# Agent-Based Modeling of Cognitive Double Auction Market Experiments

**Shu-Heng Chen**
Department of Economics
National Chengchi University
Taipei, Taiwan 116
E-mail: chen.shuheng@gmail.com

**Chung-Ching Tai**
Department of Economics
National Chengchi University
Taipei, Taiwan 116
E-mail: chungching.tai@gmail.com

**Lei-Xieng Yang**
Research Center of Brain, Mind and Learning
National Chengchi University
Taipei, Taiwan 116
E-mail: lxyang@nccu.edu.tw

**Abstract**

In this paper, we study the significance of the cognitive capacity in the context of double auction experiments. These experiments are then simulated using artificial agents in an agent-based modeling environment, and the results are compared with the results from human subjects.

**Keywords: Double Auctions, Genetic Programming, Working Memory Capacity, Market Experiments, Programmed Agents**

## 1 Motivation and Introduction

The economic significance of cognitive capacity has recently drawn intensive attention from economists. On the one hand, empirical studies indicating the significant correlation between cognitive capacity and economic performance has piled up[1]; on the other hand, experimental economists has started to treat cognitive capacity as a control variable and inquire its economic consequences based on the conducted experiments [2]. The latter has shaped an interdisciplinary research area, called *cognitive experimental economics*, by overarching economics and psychology.

Despite the active publications of cognitive economic experiments since late 1990s, most existing experiments are carried out in the context of *games*. The significance of

[1]For example, see Herrnstein and Murray (1996), Jensen (1998), Cawley, Conneely, Heckman, and Vytlacil (1997), Zax and Rees (2002), Gould (2005), Lynn and Vanhanen (2002),Lynn (2006), Weede and Kampf (2002), Jones and Schneider (2006), and Ram (2007).

[2]See Segal and Hershberger (1999), Devetag and Warglien (2003), Ohtsubo and Rapoport (2006), Casari, Ham and Kagel (2007), Cornelissen, Dewitte and Warlop (2007), Cappelletti, Guth and Ploner (2008), Devetag and Warglien (2008), Jones (2008), and Burks, Carpenter, Goette, and Rustichini (2009).

cognitive capacity in *market experiments* has, however, rarely been addressed. The lack of study in this area, to some extent, echoes with the assumption that the market mechanism is so powerful and robust that it leaves no room to participants' intelligence, learning, and whatsoever. This assumption is best known as the *zero-intelligence-plus* (ZIP) *hypothesis* (Becker, 1962; Gode and Sunder, 1993, 1994). Nevertheless, this assumption has never been tested in any traditional market experiment. Hence, no matter how sound it might be, the ZIP hypothesis is yet to be verified. In fact, some recent studies using agent-based simulation do indicate the existence of intelligence effect in the double auction market, and the difference is even up to a statistical significance level (Chen and Tai, 2003; Chen, Zeng and Yu, 2008). This paper, therefore, provides the first attempt to examine whether the cognitive capacity of subjects can matter in market experiments, and we take double auction markets as our first step.

The double auction (DA) market experiment is probably the oldest and the most intensively conducted experiment in experimental economics. However, most conventional studies are concerned with only the aggregate outcomes. Rust, Miller, and Palmer (1993, 1994) are the only few exceptions, which study the DA market from the individual perspective. We start with the DA market since it was considered as the least likely place that cognitive ability can have an effect. So, it serves as a *benchmark*.

The rest of the paper is organized as follow. In Section 2, we introduce the *cognitive double auction market experiments* implemented in this paper and analyze the experimental results. In Section 3, we propose the corresponding agent-based computational modeling of the cognitive market experiment.

## 2 Human-Subject Experiments

In the human-subject experiment, each subject is placed within a double auction experiment where his/her opponents are *all* software agents. To differentiate the *complexity* of software agents, two kinds of agents are presented: *truth tellers* and *programmed agents*. Program agents are collected from the literature. That includes the Bayesian agent (Friedman, 1991), the Easley-Ledyard's agents (Easley and Ledyard, 1993), the zero-intelligence-constrained agent (Gode and Sunder, 1993), the skeleton agent, the Kaplan agent, the Ringuette agent (Rust, Miller, and Palmer, 1993), the zero-intelligence-plus agent (Cliff and Bruten, 1997), the belief-based learning agents (Gjerstad and Dickhaut, 1998), the empirical Bayesian agent (Chan, LeBaron, Lo, and Poggio, 1999), and the mark-up agent (Zhan and Friedman, 2007). We shall call the experiments with truth tellers **Exp Series I**, and the experiments with programmed agents **Exp Series 2**.

Exp Series I was conducted with a sample of 46 subjects, whereas Exp Series II was conducted with a sample of 39 subjects. All subjects are students, either graduate students or undergraduate students. Most subjects are inexperienced, i.e., they have little experience in real auction markets, such as the stock markets. Only 13 in the Exp Series I, and 8 in Series II are experienced. The focusing attribute considered in this paper is the *working memory capacity*.

### 2.1 Working Memory Capacity

All subjects are required to perform a working memory capacity (WMC) test before they can collect the monetary reward from the experiments. Five WMC tasks from Oberauer et al. (2000) and Oberauer et al. (2003) were involved in the test.[3] They are backward digit span (BDG), memory updating (MU), operation span (OS), sentence span (SS), and spatial short-term memory (SSTM).

By following the conventional procedure in psychological tests, the score of each task are normalized using the mean and the standard deviation of the subject pool. The five standardized scores of the five tasks will then be averaged to get the WMC of a specific subject. By this procedure, we are able to further classify our human agents into two groups: those whose WMC are *above average* and those whose WMC are *below average*. This classification will facilitate our analysis. In the analysis below, only non-experienced subjects are included so as to avoid the mixture effect of experience and cognitive capacity. For the Exp Series I, we have a total of 33 non-experienced subjects, 18 of them belonging to the above-average group, whereas the rest 15 belonging to the below-average group. For the Exp Series II, we have a total of 31 non-experience subjects, 18 of them having a above-average WMC and 13 of them having a below-average WMC.

### 2.2 Demand and Supply Schedules

Three different demand and supply schedules are used in both Exp Series I and Exp Series II. To propose these schedules, we basically manipulate two things: one is the trading opportunities, and the other is the range (or scale) of possible trading prices. For the former consideration, we make subjects with extra-marginal units may still have chances to "steal" the trade from subjects with intra-marginal units. For making comparisons, these three are the same across the two series, and they are shown in Figures 4 to 6 (See Appendix A). Each subject is required to go through all these three markets, denoted by M1, M2 and M3, and for each market six iterations (six trading days).

### 2.3 Experimental Results

Their performance in terms of efficiency indexes are recorded and are shown in Figure 1. Individual subjects associated with the above-average group and the below-average group are separated by blue and red lines. In these figures, red lines and blue lines are jumping around, and it is hard to see whether working memory capacity has an effect on these individuals' performance. We, therefore, turn to see whether there is a *group difference*. To do so, we compare the average performance of the below-average group with that of the above-average group. The result is shown in Figure 2. A quick look at the results seems to indicate that there is a performance difference between the two groups. Subjects whose working memory capacities are above the average tend to perform better than subjects whose working memory capacities are below average.

Another thing deserving our attention is that in EXP 2, the case of M2 and M3, the efficiency index achieved by human subjects for both groups are even higher than 100%.

[3]For these WMC tasks, the interested reader is also referred to the web, http://eis.bris.ac.uk/ psxko/

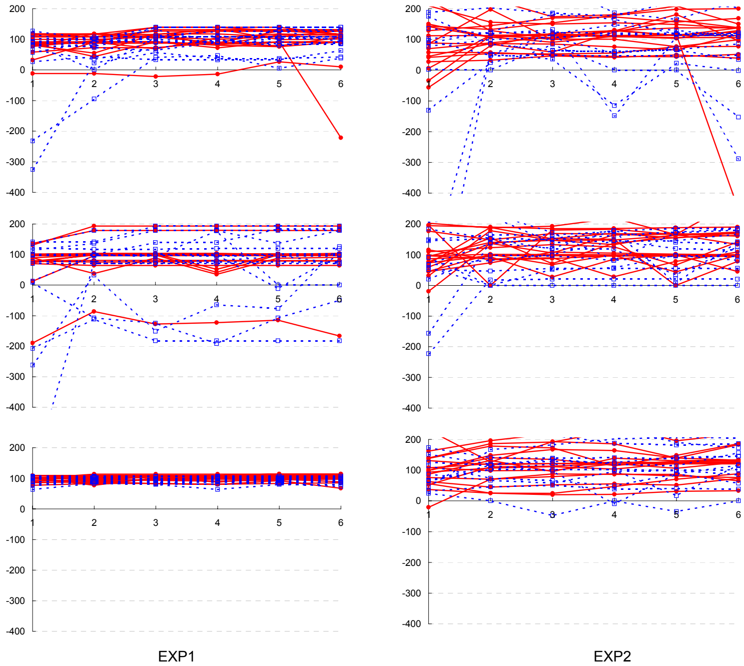

Figure 1: The Efficiency Index of Individual Subjects

This superior performance is not seen in EXP 1. This result, at first sight, may be a little counterintuitive because human agents can play even better when his opponents are more sophisticated than just truth tellers. However, when all opponents are all truth tellers, all bids and asks must be very competitive and, given this situation, it would be infeasible for human subjects with extra-marginal to "steal" trades from the those truth tellers with intra-marginal units. Alternatively, when opponents get more sophisticated and possibly more greedy, bids and asks may not be so competitive as in EXP 1, it, therefore, leaves human-subjects to develop some heuristics to steal trades from others.

Alternatively, instead of seeing the group difference, we also examine the capability of the WMC on explaining the performance at an individual level. A simple linear regression is proposed to fit the experimental data.

$$Y_{i,t} = \alpha_t + \beta_t X_i + \epsilon_{i,t}, \tag{1}$$

where $Y_{i,t}$ refers to the efficiency index of subject $i$ at period $t$, $X_{i,t}$ refers to the working memory capacity of subject $i$, and $\epsilon_{i,t}$ is the usual error term. Here, from a convergence or static viewpoint, only the last period ($t = 6$) is considered.

The regression result of the two market series are shown in Table 1. The result shows that the coefficient of the working memory capacity is insignificant in most cases. In addition to this one-period snapshot, to examine the possible influence of the WMC under learning, we also trace the correlation between the WMC and the efficiency index in each period, and the result is shown in Table 2.

The table shows two possibilities. In the first case, the WMC had the *initial effect*, but then the effect quickly or gradually disappeared when all subjects were subjected to

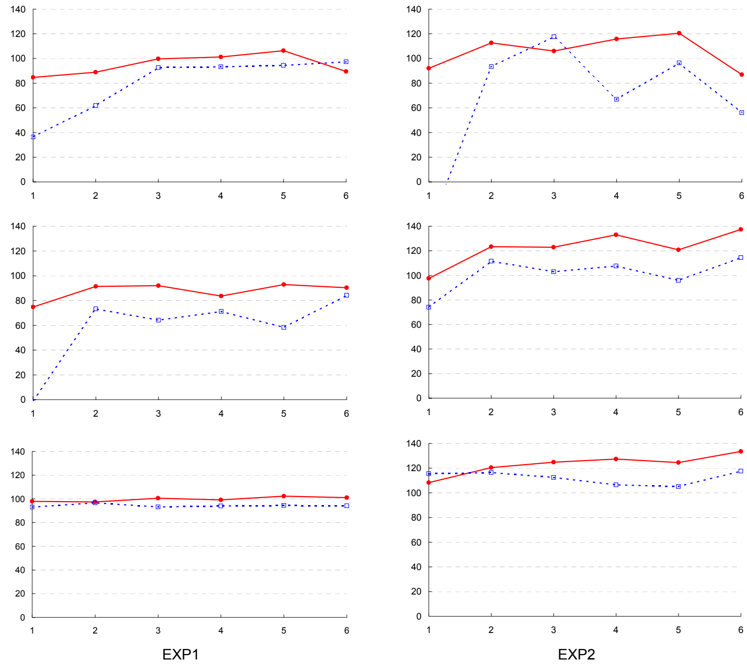


Figure 2: The Efficiency Index of Individual Subjects

learning. In the second case, the WMC had little effect even in the initial period. The only exception that we see that the WMC has a persistent positive effect is Market 3 (M3) of Exp Series I.

# 3 Agent-Based Simulations

## 3.1 Cognitive Capacity of Autonomous Agents

In the case of agent-based simulation, we replace the human agents with *autonomous agents*, the autonomous agents are modeled with *genetic programming*. Using genetic programming (GP) to model agents' adaptive behavior is not new, for example, Chen and Tai (2003). However, unlike those earlier studies, *we no longer assume that agents are equally smart*. Instead, we manipulate one control parameter of GP so that the agents' "cognitive capacity" can be "born" differently. The parameter which we manipulate in this paper is the *population size*.

The idea of using population size as a proxy variable for working memory is first proposed in Casari (2004), who literally treated the population size used in the genetic algorithm equivalent to the number of chunks that human can process at a time. According to the famous "$7 \pm 2$" rule proposed by Miller (Miller, 1956), the capacity lies between five to nine. Casari (2004) then set the population size of *genetic algorithms* to 6, "which implies that decision-makers have a hardwired limitation in processing information at six strategies at a time. (Ibid, p.261)" This is the probably the earliest article which connects the *population size* used in evolutionary computation to *working memory capacity* in cognitive psychology. However, in Casari (2004), agents are still treated as *homogeneous*.

Table 1: Regression Results of EXP1

| | | Estimate | Standard Error | t Statistic | p-value | $R^2$ | $\bar{R}^2$ |
|---|---|---|---|---|---|---|---|
| I-M1 | $\alpha$ | 93.004 | 11.280 | 8.245 | 0.000 | 0.001 | -0.030 |
| | $\beta$ | -3.868 | 18.996 | -0.204 | 0.840 | | |
| I-M2 | $\alpha$ | 87.61 | 14.79 | 5.925 | 0.000 | 0.013 | -0.018 |
| | $\beta$ | 16.13 | 24.90 | 0.648 | 0.522 | | |
| I-M3 | $\alpha$ | 97.924 | 1.730 | 56.595 | 0.000 | 0.195 | 0.169 |
| | $\beta$ | 8.001 | 2.914 | 2.746 | 0.009 | | |
| II-M1 | $\alpha$ | 73.985 | 25.320 | 2.922 | 0.006 | 0.000 | -0.034 |
| | $\beta$ | -1.729 | 41.571 | -0.042 | 0.967 | | |
| II-M2 | $\alpha$ | 127.760 | 10.439 | 12.238 | 0.000 | 0.000 | -0.034 |
| | $\beta$ | -1.659 | 17.139 | -0.097 | 0.924 | | |
| II-M3 | $\alpha$ | 126.851 | 10.588 | 11.980 | 0.000 | 0.001 | -0.033 |
| | $\beta$ | 3.019 | 17.384 | 0.174 | 0.863 | | |

Table 2: Correlation Coefficient between the WMC and the Efficient Index

| Period | EXP1-M1 | EXP1-M2 | EXP1-M3 | EXP2-M1 | EXP2-M2 | EXP2-M3 |
|---|---|---|---|---|---|---|
| 1 | 0.21 | 0.27 | 0.24 | 0.25 | 0.07 | −0.18 |
| 2 | 0.25 | 0.18 | 0.12 | −0.01 | −0.14 | −0.07 |
| 3 | 0.21 | 0.17 | 0.52 | −0.27 | 0.06 | −0.04 |
| 4 | 0.17 | 0.11 | 0.40 | 0.08 | 0.10 | 0.07 |
| 5 | 0.20 | 0.20 | 0.54 | 0.06 | 0.09 | 0.10 |
| 6 | −0.04 | 0.12 | 0.44 | −0.01 | −0.02 | 0.03 |

Like genetic algorithms, Genetic programming is a population-based algorithm, which can implement parallel processing. Hence, on the one hand, the size of the population will directly determine the capability of parallel processing. On the other hand, the human's working memory capacity is frequently tested based on the number of the cognitive tasks which humans can simultaneously process (Cappelletti, Guth and Ploner, 2008). Dual tasks have been used in hundreds of psychological experiments to measure the attentional demands of different mental activities (Pashler, 1998). Hence, *the population size seems to be an appropriate choice with regard to mimicking the working memory capacity of human agents*.

### 3.2 Simulation Designs

We consider three different population sizes: 5, 25, and 125. These artificial agents with different population sizes are therefore assumed to match human agents with different working memory capacities. These artificial agents are then placed into the same two experiment series with the same three markets used in the human subject experiments (Section 2). We shall call two simulation series, corresponding to Exp Series I and II, Sim I and Sim II. Hence, in Sim I the GP trader has truth tellers as his opponents, whereas in

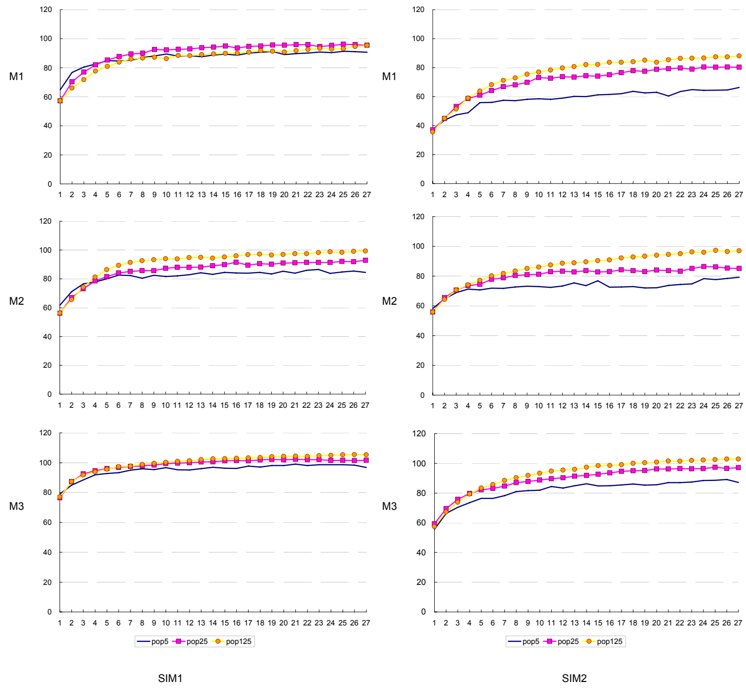


Figure 3: The Efficiency Index of Individual Subjects

Sim II the GP trader has all programmed agents as his opponents.

240 runs were conducted for each market of each Exp Series. In each single run, the position of the GP trader and its opponents are randomly determined or generated. If we consider one GP trader equivalent to one human subject, then the sample size of "subjects" used in the agent-based simulation is about eight times of the size used in the human-subject experiments.

### 3.3 Simulation Results

Since the purpose of this paper is to see whether we can successfully manipulate the artificial cognitive capacity so that it can, to some extent, lead to the same effects caused by human's cognitive capacity, we shall organize our simulation results in a way which is nicely parallel to those of human subject experiments. So, first, we ask whether different working memory capacity will lead to different earning performance. To answer this question, we organize our results so that it can be comparable to Figure 2 in human-subject experiments. This lead to the time series plot displayed in Figures 3. What were plotted in these two figures are the average efficient index taking over the 240 simulations, each lasting for 27 periods (trading days).[4] As in Figure 2, we plot the three groups together for each market so that we can easily compare their difference. Eyes-browsing

[4]This is longer than, at least superficially, than what we did for human-subject experiments, which lasts for only six days. Obviously, we can do this longer for software agents because they would only cost us very little for longer time. So, when running agent-based simulation, we may have a better chance to observe the behavior at a larger or longer span of time.

Table 3: Significance of the Differences Between GP Traders with Various Population Sizes–SIM1

| SIM I | P5 | P25 | P125 |
|---|---|---|---|
| P5 | X | | |
| P25 | 0.2168 ($M1$)<br>0.3690 ($M2$)<br>0.004758** ($M3$) | X | |
| P125 | 0.1416 ($M1$)<br>0.003733** ($M2$)<br>0.00000007873** ($M3$) | 0.3660 ($M1$)<br>0.1467 ($M2$)<br>0.0004625** ($M3$) | X |
| **SIM II** | **P5** | **P25** | **P125** |
| P5 | X | | |
| P25 | 5.538e-07** (M1)<br>0.3778 ($M2$)<br>0.004683** ($M3$) | X | |
| P125 | 2.82e-16** (M1)<br>0.0005076** ($M2$)<br>8.201e-06** (M3) | 0.0009483** ($M1$)<br>0.006392** ($M2$)<br>0.07213* ($M3$) | X |

the results indicates that cognitive capacity can affect the individuals' performance.

For a detail examination, the Wilcoxon Rank Sum Test to the data is further applied to test each pair of two working memory capacities (population sizes), say 5 vs. 25, 5 vs. 125 and 25 vs. 125, over the 3 different markets (M1 to M3) so that we can test the hypothesis that increasing in the WMC can lead to a superior performance in earning capacity. The results are given in Tables 3. In the truth-teller scenario (Sim Series I), the intelligence irrelevancy hypothesis is only consistently rejected in Market 3 (M3). For M2 and M3, the hypothesis is failed to reject on many pairs, for example, the 5-25 and 5-125 pair in M1, and the 25-125 pair in M2. In the programmed-agent scenario (Sim Series II), the null is rejected in almost all pairs and all markets with the only exception on the 5-25 pair in M1.

**Comparisons with Human-Subject Experiments** The test shows that cognitive capacity matters only in Market 3 under Sim Series I, but matters in almost all markets under Sim Series II. Comparing these results with those of human-subjects experiments, we have matches in Series I, but mismatches in Series II. For the latter, the intelligence irrelevancy hypothesis is failed to reject in all three market experiments with human subjects.

# 4 Concluding Remarks

An attempt to study the economic significance of the double auction markets is made both in human-subject experiments and agent-based simulation. While only weak evidence of the economic significance of cognitive capacity is found in human-subject experiments, our agent-based simulations indicate somewhat stronger association. In addition, human agents perform much better when all opponents are all programmed agents. They perform even better than what our autonomous agents did. Therefore, finding the incarnation of human agents in terms of software agents remains to be a challenging task.

## Acknowledgements

National Science Council Research Grant No. NSC 95-2415-H-004-002-MY3, NSC 98-2410-H-004-045-MY3 and National Chengchi University Top University Program No. 98H432 are also gratefully acknowledged.

# Appendix A Appendix

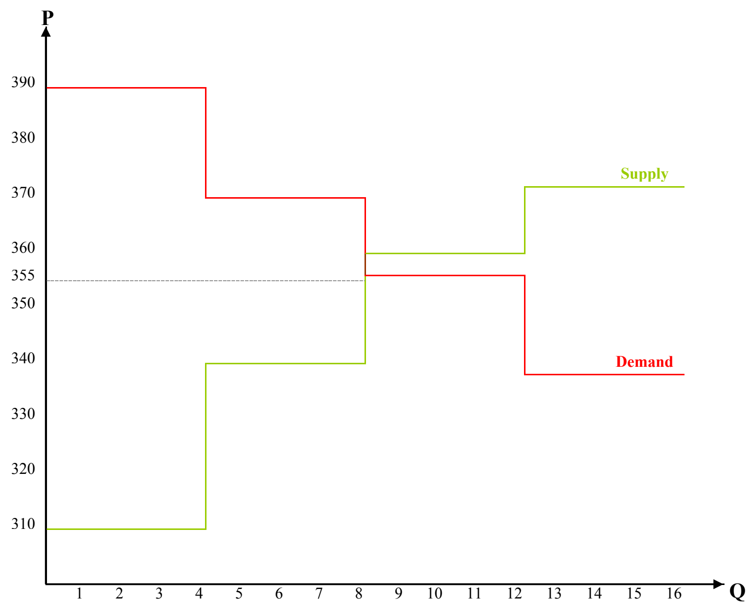


Figure 4: The Demand-Supply Schedule: M1

## References

Becker G (1962). Irrational behaviour and economic theory. Journal of Political Economy 70:1-13.

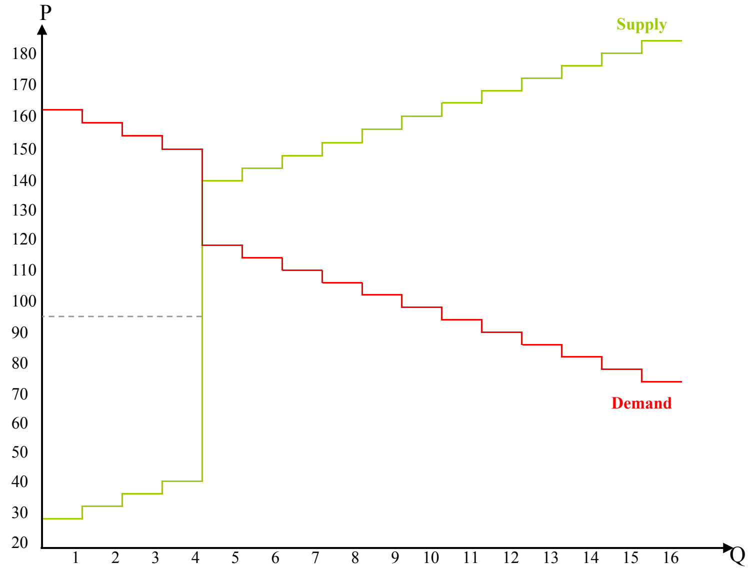

Figure 5: The Demand-Supply Schedule: M2

Burks S, Carpenter J, Goette L, Rustichini A (2009) Cognitive skills affect economic preferences, strategic behavior, and job attachment. Proceedings of the National Academy of Sciences 106(19):7745-7750.

Cappelletti D, Guth W, Ploner M (2008) Being of two minds: An ultimatum experiment investigating affective processes. Jena Economic Research Papers 2008-048, The Friedrich Schiller University and the Max Planck Institute of Economics, Jena, Germany.

Casari M (2004) Can genetic algorithms explain experimental anomalies? An application to common property resources. Computational Economics 24:257-275.

Casari M, Ham J,. Kagel J (2007) Selection bias, demographic effects, and ability effects in common value auction experiments. American Economic Review 97(4):1278-1304.

Cawley J, Conneely K, Heckman J, Vytlacil E (1997) Cognitive ability, wages, and meritocracy, in Devlin B, Fienber S, Resnick D, Roeder K (eds), Intelligence, Genes, and Success: Scientists' Respond to The Bell Curve. New York: Springer-Verlag.

Chan N T, LeBaron B, Lo A-W, Poggio T (1999) Agent-based models of financial markets: A comparison with experimental markets. MIT Artificial Markets Project, Paper No. 124, September 5, 1999. Available via CiteSeer. http://citeseer.ist.psu.edu/chan99agentbased.html. Cited 18 June 2008.

Chen S.-H. Tai C.-C (2003) Trading restrictions, price dynamics and allocative efficiency in double auction markets: Analysis based on agent-based modeling and simulations. Advances in Complex Systems 6(3):283-302.

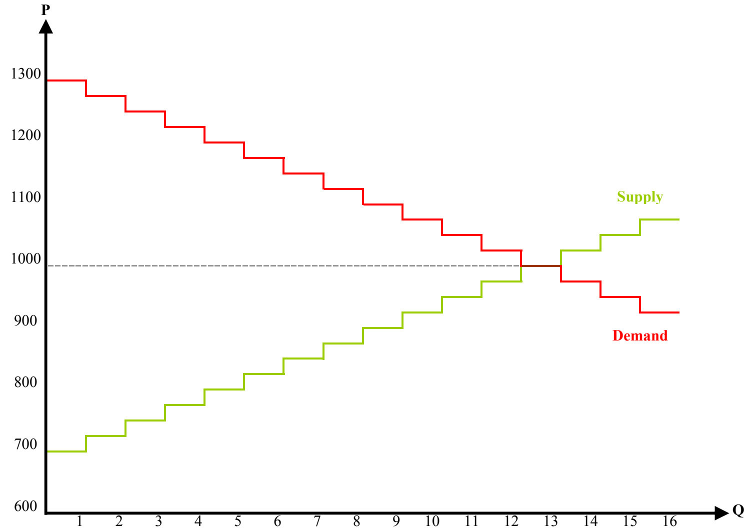

Figure 6: The Demand-Supply Schedule: M3

Chen S.-H, Zeng R.-J, Yu T (2008) Co-evolving trading strategies to analyze bounded rationality in double auction markets. In: Riolo R (ed.), Genetic programming theory and practice VI, Springer, 195–213.

Cliff D, Bruten J (1997) Zero is not enough: On the lower limit of agent intelligence for continuous double auction markets. Technical Report no. HPL-97-141, Hewlett-Packard Laboratories. (Available via CiteSeer.
http://citeseer.ist.psu.edu/cliff97zero.html. Cited 18 June 2008.

Cornelissen G, Dewitte S, Warlop L (2007). Social value orientation as a moral intuition: Decision-making in the dictator game. Economics Working Papers series. no. 1028, Department of Economics and Business, Universitat Pompeu Fabra, Spain.

Devetag G, Warglien M (2003) Games and phone numbers: Do short-term memory bounds affect strategic behavior? Journal of Economic Psychology 24:189-202.

Devetag G, Warglien M (2008) Playing the wrong game: An experimental analysis of relational complexity and strategic misrepresentation. Games and Economic Behavior 62:364-382.

Easley D, Ledyard J (1993) Theories of price formation and exchange in double oral auction. In Friedman D, Rust J (eds.), The Double Auction Market-Institutions, Theories, and Evidence. Redwood City, CA: Addison Wesley.

Friedman D (1991) A simple testable model of double auction markets. Journal of Economic Behavior and Organization 15:47–70.

Gjerstad S, Dickhaut J (1998) Price formation in double auctions. Games and Economic Behavior 22:1–29.

Gode D, Sunder S (1993) Allocative efficiency of markets with zero intelligence traders: Market as a partial substitute for individual rationality. Journal of Political Economy 101:119-137

Gode D, Sunder S (1994). Human and artificially intelligent traders in a double auction market: Experimental evidence. In: Carley K, Prietula M. (eds.), Computational Organization Theory. Lawrence Erlbaum Associates, Hillsdale, NJ, 241-262.

Gould E (2005) Inequality and ability. Labour Economics 12(2):169-189.

Herrnstein R. Murray C (1996) Bell curve: Intelligence and class structure in American life, Free Press.

Jensen A (1998) The g Factor: The science of mental ability, Praeger.

Jones G (2008) Are smarter groups more cooperative? Evidence from prisoners dilemma experiments, 1959-2003. Journal of Economic Behavior and Organization, forthcoming.

Jones G, Schneider W (2006) Intelligence, human capital, and economic growth: A Bayesian averaging of classical estimates (BACE) approach. Journal of Economic Growth 11(1):71-93.

Lynn R (2006) Race differences in intelligence: An evolutionary analysis. Washington Summit Publishers

Lynn R, Vanhanen T (2002) IQ and the Wealth of Nations, Praeger.

Miller G (1956) The magical number seven, plus or minus two: Some limits on our capacity for processing information. Psychological Review, 101(2):343-352.

Oberauer K, S*üβ* H.-M, Schulze R, Wilhelm O, Wittmann W (2000) Working memory capacity-facets of a cognitive ability construct. Personality and Individual Differences 29:1017-1045.

Oberauer K, S*üβ* H.-M, Wilhelm O, Wittmann W (2003) The multiple faces of working memory-storage, processing, supervision, and coordination. Intelligence 31:167-193.

Ohtsubo Y, Rapoport A (2006) Depth of reasoning in strategic form games. The Journal of Socio-Economics 35:31-47.

Ram R (2007) IQ and economic growth: Further augmentation of Mankiw-Romer-Weil model. Economics Letters 94(1):7-11.

Pashler H (1998). The Psychology of Attention. Cambridge, MA: MIT Press.

Rust J, Miller J, Palmer R (1993) Behavior of trading automata in a computerized double auction market, in Friedman D, Rust J (eds) Double AuctionMarkets: Theory, Institutions, and Laboratory Evidence, Redwood City, CA: Addison Wesley.

Rust J, Miller J, Palmer R (1994) Characterizing effective trading strategies: insights from a computerized double auction tournament. Journal of Economic Dynamics and Control 18:61–96.

Segal N, Hershberger S (1999) Cooperation and competition between twins: Findings from a prisoners dilemma game. Evolution and Human Behavior, 20:29V51.

Weede E, Kampf S (2002) The impact of intelligence and institutional improvements on economic growth. Kyklos 55(3):361-380.

Zax J, Rees D (2002) IQ, academic performance, environment, and earnings. Review of Economics and Statistics 84(4):600-616.

Zhan W, Friedman D (2007) Markups in double auction markets. Journal of Economic Dynamics and Control 31:2984–3005.