev, auction error term, consumers, $N(0, R)$	0.034	0.012
Parameter b_1 , firms (consumers)	0.25 (0.1)	— . — . —
Parameter b_2 , firms (consumers)	0.4 (0.1)	— . — . —
Mutation probability	0.01	
Uniform crossover probability	0.50	
Discard probability	0.025	Vriend (1995)

For all three treatments we used the same parameterization. In the first treatment (T1) we implemented the Q-model to generate baseline equilibrium distributions of firms across the quality spectrum. In the second treatment (T2) we continued to use the Q-model but inserted an opportunistic firm in the set-up. Such a “mutant” accepts consumers whose minimum quality is 12 rather than 10 points below its own quality. In the third treatment (T3) we continued to use the Q-model but inserted a “mutant” (representing a for-profit firm) that initially had higher costs but also economies of scale once a certain number of customers had been attracted; for ease of reference we refer to that model as “QES-model”.

The exact profit function for a “normal” firm (i.e., the firms populating T1, and all firms that are not “mutants” in T2 and T3) is

$$\pi = (P \cdot \min(Y, D) - C_Y \cdot Y - C_S \cdot S) \cdot \frac{Q}{100},$$

where Y is the firm’s production, D the realized demand, S the number of signals, and Q the firm’s quality. For a for-profit mutant, the cost term $C_Y \cdot Y + C_S \cdot S$ is

multiplied by 1.2. If Y is above $MES = 50$ for all runs reported in this paper, its cost is further multiplied by $\frac{MES}{Y}$.

Our parameterization is meant to capture the following facts about the cost configurations of non-profits and the kind of for-profits (colleges and universities that are part of publicly traded education companies): Traditional firms — which tend to be stand-alone, brick-and-mortar entities — do not have economies of scale in a range that is relevant (e.g., Goetz, Siegfried, and Zhang, 1991; Laband and Lentz, 2004). In fact, diseconomies of scale seem to set in fairly quickly. In contrast, publicly traded for-profits, while having significant start-up costs, are generally believed to have significant economies of scale resulting from centralization of such functions as curricular design, advertising, informational info structure, rent-seeking, and regulatory compliance (e.g., Ortman 2001; see also Ruch 2001, especially chapter 4; see also Spelling and Tucker 2006). Our claim is not easy to verify — the kind of data that one would need for it are proprietary — but a look at the 10-K form supplied by the Apollo Group on its website (www.apollogrp.edu/) is very suggestive: Average costs per student as reported there were \$5,650 in fiscal year 1995, rose nominally to about \$6,000 in 2000, and drifted down since then to \$5,000 in fiscal year 2005. In real terms, after an initial increase, average costs declined by about 30% over a decade, with student numbers having grown about ten-fold during that time. A significant part of that growth (and the economies of scale being captured) seems to be attributable to the University of Phoenix online division for which the Apollo Group issued a tracking stock for a couple of years. Unfortunately, once profits exploded — according to the UOPX 10-K the online division turned a profit of \$15 million in fiscal year 2003 and \$140 million in fiscal year 2004 — the Apollo Group killed the tracking stock (and hence an important source of data).

Following Glaeser and Shleifer (2001), we do not handicap our for-profits with taxes although — from pilot runs that we do not report here — doing so would not change our qualitative results; it would only slow down their movement up the quality ladder. We also do not explicitly model the state appropriations that (some)

nonprofits schools get, which have declined and continue to decline at a rapid pace (e.g., Breneman, 2005) and, in any case, do not affect the cost side. Overall, we feel confident that our cost assumptions are reasonable (see also Ruch 2001, p. 87).

Since scaling effects are notorious, we controlled for them by implementing treatments T1 through T3 with combinations of 12 firms/1200 consumers (Scale1) and 24 firms/2400 consumers (Scale2). Table 4 below summarizes our 3x2 design, detailing the number of runs in each cell and the number of mutants for treatments T2 and T3 across all scales. The case of 10 firms/1000 consumers was presented in Ortmann, Slobodyan, and Nordberg (2003).

Table 4

treatments\scales	Scale1: 12 firms/1200 cons	Scale2: 24 firms/2400 cons
T1: Q—model	100 runs	100 runs
T2: T1 + MH	100 runs (1 mutant)	100 runs (1 mutant)
T3: QES—model	100 runs (1 mutant)	100 runs (1 mutant)

3 Results

We restrict ourselves to what we consider the essential characteristics of the runs in a treatment cell.²¹ Before we analyze the equilibrium distribution of firms across the quality spectrum in Section 3.3, we describe convergence toward equilibrium distributions and signaling, production, and demand trends for the baseline treatment (without opportunistic firms and without for-profit firms). We focus on the baseline treatment, T1. We do not discuss the emergence of behaviorally different consumers (“hoppers” rather than “patronizers”); a discussion for a series of related runs may be found in Ortmann, Slobodyan, and Nordberg (2003). The results for the runs discussed here are qualitatively the same.

²¹A set of figures presenting all 600 runs may be obtained from the authors upon request.

3.1 Convergence toward equilibrium distributions

The results from the Q-model, in T1, show significant path dependence for both scales: the “slots” in the quality spectrum, or market niches characterized by quality ranges (which we shall also call, inspired by Zemsky, Shaman, and Iannozzi (1997), “segments”) in which a firm will find itself, typically depend on its initial rank. If a firm, for example, is initially one of the top four firms in Scale2, it is likely to end up in the top segment even if its initial quality lies significantly below the predicted quality of the cluster. The adjustment process takes between 400 (Scale2) and 200 (Scale1) iterations.

While convergence to the equilibrium location is relatively fast, we do observe — even in the absence of opportunistic firms or entrants with economies of scale — occasional eruptions and displacements in quality. A firm that moves up or down the quality spectrum typically dislodges another firm from the segment it invades. While the number and location of segments is relatively stable, there is some jockeying going on for those segments.

Opportunistic firms (T2) or entrants with economies of scale (T3) complicate the picture, generating more eruptions and displacements and slower convergence toward the equilibrium distribution. In fact, we often see cascade-like sequential convergence toward the equilibrium distribution, see Figures 1 and 2.²²

3.2 Signaling, production, and demand trends

Even though we initialize our computational experiments with quantity-signalling pairs that may be widely off-equilibrium (recall that the initial number of slots could be as high as 255, and the number of signals is initially bounded by 1023), production and demand tend to converge to their equilibrium values — which are about 100 for slots independent of scale, and 450-950 for signals dependent on scale — within

²²We have computed a measure of deviations of firms from the theoretical equilibrium — essentially the sum of squares of deviations — for all runs mentioned in Table 4. These computations give a measure of convergence beyond our informal discussion above; they are available from the authors upon request.

the first 500 periods, both in the aggregate and for individual firms (see a typical aggregate picture in Figure 3). Signaling, however, converges much more slowly and is much more volatile: a typical stochastic fluctuation in the firms' demand (10 to 15%) can lead to a much larger change in the perceived optimal signal level. This is a consequence of our signaling costs, relative to those in Vriend (1995), being fairly low. An additional source of uncertainty arises because of the local nature of the information that firms collect: they do not observe demand at other firms.

3.3 Equilibrium distributions of firms across the quality spectrum

We analyze the distribution of firms after 3000 iterations but, as the selected figures illustrate, an analysis after 1500 iterations leads qualitatively to very similar results.²³ As we will also see, segments are typically occupied by clusters of firms. We shall use the terms “segments” and “clusters” interchangeably.

Baseline treatment T1. In Appendix A we show that, theoretically, we should have 6 clusters for all firm numbers modulo 6.²⁴ Our choice of Scale1 and Scale2 was motivated by these theoretical results, although we have explored other scales, with qualitatively similar results, see Ortmann, Slobodyan, and Nordberg (2003).

Looking at 12 firms and 1200 consumers (Scale1), and 24 firms and 2400 consumers (Scale2), respectively, we do observe indeed 6 clusters of 2 and 4, as theoretically predicted. The number of firms in each cluster is essentially constant, with occasional eruptions and displacements reflecting the probabilistic nature of our modeling technique, see Figure 4. Interestingly, but in light of the results from

²³As mentioned, convergence to relatively stable configurations (6 clusters of firms) occurs in T1 within the first couple of hundred iterations (which could be thought of as semesters, trimesters, quarters or some such time unit). Even in T2 and T3, the distribution of firms closely resembles the theoretical configuration within 1500 iterations. Recall that we initialize production randomly on $[0,255]$ and therefore typically off the equilibrium of 100 units per firm. Doing runs of 3000 allows us to estimate the likelihood of disturbances and switching behavior. More details below.

²⁴We note that this number is a function of the width of the quality range and the width of the segment (to be made precise later). *Ceteris paribus*, increasing the quality range leads monotonically to a higher number of clusters. We speculate that this relation can be described by the formula $\left\lceil \frac{\text{range}(Q)}{\Delta^F + \Delta^C} \right\rceil + 1$, where $[A]$ denotes an integer part of A and $\text{range}(Q)$ is the quality range.

our model in Appendix A not surprisingly, such displacements regularly result in an exchange of members of adjacent clusters.²⁵ We note, finally, that clusters are distributed approximately equidistantly, again as predicted by the calculations in Appendix A. This result is also independent of the scale.

From the above it follows that scale is important in two respects. First, only scales modulo 6 can be accurately described by our symmetric steady state calculations. In other words, there is a large degree of freedom for scales that are not of modulo 6, especially if the number of firms is rather small. As we increase the number of firms, it becomes less important whether the number of firms is modulo 6 or not. This is good news because it means that the computational model that we propose here is rather insensitive to integer constraints. Second, as we increase scales, we find — somewhat contradicting our initial intuition — a rather stable configuration of six clusters or segments which attract whatever number of firms populate our computational laboratory.

Moral hazard treatment T2. For both Scale1 and Scale2 the opportunistic firms almost never manage to markedly increase their position in the quality spectrum after the initial adjustment process.

Table 5. Transition frequencies, T2

	Scale 1							Scale 2						
	1	2	3	4	5	6	Tot	1	2	3	4	5	6	Tot
1	33	0	0	0	0	0	33	26	0	0	0	0	0	26
2	16	3	0	0	0	0	19	16	2	0	0	0	0	18
3	8	5	2	0	0	0	15	10	3	2	0	0	0	15
4	5	2	3	2	2	0	14	14	2	3	1	0	0	20
5	2	3	2	1	2	1	11	7	8	1	0	0	0	16
6	0	0	0	4	1	3	8	0	0	0	2	1	1	3

Table 5 represents the transition frequencies separately for 100 runs of Scale1 and Scale2. Each of these runs featured one opportunistic firm (“mutant”). The right-most column in the Scale1 and Scale2 panels represents the starting position

²⁵We note that we have similar results for exploratory runs with scales of 20/2000 firms/consumers, as well as 40/4000, 48/4800, and 120/12000 all of which are not reported here.

of the mutants, with “1” denoting the bottom segment and “6” denoting the top segment.²⁶ For example, 15 firms started in both Scale1 and Scale 2 in segment “3”. In Scale1, 13 of these mutants could not hold on to their positions and moved down to segment “2” and “1”. Similarly, for Scale 2, 13 of this mutants moved to segment “2” and “1”. In fact, that is the clear-cut message: not counting the 33 and 26 firms that started in the lowest segment in Scale1 and Scale2, respectively, and got stuck there, only three moved slightly up, 12 stayed put, and 52 moved down in Scale1, whereas none moved up, 6 stayed put, and 68 moved down in Scale2. Thus, about 80% to 90% moved down, some of those that moved down did so considerably, in Scale2 more so and faster than in Scale1 (compare Figures 5 and 6).

The results reported here emerge from the very mild parameterization of moral hazard that we chose; increasing the moral hazard parameter increases the probability of downward drift systematically. If, for example, the moral hazard parameter is doubled (i.e., decreasing the quality of the worst student from $Q - 12$ to $Q - 14$), the offending firm nearly always (more than 90%) goes to the bottom of the quality spectrum.

Therefore, at least in our model, opportunism does not pay in the long run: it increases customer numbers at best marginally and leads typically to a sharp drop in quality and hence profits.

For-profit invasion treatment T3. In contrast, for both Scale1 and Scale2 and after the initial adjustment process, a for-profit firm never lowers its position in the quality spectrum.

²⁶The initial distribution was indeed symmetric and uniform over the whole quality spectrum range. In order to smooth the trajectories, we computed quality as a moving average over 50 rounds. Since opportunists tend to lose their quality rather fast, the right-most column reports distributions for Scale1 and Scale2 that are skewed to the bottom.

Table 6. Transition frequencies, T3

	Scale 1							Scale 2						
	1	2	3	4	5	6	Tot	1	2	3	4	5	6	Tot
1	26	9	3	2	0	1	41	34	0	0	0	0	12	46
2	0	8	8	2	2	21	41	0	0	0	0	0	41	41
3	0	0	0	0	2	11	13	0	0	0	0	0	11	11
4	0	0	0	0	0	5	5	0	0	0	0	0	2	2

Table 6 represents the transition frequencies separately for 100 runs of Scale1 and Scale2. Remember that each of these runs featured one for-profit firm (“mutant”). The right-most column in the Scale1 and Scale2 panels represents the starting position of the mutants, with “1” denoting the bottom segment and “6” denoting the top segment.²⁷ In fact, for-profits increase their quality most of the time, often dramatically so, as in Figures 7 and 8. Table 6 allows us to make more quantitative statements. For example, 41 firms started in both Scale1 and Scale 2 in segment “2”. In Scale1, 8 of these mutants stayed while everyone else moved up, with slightly more than 50 percent of the for-profits that started in segment “2” moving up all the way to segment “6”. Even more dramatically, for Scale 2, all for-profits that started in segment “2” moved up all the way to segment “6”; similarly for those for-profits that started in segment “2” and “4”. The interesting exception is the high number of for-profits that start out in segment “1” and do not manage to move out of that attractor. For Scale2, we even see a complete bifurcation, with about 3 mutants being stuck in segment “1” for every mutant that manages to escape and that indeed manages to escape all the way to the top. Also note that successful mutants do not necessarily start climbing the quality ladder immediately after iteration 500 when the increasing returns to scale regime is switched on in simulations.

In sum, the probability of moving up increases with the cluster number in which the for-profit finds itself immediately before the increasing returns regime (iteration 501) starts. Moreover, there is a threshold quality level somewhere between clusters 1 and 2, such that random deviation in quality of the for-profit is dampened (for-profit

²⁷The initial distribution was uniform over the lower half of the quality spectrum range, since we knew from the results reported in Ortmann, Slobodyan, and Nordberg (2003) and various pilot runs with the new cost configuration that we would see wide-spread upward drifts.

returns to its equilibrium position in cluster 1). For those above, such deviations are accelerated and for–profits stops at the very top. Some cluster 1 for–profits do move up because a particularly large positive quality deviation pushes them beyond the threshold. The existence of such a threshold could be confirmed using a simplified symmetric model of the appendix, however, the algebra becomes extremely tedious.

There are various metrics that can further quantify the trends above. Table 7 summarizes one such metric.

Table 7

		$\overline{Q}_{end} - \overline{Q}_{start}$		
		min	avg±std	max
	T1, average for all firms	-3.6	3.5±5.3	12.6
Scale1	T2, opportunistic firm	-59.5	-6.4±13.6	5.9
	T3, for–profit mutant	-10.4	30.1±26.7	75.9
	T1, average for all firms	-9.0	1.8±5.0	14.2
Scale2	T2, opportunistic firm	-58.0	-7.9±11.5	0.6
	T3, for–profit mutant	-4.1	44.9±33.6	86.7

\overline{Q}_{start} and \overline{Q}_{end} denote the average quality of a firm during periods 100 — 500 and the last 500 periods, respectively;²⁸ therefore $\overline{Q}_{end} - \overline{Q}_{start}$ is a measure of the change of a firm’s position in the quality spectrum over time. This measure quantifies in particular the default outcome of opportunistic firms moving down and for–profits moving up in quality for Scale1 and Scale2. Compare, for example, row 1 of the Scale1 section with rows 2 and 3 respectively. The average quality change for all firms in T1 (3.5) is larger than that of opportunistic firms in T2 (-6.4) and smaller than that of for–profits in T3 (30.1). Along similar lines, note that the quality change range has increased dramatically in for–profits, going from 12.6 to 75.9 at the upper limit. Similar effects can be observed for Scale2.

²⁸We do not incorporate the first 100 periods because several hundred periods are needed for the initial noise to get worked out of the system. Including the first 100 periods makes the data noisier but does not change any of the qualitative results. Excluding more initial periods would not leave enough periods for averaging before the increasing returns to scale regime of for–profits takes effect in period 501.

Turning to the difference between Scale1 and Scale2, the key result is that opportunistic firms tend to fare slightly worse in Scale2 (upper limit of the range being 0.6) than in Scale1 (5.9). Average downward movement of opportunistic firms for Scale2 is larger but not statistically different from that in Scale 1 (-7.9 vs. -6.4). Somewhat analogously, we see a stronger average upward movement of for-profits for Scale2 (44.9 vs. 30.1 for Scale1). Also, mutants' movements in Scale2 tend to be more abrupt than in Scale1: compare Figures 5 and 6 with 7 and 8, which are rather typical.

The correlation between firms' mobility and scale has a straightforward rationale: when a firm, for some reason, manages to get more than equilibrium share of its segment, its increment in quality will be proportional to profits, which are in turn proportional to the total number of consumers. Grabbing an additional 5% of a segment with 300 consumers adds 0.18 quality units to an average firm; an additional 5% of a segment with 100 consumers adds just 0.06 quality units to an average firm. Analogously, competitive advantage (of for-profits) or disadvantage (of opportunistic firms) translates readily into more pronounced quality changes and hence into more turbulent environments as the number of firms and consumers per segment increases. Less stability creates, of course, more opportunities, both positive and negative, for mutants.

4 Related literature and rationalization of our approach

The computational matching model presented above has three reference points in the literature.

First, there is the classic work by Gale and Shapley (1962) on college admissions and later related work on two-sided matching, e.g., Roth and Sotomayor (1990); Roth and Xing (1994); Roth and Xing (1997); Roth (2002); Pingle and Tesfatsion (2001); Vriend (1995); Kirman and Vriend (2001); and Weisbuch, Kirman, and Herreiner (2000). This literature has theoretically illustrated the heavy math-

emational machinery necessary to model matching processes; it has also provided compelling evidence, both theoretically and empirically, on the importance of institutional arrangements that prevent, for instance, lower-ranked market participants from “jumping the gun” on other (higher-ranked) market participants. In the context of post-secondary education in the USA, this issue has been of long-standing interest as evidenced by Avery, Fairbanks, and Zeckhauser (2003) or Kirp (2003).

Second, there is work that attempts to document the changes in higher education over the past couple of decades. What little is out there in academic journals has already been mentioned in the introduction; for the time being much of the relevant information on those developments remains available only in official SEC forms or in research reports of investment houses. The situation is changing slowly, e.g. Ruch (2001); Kirp (2003); Newman, Couturier, and Scurry (2004); Pusser (2005); Breneman, Pusser, and Turner (2006); and Laband and Lentz (2004). None of these references contains, however, a satisfying analysis of (the evolution of) post-secondary education. (The description in Laband and Lentz, for example, is based on 1996 data). Part of the problem is the proprietary nature of the data that would be necessary for a hard-headed analysis of the recent developments (e.g., data about the economies of scale resulting from centralization of curricular design, advertising, informational infrastructure, and regulatory compliance).

While not directly addressing recent changes in higher education, two academic papers deserve mention here. Rothschild and White (1995) study peer effects and show theoretically why it is imperative for colleges and universities to give out need-based financial aid to deserving students. Only by doing so will colleges and universities be able to attract those bright (but poor) students that are an indispensable input in the production process of those students that can pay but are not so smart. The analysis of these authors also makes clear why there is nothing sinister about giving merit-based (rather than need-based) financial aid to students that are deserving (whether they are poor or rich). Not surprisingly, colleges and universities do routinely monitor the comparative attractiveness of their own financial aid packages and those of their close competitors. The resulting competition

adds to the pressure to cheat on the margin (i.e., admitting “legacy” or rich students for their prospective financial rather than academic contributions.) It is this development, and the insightful analysis in Rothschild and White (1995), that motivated our moral hazard treatment. Epple and Romano (1998) theoretically and computationally study the competition between tax-financed, tuition-free public schools and competitive, tuition-financed private schools in primary and secondary education; the impact of vouchers; and peer-group effects when students differ by ability and income. The equilibrium of their model shows that schools stratify along the quality spectrum and that students in private schools, dependent on their marginal productivity, either receive tuition discounts (or have to pay tuition premia), as (implicitly) suggested in Rothschild and White (1995).

Third, there is a literature on modeling social processes through GAs and related evolutionary programming techniques. Arthur (1994) and Arthur (1991) persuasively argue the case for agent-based models of interactions of boundedly rational and heterogeneous agents. Arthur (1994) points out that such models are grounded in plenty of evidence. Indeed, much of the evidence in experimental economics (e.g., Camerer, 2003) and experimental psychology (e.g., Cosmides and Tooby, 1996; Gigerenzer, Todd, and ABC Research Group, 1999; Cowan, 2001) has reinforced the impression that Arthur gets it right — people (whether real or fictitious, such as organizations) are “intuitive statisticians” (Cosmides and Tooby 1996) who inductively keep track of the performance of a set of plausible, simple models of the world that they can cope with. When the time to make choices comes, people act upon the most credible and possibly most profitable one. The others they keep in the backs of their minds, so to speak (Arthur 1994, p.407, slightly modified). Arthur (1991) makes a similar argument but also stresses the importance of calibrating computational agents to accurately reflect how human agents learn. Not much attention has been paid during the last decade to this exhortation, although recent developments comparing the performance of human and computational agents in more or less identical settings (e.g., Chen, Duffy, and Yen, 2002; Pingle and Tesfatsion, 2001; Roth, 2002) are encouraging.

In our model, we translate this literature by assuming that both firms and consumers are boundedly rational decisions makers. Specifically, we model buyers and sellers as boundedly rational decision makers that select their best actions probabilistically by selecting a rule from the sets of rules available to them. The rules, to recall, are weighted and their weights are updated continuously to reflect their past “success”: the better they performed in the past, the more weight they get. We believe that such probabilistic choice behavior, or reinforcement learning, is a valid description of actual agents’ behavior although the behavioral rules are not derived from a well-specified dynamic optimization problem. In fact, such a derivation is not possible (unless the decision problem is simplified so as to become almost unrecognizable).

Take consumers: Even US educators in economics are unable to partition US Ph.D. granting economics departments into more than 17 segments, see Thursby (2000). Never mind that the professors of economics who were the experts in that study are much more sophisticated decision-makers than the average consumer (a prospective student), consumers of postsecondary education evaluate colleges and universities along many more criteria than the quality of a Ph.D. (as discussed delightfully in Boylan, 1998). In fact, a veritable cottage industry has emerged that guides those applicants that can afford to shell out thousands of dollars in the admissions game (Kirp 2003).

Take firms: To adjust the production/signaling pair for the next period exactly, the firm needs to know the number of competitors, their exact quality, how their quality changes given this period’s profits, how their student mix is expected to change (and this, in turn, depends on how many students were turned away by the competitors and the exact weight structure of rules of these students), and what will be the student mix in the next period (this changes from period to period, as every firm changes its quality, which in turn makes some new students desirable or willing to consider this firm; properties of these students, i.e., weights of their rules and therefore the probability of them selecting PATR or KNOWN, is unknown to the firm). In other words, a gradual adjustment of production/signaling pairs,

combined with an occasional change of the strategy (selecting a different rule) seems reasonable.

The model’s structure is extremely complicated, and the inference problems of firms correspondingly daunting, especially when the model is still converging to the equilibrium with 6 clusters with stable membership. By selecting a production/signaling pair, every firm is influencing its quality and thus trying to find a steady state quality subject to all other firms doing the same. This optimization problem might be well-behaved locally (there is a single maximum), but not globally: it is always possible to switch places with a firm from the higher segment. Trying different starting points (different rules) for the Hill Climbing algorithm gives the firm a chance of getting into a neighborhood of a better local optimum, or a better initial position around the present local maximum, thus improving the convergence speed.

As to the interaction of firms and consumers, the matching problem is extremely complicated, especially when both firms and consumers are still far from learning a symmetric equilibrium with 6 clusters or something similar to it. It would be extremely naive to expect that any individual firm or consumer would be able to set up a dynamic optimization problem and solve it, especially far from the equilibrium. In general, learning optimal behavior in a dynamic problem is not an easy task: see, e.g., Lettau and Uhlig (1999) for a consonant view on using rules of thumb.

Even though our model lacks explicit intertemporal links such as savings or investment, it is dynamic: firms’ quality tomorrow depends on their behavior today, firms’ behavior influences what the consumers are doing tomorrow, and consumers’ actions (decision to go to this or that firm, to use a signal or disregard it and go for PATR action, etc.) also influence a firms’ future behavior through profitability of different production/signaling pairs. This dynamic nature is the reason we decided to include “bucket brigading” into our Classifier System, as described in Section 2.2. Lettau and Uhlig (1999) describe a connection between “bucket brigading” in Classifier System and dynamic programming approach to maximization.

We note that “bucket brigading” is not really necessary, as we observed conver-

gence to a neighborhood of the symmetric 6-cluster equilibrium without it; however, it does improve the convergence speed. We also note that despite the complexity of the problem the firms and consumers are facing, they are able to locate the equilibrium across a large variety of different variations in the model parameters, such as presence of “bucket brigading”, details of implementation of the stochastic auction, precise weight on the profits in the firm’s quality update equation, details of the matching protocol (in early versions of this paper, consumers with action KNOWN randomly selected a single firm among those that have sent a signal instead of forming a list of all such firms as in the current version), etc. This behavior of our model allows us to say that equilibrium with firms located in 6 equidistant clusters is a “deep minimum”, and so any sensible specification of the learning behavior by firms and households will probably converge to it.

A standard objection to agent-based modeling is, “why not model the matching process the good old-fashioned way, i.e. using an equilibrium search model with perfectly rational agents?” Firstly, drawing on Arthur’s arguments and the experimental literature already mentioned, we believe it is self-evident that agents (including aggregate agents such as firms) are not perfectly rational. (That, of course, does not mean that they do not optimize. Surely our agents optimize, noisily, within the constraints they have been given.) Secondly, we (see also Roth (2002) for a similar argument) simply do not see a way to model the issues we have addressed above in the good-old fashioned way. (That said, we stress that in the appendix we provide analytic results for a simplified version of the problem we analyzed above).

We do acknowledge that the sensitivity and arbitrariness of agent-based modeling is an important issue. An important feature of our technique is that we can easily incorporate competing behavioral assumptions, and that we can do so even by modelling various types. Think of this feature as a list of assumptions that is initially given to a reader who then might (dis)agree with their (un)reasonableness. To illustrate, recall how we conceptualized the consumer’s decision: Consumers randomly draw a firm from their lists of desired firms that have signaled them. One might argue in favor of a higher degree of rationality and have consumers select the

best firm instead. It is obvious that checking the sensitivity of the model to this change in assumption would require only a couple of key strokes and another set of runs.

Relatedly, there is the question of the specifics of the evolutionary modeling technique, such as the specific parameter values of the standard deviation of the auction's error terms, the speed with which that standard deviation decreases over time, and the discard probability. Specifically, a number of authors such as Michalewicz (1999); Mitchell (1996); and Chattoe (1998), have voiced a concern that the degrees of freedom inherent in evolutionary modeling techniques — similar to the degrees of freedom of the design and implementation of human experiments — subject any computational model to the real danger of being a mere example, and one, for that matter, that may be rather unrepresentative as regards the complete set of sensible parameterizations. This issue is admittedly an important and tricky one. For now, we have solved it by relying almost exclusively on the implementation details in Vriend (1995).

Axelrod (1997) has enumerated some of the problems that complicate replication of computational simulations (and re-engineering of extant models). Our own experience supports Axelrod's exhortation to make a model description and presentation of results as unambiguous and complete as possible, and to facilitate other researchers' attempts to re-engineer one's model; see also the related discussion in Valente and Andersen (2002) although we have our own reservations about the approach they propose. Replicability, on that we agree, is the hallmark of good science among experimental economists and psychologists alike (see Roth (2002) and the commentaries on that article) and it seems worthwhile to establish it as a fundamental methodological tenet in agent-based modeling, too;²⁹ for other tenets see Hollenbeck (2000).

²⁹In this spirit, we will make our code available to interested researchers.

5 Concluding remarks

We have proposed a computational model to study (the evolution of) post-secondary education. Although our model is motivated by developments in the USA, the insights it generates should be easily transferable to related developments in other countries. While, in principle, we could calibrate our model with data from the USA (or other countries, for that matter), and while we believe that it captures key aspects of post-secondary education – e.g., our results on the detrimental consequences of opportunism and the likely advances of the for-profit enterprises we were interested in — we prefer to think about our model primarily as a computational laboratory. It is useful to conceptualize such a laboratory as a culture-dish, as Tesfatsion (2002) does, which allows us to explore how macro regularities might emerge from the bottom up through the repeated local interactions of boundedly rational, heterogeneous agents and that in turn allows us to explore systematically the consequences of various strategies that sellers might try.

We do believe that our computational agents’ decision making is a reasonable approximation of real agents’ decision making. Rendering our computational model a more reliable laboratory of post-secondary education requires, in our view, not so much a more refined calibration of our computational agents as a more refined mapping of post-secondary education to our computational model.

References

- ADELMAN, C. (2000): “A Parallel Universe: Certification in the Information Technology Guild,” *Change*, May/June, 20 – 29.
- ALTONJI, J. G., AND T. DUNN (1996): “The Effects of Family Characteristics on the Return to Education,” *Review of Economics and Statistics*, 78, 692–704.
- ANONYMOUS (2006): “Duke Professor Plans for-Profit College with Classic Curriculum,” *The Wilmington Star on the Web*, July 23, 2006.

- ARTHUR, W. B. (1991): “Designing Economic Agents That Act Like Human Agents: A Behavioral Approach to Bounded Rationality,” *American Economic Review*, 81(2), 353–59.
- (1994): “Inductive Reasoning and Bounded Rationality,” *American Economic Review*, 84(2), 406–11.
- AVERY, C., A. FAIRBANKS, AND R. ZECKHAUSER (2003): *The Early Admissions Game: Joining the Elite*. Harvard University Press, Cambridge, MA.
- AXELROD, R. (1997): *The Complexity of Cooperation: Agent-Based Models of Conflict and Cooperation*. The Princeton University Press, Princeton, N.J.
- BEHRMAN, J. R., M. R. ROSENZWEIG, AND P. TAUBMAN (1996): “College Choice and Wages: Estimates Using Data on Female Twins,” *Review of Economics and Statistics*, 78, 672–85.
- BOYLAN, J. F. (1998): *Getting In*. Warner Books, New York, N.Y.
- BRENEMAN, D. W. (1994): *Liberal Art Colleges: Thriving, Surviving, or Endangered?* The Brookings Institution, Washington, DC.
- BRENEMAN, D. W. (2005): “Entrepreneurship in Higher Education,” in *Arenas of Entrepreneurship: Where Nonprofit and For-Profit Institutions Compete*, ed. by B. Pusser, pp. 3–9. Jossey-Bass., San Francisco, CA.
- BRENEMAN, D. W., B. PUSSER, AND S. TURNER (eds.) (2006): *Earnings from Learning. The Rise of For-Profit Universities* Albany, NY. State University of New York Press.
- CAMERER, C. (2003): *Behavioral Game Theory*. Princeton University Press, Princeton, NJ.
- CHATTOE, E. (1998): “Just How (Un)Realistic are Evolutionary Algorithms as Representation of Social Process?,” *Journal of Artificial Societies and Social Simulation*, 1(3).

- CHEN, S.-H., J. DUFFY, AND C.-H. YEN (2002): “Equilibrium Selection Via Adaptation: Using Genetic Programming to Model Learning in a Coordination Game,” *The Electronic Journal of Evolutionary Modeling and Economic Dynamics*, 1, 1–44.
- COSMIDES, L., AND J. TOOBY (1996): “Are Humans Good Intuitive Statisticians After All? Rethinking some Conclusions of the Literature on Judgment under Uncertainty,” *Cognition*, 58, 1–73.
- COWAN, N. (2001): “The Magical Number 4 in Short–Term Memory: A Reconsideration of Mental Storage Capacity,” *Behavioral and Brain Sciences*, 24, 87–185.
- EPPLE, D., AND R. E. ROMANO (1998): “Competition Between Private and Public Schools, Vouchers, and Peer–Group Effects,” *American Economic Review*, 88, 33–62.
- FACCHINA, B., E. SHOWELL, AND J. STONE (1993): “Privileges and Exemptions Enjoyed by Nonprofit Organizations,” *University of San Francisco Law Review*, 28, 85–121.
- GALE, D., AND L. SHAPLEY (1962): “College Admissions and Stability of Marriage,” *American Mathematical Monthly*, 69, 9–15.
- GIGERENZER, G., P. M. TODD, AND ABC RESEARCH GROUP (1999): *Simple Heuristics That Make Us Smart*. Oxford University Press, Oxford, UK.
- GLAESER, E. L., AND A. SHLEIFER (2001): “Not–for–Profit Entrepreneurs,” *Journal of Public Economics*, 81, 99–115.
- GOETZ, M., J. J. SIEGFRIED, AND H. ZHANG (1991): “Estimating Economies of Scale in Higher Education,” *Economics Letters*, 17, 203 – 208.
- HOLLAND, J. H. (1992): *Adaptation in Natural and Artificial Systems*. MIT Press, Cambridge, MA.

- HOLLENBECK, J. R. (2000): “Consequences of Organizational Reward Systems,” in *Computational Modeling of Behavior in Organizations : The Third Scientific Discipline*, ed. by D. R. Ilgen, and C. L. Hulin. APA Books, Washington, D.C.
- KIRMAN, A. P., AND N. J. VRIEND (2001): “Evolving Market Structure: An ACE Model of Price Dispersion and Loyalty,” *Journal of Economic Dynamics and Control*, 25, 459–502.
- KIRP, D. L. (2003): *Shakespeare, Einstein, and the Bottom Line. The Marketing of Higher Education*. Harvard University Press, Cambridge, MA.
- KRAFT, E., AND M. VODOPOVIEC (2003): “The New Kids on the Block: The Entry of Private Business Schools in Transition Economies,” *Education Economics*, 11, 239–257.
- LABAND, D. N., AND B. F. LENTZ (2004): “Do Costs Differ Between For–Profit and Not–for–Profit Producers of Higher Education?,” *Research in Higher Education*, 45(1), 429–41.
- LAZEAR, E. P. (2001): “Educational Production,” *Quarterly Journal of Economics*, 116, 777–803.
- LETTAU, M., AND H. UHLIG (1999): “Rules of Thumb versus Dynamic Programming,” *American Economic Review*, 89, 148–74.
- MICHALEWICZ, Z. (1999): *Genetic Algorithms + Data Structures = Evolution Programs*. Springer-Verlag, Berlin Heidelberg New York.
- MITCHELL, M. (1996): *An Introduction to Genetic Algorithms*. MIT Press/Bradford, Cambridge, MA.
- NEWMAN, F., L. COUTURIER, AND J. SCURRY (2004): *The Future of Higher Education: Rhetoric, Reality, and the Risks of the Market*. Jossey–Bass, San Francisco, CA.

- ORTMANN, A. (1997): “How to Survive in Post-Industrial Environments: Adam Smith’s Advice for Today Colleges and Universities,” *Journal of Higher Education*, 68, 483–501.
- (2001): “Capital Romance: Why Wall Street Fell in Love with Higher Education,” *Education Economics*, 9(3), 293–311.
- ORTMANN, A. (2006): “Capital Romance. Why Wall Street Fell in Love With Higher Education,” in *Earnings from Learning: The Rise of For-Profit Universities*, ed. by D. W. Breneman, B. Pusser, and S. E. Turner, pp. 145–166. SUNY Press, Albany, NY.
- ORTMANN, A., S. SLOBODYAN, AND S. S. NORDBERG (2003): “(The Evolution of) Post-Secondary Education: A Computational Model and Experiments,” CERGE–EI Working Paper 208.
- PINGLE, M., AND L. TESHATSION (2001): “Non-Employment and the Evolution of Worker-Employer Cooperation: Experiments with Real and Computational Agents,” mimeo.
- PUSSER, B. (2005): *Arenas of Entrepreneurship: Where Nonprofit and For-Profit Institutions Compete*. Jossey-Bass., San Francisco, CA.
- ROTH, A. E. (2002): “The Economist as Engineer: Game Theory, Experimentation, and Computation as Tools for Design Economics,” *Econometrica*, 70, 1341–78.
- ROTH, A. E., AND M. A. O. SOTOMAYOR (1990): *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis*. Cambridge University Press, Cambridge, MA.
- ROTH, A. E., AND X. XING (1994): “Jumping the Gun: Imperfections and Institutions Related to the Timing of Market Transactions,” *American Economic Review*, 84, 992–1044.

- (1997): “Turnaround Time and Bottlenecks in Market Clearing: Decentralized Matching in the Market for Clinical Psychologists,” *Journal of Political Economy*, 105, 284–329.
- ROTHSCHILD, M., AND L. J. WHITE (1995): “The Analytics of the Pricing of Higher Education and Other Services in Which the Customers are Inputs,” *Journal of Political Economy*, 103, 573–623.
- RUCH, R. S. (2001): *Higher Ed, Inc.: The Rise of the For-Profit University*. John Hopkins University Press, Baltimore London.
- SPERLING, J. G. (2000): *Rebel With a Cause: The Entrepreneur Who Created the University of Phoenix and the For-Profit Revolution in Higher Education*. John Wiley & Sons, New York, NY.
- SPERLING, J. G., AND R. TUCKER (2006): *For-Profit Higher Education: Developing a World-Class Adult Workforce*. Transaction Publishers, Phoenix, AZ.
- TAMURA, R. (2001): “Teachers, Growth, and Convergence,” *Journal of Political Economy*, 109(5), 1021–59.
- TESFATSION, L. (2002): “Agent-Based Computational Economics: Growing Economies from the Bottom Up,” mimeo.
- THURSBY, J. G. (2000): “What Do We Say About Ourselves and What Does It Mean? Yet Another Look at Economics Department Research,” *Journal of Economic Literature*, 38, 383–404.
- VALENTE, M. (2002): “Comments On "Equilibrium Selection Via Adaptation: Using Genetic Programming to Model Learning in a Coordination", by Chen, Duffy and Yeh,” *Electronic Journal of Evolutionary Modeling and Economic Dynamics*, <http://www.e-jemed.org>, 1021.

- VALENTE, M., AND E. S. ANDERSEN (2002): “A Hands-on Approach to Evolutionary Simulation: Nelson-Winter Models in the Laboratory for Simulation Development,” *Electronic Journal of Evolutionary Modelling and Economic Dynamics*, [http:// www.e-jemed.org](http://www.e-jemed.org), 1003.
- VRIEND, N. J. (1995): “Self-Organization of Markets: An Example of a Computational Approach,” *Computational Economics*, 8, 205–31.
- WASHBURN, J. (2005): *University, Inc.: The Corporate Corruption of American Higher Education*. Basic Books, New York, NY.
- WEISBUCH, G., A. KIRMAN, AND D. HERREINER (2000): “Market Organisation and Trading Relationships,” *The Economic Journal*, 110, 411–36.
- WHITLEY, D. (1989): “The GENITOR Algorithm and Selection Pressure: Why Rank-Based Allocation of Reproductive Trials is Best,” in *Proceedings of the Third International Conference on Genetic Algorithms*, ed. by J. D. Schaffer, pp. 116–21, San Mateo, CA. Morgan Kaufman.
- ZEMSKY, R., S. SHAMAN, AND M. IANNOZZI (1997): “In Search of Strategic Perspective: A Tool for Mapping the Market in Post Secondary Education,” *Change*, May/June, 23 – 38.

A Theoretical equilibrium configurations

In the following we describe the symmetric steady state of the Q-model. “Symmetric steady state” denotes situations where every firm serves the same number of agents, where every firm has the same profit share (which is indeed what we observe empirically), and where the quality of the firm equals the average quality of its consumers. We calibrate the model so that the average firm quality \bar{Q} equals the average consumer quality of 50 (which, given our assumption of uniform distribution of consumers along the quality spectrum $[0,100]$, is what we can expect on average). We note that symmetric steady state implies $\bar{Q} = 50$ but that the reverse implication does not necessarily hold. Since firm quality is defined as a weighted sum of both average consumer quality and profits, we begin with the profit weight calibration before proceeding with an analysis of the equilibrium number of clusters and, in fact, the exact location of the clusters (cluster configuration).

A.1 Profit weight calibration

A firm’s quality is updated according to the following rule:

$$Q = w_1 \cdot Q_{avg} + w_2 \cdot \pi, \quad (1)$$

where Q is the firm’s quality, Q_{avg} the average quality of its consumers, π is the firm’s profits, and w_1 and w_2 are weights. Symmetric steady state profits are given by

$$\pi = \frac{N}{[Q]} \alpha Q,$$

where N is the number of consumers per firm (100 in all runs), $[Q]$ the quality range (100 in all runs), and α the profit share (average α is 0.46 – 0.48 for different configurations, with a standard deviation 0.02 – 0.03). Note that α is determined experimentally.

The requirement that $Q = Q_{avg}$ amounts to $Q(1 - \alpha w_2) = Q w_1$, or $w_2 = (1 - w_1)/\alpha$. The empirical value of w_2 which prompts $\bar{Q} = 50$ is indeed very close to the

one just derived. For example, with 24 firms and $\alpha = 0.48$, $w_1 = 0.95$, the derived value is $w_2 = 0.107$, while $\bar{Q} \approx 50$ requires an empirical value of $w_2 \approx 0.104$.

A.2 Equilibrium number of clusters

In this subsection we show why the configuration that we observe in most runs with the number of firms modulo 6 (6 relatively tight clusters of firms) is a stable symmetric steady state for our choice of the quality range. For the sake of argument, assume that firms' quality is adjusted according to (1) with $w_1 = 1$ and $w_2 = 0$, *i.e.*, a firm's quality equals the average quality of its consumers. (Runs with this quality adjustment rule reveal the same distribution of 6 relatively tight clusters of firms). Additionally assume that if a given number of T consumers can be served by n firms, then T/n of them will be served by every firm, that is, competition leads to even distribution of consumers among firms in equilibrium. A firm will accept customers who are at least of quality $Q - \Delta$, where Q is the firm's quality. A customer will accept a firm that has at least quality of $Q_{cust} - \Delta$ where Q_{cust} is the customer's quality. Therefore, a firm with quality Q can serve only customers in the quality interval $[Q - \Delta, Q + \Delta]$.

Assume that all firms are in steady state. Assume next that one of them has, by some random disturbance, its quality adjusted upwards by dq . The firm under consideration loses some consumers at the lower end of its segment at quality $Q - \Delta$ but also obtains some consumers at the upper end of its segment at quality $Q + \Delta$. If n other firms are competing at the lower end and $n + j$ other firms can serve consumers at the upper end, the number of consumers lost and obtained are respectively $\frac{1}{n+1} \frac{dq}{[Q]} N_{tot}$ and $\frac{1}{n+j+1} \frac{dq}{[Q]} N_{tot}$, where N_{tot} is the total number of consumers. Thus, the new average quality of the firm is given by

$$\tilde{Q} = \frac{\sum Q - \frac{1}{n+1} \frac{dq}{[Q]} N_{tot} \cdot (Q - \Delta) + \frac{1}{n+j+1} \frac{dq}{[Q]} N_{tot} \cdot (Q + \Delta)}{N - \frac{1}{n+1} \frac{dq}{[Q]} N_{tot} + \frac{1}{n+j+1} \frac{dq}{[Q]} N_{tot}},$$

where $\sum Q$ is the sum of the firm's consumers' quality in steady state and equal to $Q \cdot N$ by assumption. Dividing the numerator and denominator by $\sum Q$ and N

respectively and using the fact that $\frac{1+x}{1+y} \approx 1 + x - y$ for $x \ll 1, y \ll 1$, one obtains

$$\tilde{Q} \approx Q + \frac{dq}{[Q]} \frac{N_{tot}}{N} \left\{ \frac{Q + \Delta}{n + j + 1} - \frac{Q - \Delta}{n + 1} - \frac{Q}{n + j + 1} + \frac{Q}{n + 1} \right\},$$

or

$$dq' = \tilde{Q} - Q \approx \frac{dq}{[Q]} \frac{N_{tot}}{N} \cdot \Delta \cdot \frac{n + 1 + n + j + 1}{(n + 1)(n + j + 1)}.$$

The steady state is stable if random fluctuations in quality are dampened over time, or $|dq'| < |dq|$.³⁰ Therefore, the stability of the steady state depends on the magnitude of the following term:

$$\frac{\Delta}{[Q]} N_f \cdot \frac{n + 1 + n + j + 1}{(n + 1)(n + j + 1)}, \quad (2)$$

where $N_f = \frac{N_{tot}}{N}$ is the number of firms in the economy.

Let us consider some special cases of (2). Suppose that a firm in steady state does not have any competition at the lower end of its segment, $n = 0$. Then (2) becomes $\frac{\Delta}{[Q]} N_f \cdot \frac{2+j}{1+j}$, and for parameter values ($\Delta = 10, [Q] = 100$) this expression is greater than one for any $j > 0$, and any $N_f \geq 10$. In other words, any steady state that implies no competition at the lower end is not stable, because a random upward quality movement is amplified. Similarly, suppose that there is no competition at the upper end of a firm's segment. In this case, $j = -n$, and (2) is $\frac{\Delta}{[Q]} N_f \cdot \frac{2+n}{1+n}$ which is again greater than one for any $n > 0$, and any $N_f \geq 10$. Therefore, a steady state involving zero competition at the upper end cannot be stable.

The preceding result demonstrates that steady states with fewer than five clusters are unstable, because they necessarily involve zero competition either at the lower or at the upper end of the quality segment. How about five segments then? Assume a steady state with five firm clusters, numbered in ascending quality order. Given the parameter values that we used for our treatments, $\Delta = 10, [Q] = 100$, the five firm clusters will be located at qualities 10, 30, 50, 70, and 90. Suppose now that clusters number 2 and 4 move down and up, respectively. In this case, a firm

³⁰In other words, we want the eigenvalue of the difference equation $Q_{n+1} = f(Q_n)$, linearized around the steady state, to be less than one. It is always positive, therefore oscillating dynamics around the steady state are impossible.

from cluster 3 that randomly increased its quality by dq will have $dq' > dq$, while the one that had its quality decreased by dq will have $|dq'| > |dq|$. (Recall that dq' is the deviation from steady state after one iteration.) In other words, cluster 3 will be torn apart by any non-negligible simultaneous movements of clusters 2 and 4. Therefore, a configuration with 5 clusters is stable but the associated basin of attraction is very small. In numerical simulations with $N_f = 10$ we have observed stable constellations with 5 clusters of firms only once or twice every 100 runs.

Why, then, do we observe constellations of 6 clusters for runs with a large number of firms, say 24 and 48? And why do we observe constellations with between 6 and 8 clusters for runs with 10 firms? Compare two steady states, one with C clusters and another with $C + 1$, where $10 > C > 5$. A firm that moved up by dq faces the same competition at its lower end from members of its own cluster and the lower one, with the total number of other firms given by $\frac{N_f}{C} - 1 + \frac{N_f}{C}$ (disregarding integer constraints). On the other hand, at the upper end of its segment, competition from members of its own cluster disappears and only that from the upper cluster remains. Therefore, $j = 1 - \frac{N_f}{C}$. (2) is now proportional to $\frac{\Delta}{[Q]} N_f \cdot \frac{3^{\frac{N_f}{C}+1}}{2^{\frac{N_f}{C}} \cdot (2^{\frac{N_f}{C}+1})}$ or

$$\frac{\Delta}{[Q]} N_f \cdot \frac{C \cdot (3N_f + C)}{2N_f \cdot (2N_f + C)}. \quad (3)$$

The partial derivative of the preceding expression with respect to C is proportional to

$$\frac{2N_f \cdot (6N_f^2 + 4CN_f + C^2)}{4N_f^2 \cdot (2N_f + C)^2},$$

which is always positive. Therefore, the movement to a higher number of clusters implies a larger eigenvalue, and hence a less stable steady state.³¹

Summarizing the results, we see that configurations with 4 clusters are unstable and those with 5 clusters are likely to be destroyed even by small fluctuations. Furthermore, configurations with more than 6 clusters are less stable than those with 6, and indeed they are increasingly less stable as the number of clusters goes

³¹A similar result is true for any number of clusters. The math, however, becomes tedious.

up. Therefore, in numerical simulations one is likely to observe a configuration with 6 clusters.

Finally, observe that with $C = 6$, (3) equals 0.42 for $N_f = 12$, 0.433 with $N_f = 24$, and approaches 0.45 as $N_f \rightarrow \infty$. This means that configurations of 6 clusters are always stable for any number of firms.

A.3 Cluster configurations

Having established theoretically the most likely distribution of clusters, we next calculate their exact location in the symmetric steady state with C clusters. We assume that there is an equal number of firms in each cluster. Under the symmetric steady state assumptions spelled out in the previous subsection, calculations are the same for one or n firms in a cluster; we thus restrict our discussion to one firm per cluster.

Order quality locations in a symmetric steady state in ascending order from Q_1 to Q_C . For $10 \geq C \geq 5$, the first firm (remember we restrict our discussion to one firm per cluster) has no competition at its lower end and competition from the second firm only at the upper end. Denote as D the density of customers per unit of quality. Then the first firm will serve customers located in $[0, Q_2 - \Delta]$ alone and those in $[Q_2 - \Delta, Q_1 + \Delta]$ together with the second firm. Since in the symmetric steady state the average quality of consumers equals its own quality, we have

$$Q_1 = \frac{D \int_0^{Q_2 - \Delta} Q dQ + \frac{1}{2} D \int_{Q_2 - \Delta}^{Q_1 + \Delta} Q dQ}{D \int_0^{Q_2 - \Delta} dQ + \frac{1}{2} D \int_{Q_2 - \Delta}^{Q_1 + \Delta} dQ} = \frac{1}{2} \frac{(Q_2 - \Delta)^2 + (Q_1 + \Delta)^2}{Q_2 - \Delta + Q_1 + \Delta} = \frac{(Q_2 - \Delta)^2 + (Q_1 + \Delta)^2}{2 \cdot (Q_2 + Q_1)}. \quad (4)$$

Consider now the second firm. It is the sole provider to consumers in $[Q_1 + \Delta, Q_3 - \Delta]$ and a joint provider with first and third firm in $[Q_2 - \Delta, Q_1 + \Delta]$ and

$[Q_3 - \Delta, Q_2 + \Delta]$, respectively. The symmetric steady state condition then becomes

$$Q_2 = \frac{\frac{1}{2}D \int_{Q_2-\Delta}^{Q_1+\Delta} QdQ + D \int_{Q_1+\Delta}^{Q_3-\Delta} QdQ + \frac{1}{2}D \int_{Q_3-\Delta}^{Q_2+\Delta} QdQ}{\frac{1}{2}D \int_{Q_2-\Delta}^{Q_1+\Delta} dQ + D \int_{Q_1+\Delta}^{Q_3-\Delta} dQ + \frac{1}{2}D \int_{Q_3-\Delta}^{Q_2+\Delta} dQ} = \frac{(Q_3 - \Delta)^2 + (Q_2 + \Delta)^2 - (Q_2 - \Delta)^2 - (Q_1 + \Delta)^2}{2 \cdot [(Q_3 - \Delta) + (Q_2 + \Delta) - (Q_2 - \Delta) - (Q_1 + \Delta)]}. \quad (5)$$

After some algebra, (5) transforms into

$$Q_2 = \frac{Q_1 + Q_3}{2}, \quad (6)$$

which says that the symmetric steady state location of the second firm is exactly between the first firm and the third firm. It is trivial to show that a similar result holds for all other firms located in the interior of the quality spectrum,

$$Q_3 = \frac{Q_2 + Q_4}{2}, \quad (7a)$$

$$Q_4 = \frac{Q_3 + Q_5}{2}, \quad (7b)$$

$$\dots \quad (7c)$$

$$Q_{C-1} = \frac{Q_{C-2} + Q_C}{2}. \quad (7d)$$

Finally, for the last firm C , the symmetric steady state condition is given by

$$Q_C = \frac{2 \cdot [Q]^2 - (Q_C - \Delta)^2 - (Q_{C-1} + \Delta)^2}{2 \cdot [Q] - Q_{C-1} - Q_C}. \quad (8)$$

Combining (6) and (7) we obtain $Q_4 = 3Q_2 - 2Q_1$, $Q_3 = 2Q_2 - Q_1$ or $Q_3 = Q_2 + (Q_2 - Q_1)$, $Q_4 = Q_3 + 2(Q_2 - Q_1)$. In other words, firms are located at equal distance $\delta = (Q_2 - Q_1)$ from each other. The problem of finding symmetric steady state locations is thus reduced to solving a system of two quadratic equations, (4) and (8), in two unknowns, Q_1 and δ (remember that $Q_2 = Q_1 + \delta$, $Q_{C-1} = Q_1 + (C - 2) \cdot \delta$, $Q_C = Q_1 + (C - 1) \cdot \delta$). The solution can be found numerically when C , the number of clusters in symmetric equilibrium, is given.

In the previous subsection we have argued that, given our parameter values Δ and $[Q]$, the symmetric steady state with 6 clusters should be the most stable one.

Steady state positions with 6 clusters are given by [8.48; 25.09; 41.70; 58.30; 74.91; 91.52].

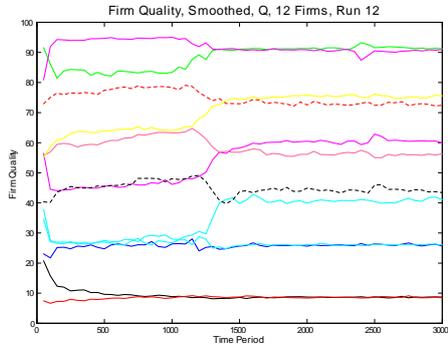


Figure 1

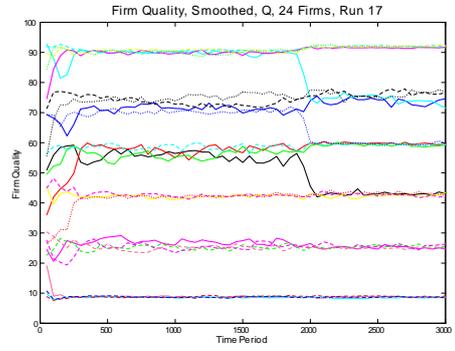


Figure 2

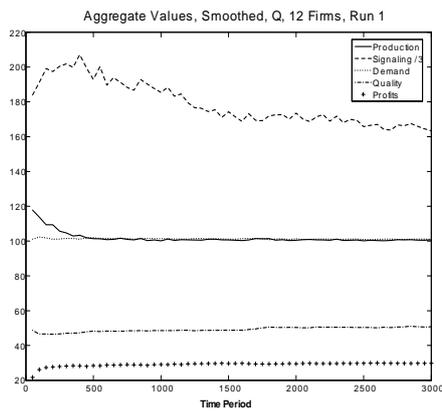


Figure 3

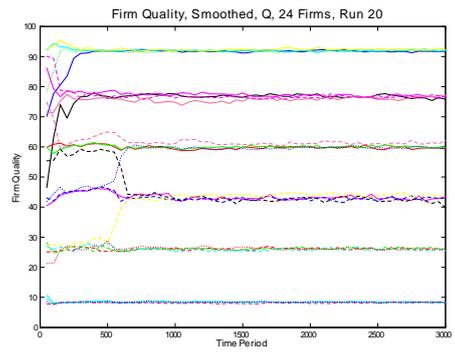


Figure 4

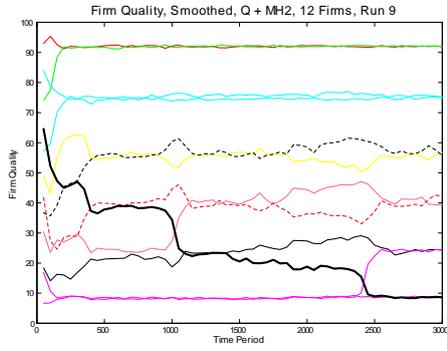


Figure 5. Thick line: mutant.

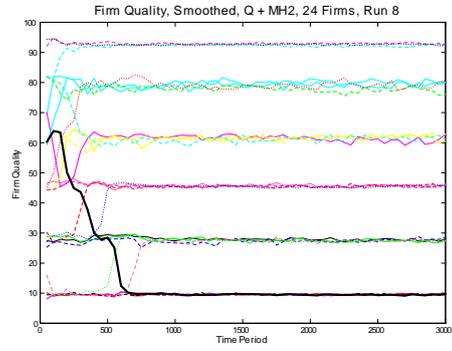


Figure 6. Thick line: mutant.

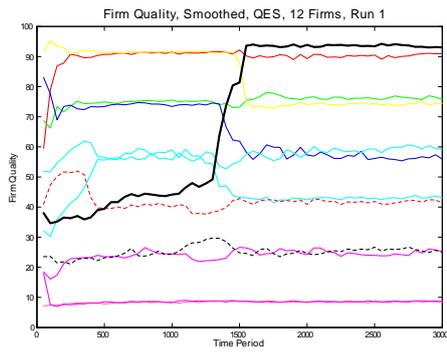


Figure 7. Thick line: mutant.

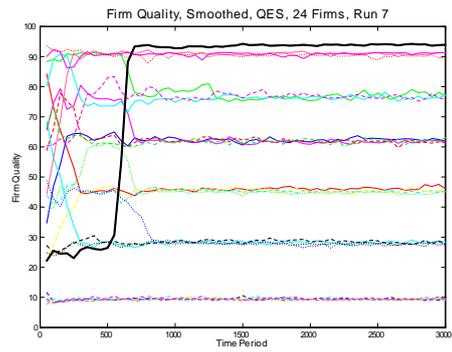


Figure 8. Thick line: mutant.

Individual researchers, as well as the on-line and printed versions of the CERGE-EI Working Papers (including their dissemination) were supported from the following institutional grants:

- Economic Aspects of EU and EMU Entry [Ekonomické aspekty vstupu do Evropské unie a Evropské měnové unie], No. AVOZ70850503, (2005-2010);
- Economic Impact of European Integration on the Czech Republic [Ekonomické dopady evropské integrace na ČR], No. MSM0021620846, (2005-2011);

Specific research support and/or other grants the researchers/publications benefited from are acknowledged at the beginning of the Paper.

(c) Andreas Ortmann, Sergey Slobodyan, 2008.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical or photocopying, recording, or otherwise without the prior permission of the publisher.

Published by
Charles University in Prague, Center for Economic Research and Graduate Education (CERGE)
and
Economics Institute ASCR, v. v. i. (EI)
CERGE-EI, Politických vězňů 7, 111 21 Prague 1, tel.: +420 224 005 153, Czech Republic.
Printed by CERGE-EI, Prague
Subscription: CERGE-EI homepage: <http://www.cerge-ei.cz>

Editors: Directors of CERGE and EI
Managing editors: Deputy Directors for Research of CERGE and EI

ISSN 1211-3298
ISBN 978-80-7343-155-6 (Univerzita Karlova. Centrum pro ekonomický výzkum
a doktorské studium)
ISBN 978-80-7344-144-9 (Národohospodářský ústav AV ČR, v. v. i.)



CERGE-EI
P.O.BOX 882
Politických vězňů 7
111 21 Praha 1
Czech Republic
<http://www.cerge-ei.cz>